WEAK GRAVITATIONAL LENSING

MATTHIAS BARTELMANN[1] AND PETER SCHNEIDER[1,2]
[1] MAX-PLANCK-INSTITUT FÜR ASTROPHYSIK, P.O. BOX 1523, D–85740 GARCHING, GERMANY
[2] INSTITUT FÜR ASTROPHYSIK UND EXTRATERRESTRISCHE FORSCHUNG, UNIVERSITÄT BONN
AUF DEM HÜGEL 71, D–53121 BONN, GERMANY

to appear in Physics Reports

ABSTRACT

We review theory and applications of weak gravitational lensing. After summarising Friedmann-Lemaître cosmological models, we present the formalism of gravitational lensing and light propagation in arbitrary space-times. We discuss how weak-lensing effects can be measured. The formalism is then applied to reconstructions of galaxy-cluster mass distributions, gravitational lensing by large-scale matter distributions, QSO-galaxy correlations induced by weak lensing, lensing of galaxies by galaxies, and weak lensing of the cosmic microwave background.

1. INTRODUCTION

1.1. *Gravitational Light Deflection*

Light rays are deflected when they propagate through an inhomogeneous gravitational field. Although several researchers had speculated about such an effect well before the advent of General Relativity (see Schneider et al. 1992 for a historical account), it was Einstein's theory which elevated the deflection of light by masses from a hypothesis to a firm prediction. Assuming light behaves like a stream of particles, its deflection can be calculated within Newton's theory of gravitation, but General Relativity predicts that the effect is twice as large. A light ray grazing the surface of the Sun is deflected by 1.75 arc seconds compared to the 0.87 arc seconds predicted by Newton's theory. The confirmation of the larger value in 1919 was perhaps the most important step towards accepting General Relativity as the correct theory of gravity (Eddington 1920).

Cosmic bodies more distant, more massive, or more compact than the Sun can bend light rays from a single source sufficiently strongly so that multiple light rays can reach the observer. The observer sees an image in the direction of each ray arriving at their position, so that the source appears multiply imaged. In the language of General Relativity, there may exist more than one null geodesic connecting the world-line of a source with the observation event. Although predicted long before, the first multiple-image system was discovered only in 1979 (Walsh et al. 1979). From then on, the field of *gravitational lensing* developed into one of the most active subjects of astrophysical research. Several dozens of multiply-imaged sources have since been found. Their quantitative analysis provides accurate masses of, and in some cases detailed information on, the deflectors. An example is shown in Fig. 1.

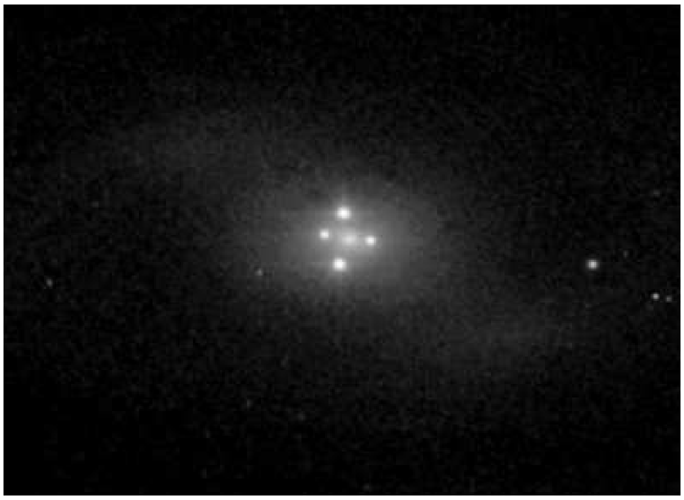

FIG. 1.—The gravitational lens system 2237+0305 consists of a nearby spiral galaxy at redshift $z_{\rm d} = 0.039$ and four images of a background quasar with redshift $z_{\rm s} = 1.69$. It was discovered by Huchra et al. (1985). The image was taken by the *Hubble Space Telescope* and shows only the innermost region of the lensing galaxy. The central compact source is the bright galaxy core, the other four compact sources are the quasar images. They differ in brightness because they are magnified by different amounts. The four images roughly fall on a circle concentric with the core of the lensing galaxy. The mass inside this circle can be determined with very high accuracy (Rix et al. 1992). The largest separation between the images is $1.8''$.

Tidal gravitational fields lead to differential deflection of light bundles. The size and shape of their cross sections are therefore changed. Since photons are neither emitted nor absorbed in the process of gravitational light deflection, the surface brightness of lensed sources remains unchanged. Changing the size of the cross section of a light bundle therefore changes the flux observed from a source. The different images in multiple-image systems generally have different fluxes. The images of extended sources, i.e. sources which can observationally be resolved, are deformed by the gravitational tidal field. Since astronomical sources like galaxies are not intrinsically circular, this deformation is generally very difficult to identify in individual images. In some cases, however, the distortion is strong enough to be readily recognised, most noticeably in the case of *Einstein rings* (see Fig. 2) and *arcs* in galaxy clusters (Fig. 3).

If the light bundles from some sources are distorted so strongly that their images appear as giant luminous arcs, one may expect many more sources behind a cluster whose images are only weakly distorted. Although weak distortions in individual images can hardly be recognised, the net distortion averaged over an ensemble of images can still be detected. As we shall describe in Sect. 2.3., deep optical exposures reveal a dense population of faint galaxies on the sky. Most of these galaxies are at high redshift, thus distant, and their image shapes can be utilised to probe the tidal gravitational field of intervening mass concentrations. Indeed, the tidal gravitational field can be reconstructed from the coherent distortion apparent in images of the faint galaxy population, and from that the density profile of intervening clusters of

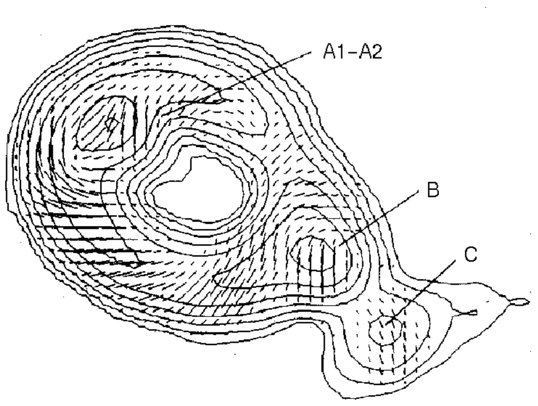

FIG. 2.—The radio source MG 1131+0456 was discovered by Hewitt et al. (1988) as the first example of a so-called *Einstein ring*. If a source and an axially symmetric lens are co-aligned with the observer, the symmetry of the system permits the formation of a ring-like image of the source centred on the lens. If the symmetry is broken (as expected for all realistic lensing matter distributions), the ring is deformed or broken up, typically into four images (see Fig. 1). However, if the source is sufficiently extended, ring-like images can be formed even if the symmetry is imperfect. The 6 cm radio map of MG 1131+0456 shows a closed ring, while the ring breaks up at higher frequencies where the source is smaller. The ring diameter is $2.1''$.

galaxies can be inferred (see Sect. 4.).

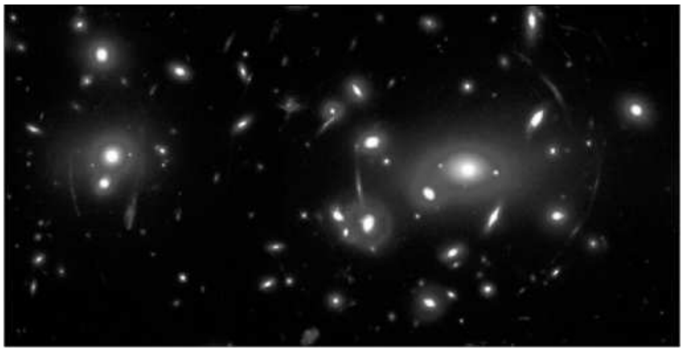

FIG. 3.—The cluster Abell 2218 hosts one of the most impressive collections of arcs. This *HST* image of the cluster's central region shows a pattern of strongly distorted galaxy images tangentially aligned with respect to the cluster centre, which lies close to the bright galaxy in the upper part of this image. The frame measures about $80'' \times 160''$. (courtesy of J.-P. Kneib)

1.2. Weak Gravitational Lensing

This review deals with *weak gravitational lensing*. There is no generally applicable definition of weak lensing despite the fact that it constitutes a flourishing area of research. The common aspect of all studies of weak gravitational lensing is that measurements of its effects are statistical in nature. While a single multiply-imaged source provides information on the mass distribution of the deflector, weak lensing effects show up only across ensembles of sources. One example was given above: The shape distribution of an ensemble of galaxy images is changed close to a massive galaxy cluster in the foreground, because the cluster's tidal field polarises the images. We shall see later that the size distribution of the background galaxy population is also locally changed in the neighbourhood of a massive intervening mass concentration.

Magnification and distortion effects due to weak lensing can be used to probe the statistical properties of the matter distribution between us and an ensemble of distant sources, provided some assumptions on the source properties can be made. For example, if a *standard candle*[1] at high redshift is identified, its flux can be used to estimate the magnification along its line-of-sight. It can be assumed that the orientation of faint distant galaxies is random. Then, any coherent alignment of images signals the presence of an intervening tidal gravitational field. As a third example, the positions on the sky of cosmic objects at vastly different distances from us should be mutually independent. A statistical association of foreground objects with background sources can therefore indicate the magnification caused by the foreground objects on the background sources.

All these effects are quite subtle, or weak, and many of the current challenges in the field are observational in nature. A coherent alignment of images of distant galaxies *can* be due to an intervening tidal gravitational field, but *could* also be due to propagation effects in the Earth's atmosphere or in the telescope. A variation in the number density of background sources around a foreground object *can* be due to a magnification effect, but *could* also be due to non-uniform photometry or obscuration effects. These potential systematic effects have to be controlled at a level well below the expected weak-lensing effects. We shall return to this essential point at various places in this review.

1.3. Applications of Gravitational Lensing

Gravitational lensing has developed into a versatile tool for observational cosmology. There are two main reasons:

1. The deflection angle of a light ray is determined by the gravitational field of the matter distribution along its path. According to Einstein's theory of General Relativity, the gravitational field is in turn determined by the stress-energy tensor of the matter distribution. For the astrophysically most relevant case of non-relativistic matter, the latter is characterised by the density distribution alone. Hence, the gravitational field, and thus the deflection angle, depend neither on the nature of the matter nor on its physical state. Light deflection probes the total matter density, without distinguishing between ordinary (baryonic) matter or dark matter. In contrast to other dynamical methods for probing gravitational fields, no assumption needs to be made on the dynamical state of the matter. For example, the interpretation of radial velocity measurements in terms of the gravitating mass requires the applicability of the virial theorem (i.e., the physical system is assumed to be in virial equilibrium), or knowledge of the orbits (such as the circular orbits in disk galaxies). However, as will be discussed in Sect. 3., lensing measures only the mass distribution projected along the line-of-sight, and is therefore insensitive to the extent of the mass distribution *along* the light rays, as long as this extent is small compared to the distances from the observer and the source to the deflecting mass. Keeping this in mind, mass determinations by lensing do not depend on any symmetry assumptions.

[1]The term *standard candle* is used for any class of astronomical objects whose intrinsic luminosity can be inferred independently of the observed flux. In the simplest case, all members of the class have the same luminosity. More typically, the luminosity depends on some other known and observable parameters, such that the luminosity can be inferred from them. The luminosity distance to any standard candle can directly be inferred from the square root of the ratio of source luminosity and observed flux. Since the luminosity distance depends on cosmological parameters, the geometry of the Universe can then directly be investigated. Probably the best current candidates for standard candles are supernovae of Type Ia. They can be observed to quite high redshifts, and thus be utilised to estimate cosmological parameters (e.g. Riess et al. 1998).

2. Once the deflection angle as a function of impact parameter is given, gravitational lensing reduces to simple geometry. Since most lens systems involve sources (and lenses) at moderate or high redshift, lensing can probe the geometry of the Universe. This was noted by Refsdal (1964), who pointed out that lensing can be used to determine the Hubble constant and the cosmic density parameter. Although this turned out later to be more difficult than anticipated at the time, first measurements of the Hubble constant through lensing have been obtained with detailed models of the matter distribution in multiple-image lens systems and the difference in light-travel time along the different light paths corresponding to different images of the source (e.g., Kundić et al. 1997; Schechter et al. 1997; Biggs et al. 1999). Since the volume element per unit redshift interval and unit solid angle also depends on the geometry of space-time, so does the number of lenses therein. Hence, the lensing probability for distant sources depends on the cosmological parameters (e.g., Press & Gunn 1973). Unfortunately, in order to derive constraints on the cosmological model with this method, one needs to know the evolution of the lens population with redshift. Nevertheless, in some cases, significant constraints on the cosmological parameters (Kochanek 1993, 1996; Maoz & Rix 1993; Bartelmann et al. 1998; Falco et al. 1998), and on the evolution of the lens population (Mao & Kochanek 1994) have been derived from multiple-image and arc statistics. (See also the review by Chiba & Futamase 1999.)

The possibility to directly investigate the dark-matter distribution led to substantial results over recent years. Constraints on the size of the dark-matter halos of spiral galaxies were derived (e.g., Brainerd et al. 1996), the presence of dark-matter halos in elliptical galaxies was demonstrated (e.g., Maoz & Rix 1993; Griffiths et al. 1996), and the projected total mass distribution in many cluster of galaxies was mapped (e.g., Kneib et al. 1996; Hoekstra et al. 1998; Kaiser et al. 1998). These results directly impact on our understanding of structure formation, supporting hierarchical structure formation in cold dark matter (CDM) models. Constraints on the nature of dark matter were also obtained. Compact dark-matter objects, such as black holes or brown dwarfs, cannot be very abundant in the Universe, because otherwise they would lead to observable lensing effects (e.g., Schneider 1993; Dalcanton et al. 1994). Galactic microlensing experiments constrained the density and typical mass scale of massive compact halo objects in our Galaxy (see Paczyński 1996, Roulet & Mollerach 1997 and Mao 2000 for reviews). We refer the reader to the reviews by Blandford & Narayan (1992), Schneider (1996a) and Narayan & Bartelmann (1999) for a detailed account of the cosmological applications of gravitational lensing.

We shall concentrate almost entirely on weak gravitational lensing here. Hence, the flourishing fields of multiple-image systems and their interpretation, Galactic microlensing and its consequences for understanding the nature of dark matter in the halo of our Galaxy, and the detailed investigations of the mass distribution in the inner parts of galaxy clusters through arcs, arclets, and multiply imaged background galaxies, will not be covered in this review. In addition to the references given above, we would like to point the reader to Refsdal & Surdej (1994), Fort & Mellier (1994), Wu (1996), and Hattori et al. (1999) for more recent reviews on various aspects of gravitational lensing, to Mellier (1999) for a very recent review on weak lensing, and to the monograph (Schneider et al. 1992) for a detailed account of the theory and applications of gravitational lensing.

1.4. Structure of this Review

Many aspects of weak gravitational lensing are intimately related to the cosmological model and to the theory of structure formation in the Universe. We therefore start the review by giving some cosmological background in Sect. 2.. After summarising Friedmann-Lemaître-Robertson-Walker models, we sketch the theory of structure formation, introduce astrophysical objects like QSOs, galaxies, and galaxy clusters, and finish the Section with a general discussion of correlation functions, power spectra, and their projections. Gravitational light deflection in general is the subject of Sect. 3., and the specialisation to weak lensing is described in Sect. 4.. One of the main aspects there is how weak lensing effects can be quantified and measured. The following two sections describe the theory of weak lensing by galaxy clusters (Sect. 5.) and cosmological mass distributions (Sect. 6.). Apparent correlations between background QSOs and foreground galaxies due to the magnification bias caused by large-scale matter distributions are the subject of Sect. 7.. Weak lensing effects of foreground galaxies on background galaxies are reviewed in Sect. 8., and Sect. 9. finally deals with weak lensing of the most distant and most extended source possible, i.e. the Cosmic Microwave Background. We present a brief summary and an outlook in Sect. 10..

We use standard astronomical units throughout: $1 M_\odot = 1\,\mathrm{solar\ mass} = 2 \times 10^{33}\,\mathrm{g}$; $1\,\mathrm{Mpc} = 1\,\mathrm{megaparsec} = 3.1 \times 10^{24}\,\mathrm{cm}$.

2. COSMOLOGICAL BACKGROUND

We review in this section those aspects of the standard cosmological model which are relevant for our further discussion of weak gravitational lensing. This standard model consists of a description for the cosmological background which is a homogeneous and isotropic solution of the field equations of General Relativity, and a theory for the formation of structure.

The background model is described by the Robertson-Walker metric (Robertson 1935; Walker 1935), in which hypersurfaces of constant time are homogeneous and isotropic three-spaces, either flat or curved, and change with time according to a scale factor which depends on time only. The dynamics of the scale factor is determined by two equations which follow from Einstein's field equations given the highly symmetric form of the metric.

Current theories of structure formation assume that structure grows via gravitational instability from initial seed perturbations whose origin is yet unclear. Most common hypotheses lead to the prediction that the statistics of the seed fluctuations is Gaussian. Their amplitude is low for most of their evolution so that linear perturbation theory is sufficient to describe their growth until late stages. For general references on the cosmological model and on the theory of structure formation, cf. Weinberg (1972), Misner et al. (1973), Peebles (1980), Börner (1988), Padmanabhan (1993), Peebles (1993), and Peacock (1999).

2.1. Friedmann-Lemaître Cosmological Models

2.1.1. Metric

Two postulates are fundamental to the standard cosmological model, which are:

1. *When averaged over sufficiently large scales, there exists a mean motion of radiation and matter in the Universe with respect to which all averaged observable properties are isotropic.*

2. *All fundamental observers, i.e. imagined observers which follow this mean motion, experience the same history of the Universe, i.e. the same averaged observable properties, provided they set their clocks suitably.* Such a universe is called *observer-homogeneous.*

General Relativity describes space-time as a four-dimensional manifold whose metric tensor $g_{\alpha\beta}$ is considered as a dynamical field. The dynamics of the metric is governed by Einstein's field equations, which relate the Einstein tensor to the stress-energy tensor of the matter contained in space-time. Two events in space-time with coordinates differing by $\mathrm{d}x^\alpha$ are separated by $\mathrm{d}s$, with $\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$. The *eigentime* (proper time) of an observer who travels by $\mathrm{d}s$ changes by $c^{-1}\mathrm{d}s$. Greek indices run over $0\ldots3$ and Latin indices run over the spatial indices $1\ldots3$ only.

The two postulates stated above considerably constrain the admissible form of the metric tensor. Spatial coordinates which are constant for fundamental observers are called comoving coordinates. In these coordinates, the mean motion is described by $\mathrm{d}x^i = 0$, and hence $\mathrm{d}s^2 = g_{00}\mathrm{d}t^2$. If we require that the *eigentime* of fundamental observers equal the cosmic time, this implies $g_{00} = c^2$.

Isotropy requires that clocks can be synchronised such that the space-time components of the metric tensor vanish, $g_{0i} = 0$. If this was impossible, the components of g_{0i} identified one particular direction in space-time, violating isotropy. The metric can therefore be written

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 + g_{ij}\mathrm{d}x^i\mathrm{d}x^j \,, \tag{2.1}$$

where g_{ij} is the metric of spatial hypersurfaces. In order not to violate isotropy, the spatial metric can only isotropically contract or expand with a scale function $a(t)$ which must be a function of time only, because otherwise the expansion would be different at different places, violating homogeneity. Hence the metric further simplifies to

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\mathrm{d}l^2 \,, \tag{2.2}$$

where $\mathrm{d}l$ is the line element of the homogeneous and isotropic three-space. A special case of the metric (2.2) is the Minkowski metric, for which $\mathrm{d}l$ is the Euclidean line element and $a(t)$ is a constant. Homogeneity also implies that all quantities describing the matter content of the Universe, e.g. density and pressure, can be functions of time only.

The spatial hypersurfaces whose geometry is described by $\mathrm{d}l^2$ can either be flat or curved. Isotropy only requires them to be spherically symmetric, i.e. spatial surfaces of constant distance from an arbitrary point need to be two-spheres. Homogeneity permits us to choose an arbitrary point as coordinate origin. We can then introduce two angles θ, ϕ which uniquely identify positions on the unit sphere around the origin, and a radial coordinate w. The most general admissible form for the spatial line element is then

$$\mathrm{d}l^2 = \mathrm{d}w^2 + f_K^2(w)\left(\mathrm{d}\phi^2 + \sin^2\theta\mathrm{d}\theta^2\right) \equiv \mathrm{d}w^2 + f_K^2(w)\mathrm{d}\omega^2 \,. \tag{2.3}$$

Homogeneity requires that the radial function $f_K(w)$ is either a trigonometric, linear, or hyperbolic function of w, depending on whether the curvature K is positive, zero, or negative. Specifically,

$$f_K(w) = \begin{cases} K^{-1/2}\sin(K^{1/2}w) & (K>0) \\ w & (K=0) \\ (-K)^{-1/2}\sinh[(-K)^{1/2}w] & (K<0) \end{cases} \,. \tag{2.4}$$

Note that $f_K(w)$ and thus $|K|^{-1/2}$ have the dimension of a length. If we define the radius r of the two-spheres by $f_K(w) \equiv r$, the metric $\mathrm{d}l^2$ takes the alternative form

$$\mathrm{d}l^2 = \frac{\mathrm{d}r^2}{1-Kr^2} + r^2\mathrm{d}\omega^2 \,. \tag{2.5}$$

2.1.2. Redshift

Due to the expansion of space, photons are redshifted while they propagate from the source to the observer. Consider a comoving source emitting a light signal at t_e which reaches a comoving observer at the coordinate origin $w = 0$ at time t_o. Since $\mathrm{d}s = 0$ for light, a backward-directed radial light ray propagates according to $|c\mathrm{d}t| = a\mathrm{d}w$, from the metric. The (comoving) coordinate distance between source and observer is constant by definition,

$$w_\mathrm{eo} = \int_\mathrm{o}^\mathrm{e}\mathrm{d}w = \int_{t_\mathrm{e}}^{t_\mathrm{o}(t_\mathrm{e})}\frac{c\mathrm{d}t}{a} = \text{constant} \,, \tag{2.6}$$

and thus in particular the derivative of w_eo with respect to t_e is zero. It then follows from eq. (2.6)

$$\frac{\mathrm{d}t_\mathrm{o}}{\mathrm{d}t_\mathrm{e}} = \frac{a(t_\mathrm{o})}{a(t_\mathrm{e})} \,. \tag{2.7}$$

Identifying the inverse time intervals $(\mathrm{d}t_\mathrm{e,o})^{-1}$ with the emitted and observed light frequencies $\nu_\mathrm{e,o}$, we can write

$$\frac{\mathrm{d}t_\mathrm{o}}{\mathrm{d}t_\mathrm{e}} = \frac{\nu_\mathrm{e}}{\nu_\mathrm{o}} = \frac{\lambda_\mathrm{o}}{\lambda_\mathrm{e}} \,. \tag{2.8}$$

Since the redshift z is defined as the relative change in wavelength, or $1+z = \lambda_\mathrm{o}\lambda_\mathrm{e}^{-1}$, we find

$$1+z = \frac{a(t_\mathrm{o})}{a(t_\mathrm{e})} \,. \tag{2.9}$$

This shows that light is redshifted by the amount by which the Universe has expanded between emission and observation.

2.1.3. Expansion

To complete the description of space-time, we need to know how the scale function $a(t)$ depends on time, and how the curvature K depends on the matter which fills space-time. That is, we ask for the dynamics of the space-time. Einstein's field equations relate the Einstein tensor $G_{\alpha\beta}$ to the stress-energy tensor $T_{\alpha\beta}$ of the matter,

$$G_{\alpha\beta} = \frac{8\pi G}{c^2}T_{\alpha\beta} + \Lambda g_{\alpha\beta} \,. \tag{2.10}$$

The second term proportional to the metric tensor $g_{\alpha\beta}$ is a generalisation introduced by Einstein to allow static cosmological solutions of the field equations. Λ is called the cosmological constant. For the highly symmetric form of the metric given by (2.2) and (2.3), Einstein's equations imply that $T_{\alpha\beta}$ has to have the form of the stress-energy tensor of a homogeneous perfect fluid, which is characterised by its density $\rho(t)$ and its pressure $p(t)$. Matter density and pressure can only depend on time because of homogeneity. The field equations then simplify to the two independent equations

$$\left(\frac{\dot a}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3} \tag{2.11}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}\ . \tag{2.12}$$

The scale factor $a(t)$ is determined once its value at one instant of time is fixed. We choose $a = 1$ at the present epoch t_0. Equation (2.11) is called *Friedmann's equation* (Friedmann 1922, 1924). The two equations (2.11) and (2.12) can be combined to yield the *adiabatic equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[a^3(t)\rho(t)c^2\right] + p(t)\frac{\mathrm{d}a^3(t)}{\mathrm{d}t} = 0\ , \tag{2.13}$$

which has an intuitive interpretation. The first term $a^3\rho$ is proportional to the energy contained in a fixed comoving volume, and hence the equation states that the change in 'internal' energy equals the pressure times the change in proper volume. Hence eq. (2.13) is the first law of thermodynamics in the cosmological context.

A metric of the form given by eqs. (2.2), (2.3), and (2.4) is called the Robertson-Walker metric. If its scale factor $a(t)$ obeys Friedmann's equation (2.11) and the adiabatic equation (2.13), it is called the Friedmann-Lemaître-Robertson-Walker metric, or the Friedmann-Lemaître metric for short. Note that eq. (2.12) can also be derived from Newtonian gravity except for the pressure term in (2.12) and the cosmological constant. Unlike in Newtonian theory, pressure acts as a source of gravity in General Relativity.

2.1.4. Parameters

The relative expansion rate $\dot{a}a^{-1} \equiv H$ is called the *Hubble parameter*, and its value at the present epoch $t = t_0$ is the *Hubble constant*, $H(t_0) \equiv H_0$. It has the dimension of an inverse time. The value of H_0 is still uncertain. Current measurements roughly fall into the range $H_0 = (50-80)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (see Freedman 1996 for a review), and the uncertainty in H_0 is commonly expressed as $H_0 = 100\,h\,\mathrm{km\ s^{-1}\ Mpc^{-1}}$, with $h = (0.5-0.8)$. Hence

$$H_0 \approx 3.2\times10^{-18}\,h\,\mathrm{s^{-1}} \approx 1.0\times10^{-10}\,h\,\mathrm{yr^{-1}}\ . \tag{2.14}$$

The time scale for the expansion of the Universe is the inverse Hubble constant, or $H_0^{-1} \approx 10^{10}\,h^{-1}$ years.

The combination

$$\frac{3H_0^2}{8\pi G} \equiv \rho_{\rm cr} \approx 1.9\times10^{-29}\,h^2\,\mathrm{g\,cm^{-3}} \tag{2.15}$$

is the *critical density* of the Universe, and the density ρ_0 in units of $\rho_{\rm cr}$ is the *density parameter* Ω_0,

$$\Omega_0 = \frac{\rho_0}{\rho_{\rm cr}}\ . \tag{2.16}$$

If the matter density in the universe is critical, $\rho_0 = \rho_{\rm cr}$ or $\Omega_0 = 1$, and if the cosmological constant vanishes, $\Lambda = 0$, spatial hypersurfaces are flat, $K = 0$, which follows from (2.11) and will become explicit in eq. (2.30) below. We further define

$$\Omega_\Lambda \equiv \frac{\Lambda}{3H_0^2}\ . \tag{2.17}$$

The *deceleration parameter* q_0 is defined by

$$q_0 = -\frac{\ddot{a}a}{\dot{a}^2} \tag{2.18}$$

at $t = t_0$.

2.1.5. Matter Models

For a complete description of the expansion of the Universe, we need an equation of state $p = p(\rho)$, relating the pressure to the energy density of the matter. Ordinary matter, which is frequently called *dust* in this context, has $p \ll \rho c^2$, while $p = \rho c^2/3$ for radiation or other forms of relativistic matter. Inserting these expressions into eq. (2.13), we find

$$\rho(t) = a^{-n}(t)\,\rho_0\ , \tag{2.19}$$

with

$$n = \begin{cases} 3 & \text{for dust, } p = 0 \\ 4 & \text{for relativistic matter, } p = \rho c^2/3 \end{cases}\ . \tag{2.20}$$

The energy density of relativistic matter therefore drops more rapidly with time than that of ordinary matter.

2.1.6. Relativistic Matter Components

There are two obvious candidates for relativistic matter today, photons and neutrinos. The energy density contained in photons today is determined by the temperature of the Cosmic Microwave Background, $T_{\rm CMB} = 2.73\,\mathrm{K}$ (Fixsen et al. 1996). Since the CMB has an excellent black-body spectrum, its energy density is given by the Stefan-Boltzmann law,

$$\rho_{\rm CMB} = \frac{1}{c^2}\frac{\pi^2}{15}\frac{(kT_{\rm CMB})^4}{(\hbar c)^3} \approx 4.5\times10^{-34}\,\mathrm{g\,cm^{-3}}\ . \tag{2.21}$$

In terms of the cosmic density parameter Ω_0 [eq. (2.16)], the cosmic density contributed by the photon background is

$$\Omega_{\rm CMB,0} = 2.4\times10^{-5}\,h^{-2}\ . \tag{2.22}$$

Like photons, neutrinos were produced in thermal equilibrium in the hot early phase of the Universe. Interacting weakly, they decoupled from the cosmic plasma when the temperature of the Universe was $kT \approx 1\,\mathrm{MeV}$ because later the time-scale of their leptonic interactions became larger than the expansion time-scale of the Universe, so that equilibrium could no longer be maintained. When the temperature of the Universe dropped to $kT \approx 0.5\,\mathrm{MeV}$, electron-positron pairs annihilated to produce γ rays. The annihilation heated up the photons but not the neutrinos which had decoupled earlier. Hence the neutrino temperature is lower than the photon temperature by an amount determined by entropy conservation. The entropy $S_{\rm e}$ of the electron-positron pairs was dumped completely into the entropy of the photon background S_γ. Hence,

$$(S_{\rm e} + S_\gamma)_{\rm before} = (S_\gamma)_{\rm after}\ , \tag{2.23}$$

where "before" and "after" refer to the annihilation time. Ignoring constant factors, the entropy per particle species is $S \propto gT^3$, where g is the statistical weight of the species. For bosons $g = 1$, and for fermions $g = 7/8$ per spin state. Before annihilation, we thus have $g_{\rm before} = 4\cdot7/8+2 = 11/2$, while after the annihilation $g = 2$ because only photons remain. From eq. (2.23),

$$\left(\frac{T_{\rm after}}{T_{\rm before}}\right)^3 = \frac{11}{4}\ . \tag{2.24}$$

After the annihilation, the neutrino temperature is therefore lower than the photon temperature by the factor $(11/4)^{1/3}$. In particular, the neutrino temperature today is

$$T_{\nu,0} = \left(\frac{4}{11}\right)^{1/3} T_{\rm CMB} = 1.95\,\mathrm{K}\ . \tag{2.25}$$

Although neutrinos have long been out of thermal equilibrium, their distribution function remained unchanged since they decoupled, except that their temperature gradually dropped in the course of cosmic expansion. Their energy density can thus be computed from a Fermi-Dirac distribution with temperature T_ν, and be converted to the equivalent cosmic density parameter as for the photons. The result is

$$\Omega_{\nu,0} = 2.8 \times 10^{-6} h^{-2} \tag{2.26}$$

per neutrino species.

Assuming three relativistic neutrino species, the total density parameter in relativistic matter today is

$$\Omega_{R,0} = \Omega_{CMB,0} + 3 \times \Omega_{\nu,0} = 3.2 \times 10^{-5} h^{-2} \,. \tag{2.27}$$

Since the energy density in relativistic matter is almost five orders of magnitude less than the energy density of ordinary matter today if Ω_0 is of order unity, the expansion of the Universe today is matter-dominated, or $\rho = a^{-3}(t)\rho_0$. The energy densities of ordinary and relativistic matter were equal when the scale factor $a(t)$ was

$$a_{eq} = \frac{\Omega_{R,0}}{\Omega_0} = 3.2 \times 10^{-5} \Omega_0^{-1} h^{-2} \,, \tag{2.28}$$

and the expansion was radiation-dominated at yet earlier times, $\rho = a^{-4}\rho_0$. The epoch of equality of matter and radiation density will turn out to be important for the evolution of structure in the Universe discussed below.

2.1.7. Spatial Curvature and Expansion

With the parameters defined previously, Friedmann's equation (2.11) can be written

$$H^2(t) = H_0^2 \left[a^{-4}(t)\Omega_{R,0} + a^{-3}(t)\Omega_0 - a^{-2}(t)\frac{Kc^2}{H_0^2} + \Omega_\Lambda \right] \,. \tag{2.29}$$

Since $H(t_0) \equiv H_0$, and $\Omega_{R,0} \ll \Omega_0$, eq. (2.29) implies

$$K = \left(\frac{H_0}{c}\right)^2 (\Omega_0 + \Omega_\Lambda - 1) \,, \tag{2.30}$$

and eq. (2.29) becomes

$$H^2(t) = H_0^2 \left[\frac{\Omega_{R,0}}{a^4(t)} + \frac{\Omega_0}{a^3(t)} + \frac{1 - \Omega_0 - \Omega_\Lambda}{a^2(t)} + \Omega_\Lambda \right] \,. \tag{2.31}$$

The curvature of spatial hypersurfaces is therefore determined by the sum of the density contributions from matter, Ω_0, and from the cosmological constant, Ω_Λ. If $\Omega_0 + \Omega_\Lambda = 1$, space is flat, and it is closed or hyperbolic if $\Omega_0 + \Omega_\Lambda$ is larger or smaller than unity, respectively. The spatial hypersurfaces of a low-density universe are therefore hyperbolic, while those of a high-density universe are closed [cf. eq. (2.4)]. A Friedmann-Lemaître model universe is thus characterised by four parameters: the expansion rate at present (or Hubble constant) H_0, and the density parameters in matter, radiation, and the cosmological constant.

Dividing eq. (2.12) by eq. (2.11), using eq. (2.30), and setting $p = 0$, we obtain for the deceleration parameter q_0

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda \,. \tag{2.32}$$

The age of the universe can be determined from eq. (2.31). Since $dt = da\,\dot{a}^{-1} = da(aH)^{-1}$, we have, ignoring $\Omega_{R,0}$,

$$t_0 = \frac{1}{H_0} \int_0^1 da \left[a^{-1}\Omega_0 + (1 - \Omega_0 - \Omega_\Lambda) + a^2\Omega_\Lambda \right]^{-1/2} \,. \tag{2.33}$$

It was assumed in this equation that $p = 0$ holds for all times t, while pressure is not negligible at early times. The corresponding error, however, is very small because the universe spends only a very short time in the radiation-dominated phase where $p > 0$.

Figure 4 shows t_0 in units of H_0^{-1} as a function of Ω_0, for $\Omega_\Lambda = 0$ (solid curve) and $\Omega_\Lambda = 1 - \Omega_0$ (dashed curve). The model universe is older for lower Ω_0 and higher Ω_Λ because the deceleration decreases with decreasing Ω_0 and the acceleration increases with increasing Ω_Λ.

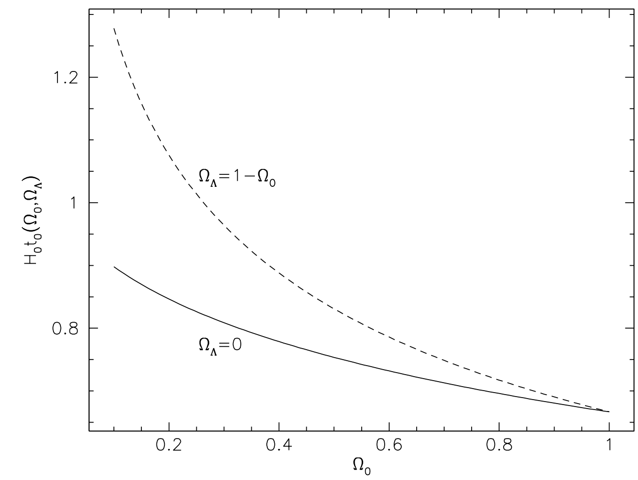

FIG. 4.—Cosmic age t_0 in units of H_0^{-1} as a function of Ω_0, for $\Omega_\Lambda = 0$ (solid curve) and $\Omega_\Lambda = 1 - \Omega_0$ (dashed curve).

In principle, Ω_Λ can have either sign. We have restricted ourselves in Fig. 4 to non-negative Ω_Λ because the cosmological constant is usually interpreted as the energy density of the vacuum, which is positive semi-definite.

The time evolution (2.31) of the Hubble function $H(t)$ allows one to determine the dependence of Ω and Ω_Λ on the scale function a. For a matter-dominated universe, we find

$$\begin{aligned} \Omega(a) &= \frac{8\pi G}{3H^2(a)} \rho_0 a^{-3} = \frac{\Omega_0}{a + \Omega_0(1-a) + \Omega_\Lambda(a^3 - a)} \,, \\ \Omega_\Lambda(a) &= \frac{\Lambda}{3H^2(a)} = \frac{\Omega_\Lambda a^3}{a + \Omega_0(1-a) + \Omega_\Lambda(a^3 - a)} \,. \end{aligned} \tag{2.34}$$

These equations show that, whatever the values of Ω_0 and Ω_Λ are at the present epoch, $\Omega(a) \to 1$ and $\Omega_\Lambda \to 0$ for $a \to 0$. This implies that for sufficiently early times, all matter-dominated Friedmann-Lemaître model universes can be described by Einstein-de Sitter models, for which $K = 0$ and $\Omega_\Lambda = 0$. For $a \ll 1$, the right-hand side of Friedmann's equation (2.31) is therefore dominated by the matter and radiation terms because they contain the strongest dependences on a^{-1}. The Hubble function $H(t)$ can then be approximated by

$$H(t) = H_0 \left[\Omega_{R,0} a^{-4}(t) + \Omega_0 a^{-3}(t) \right]^{1/2} \,. \tag{2.35}$$

Using the definition of a_{eq}, $a_{eq}^{-4}\Omega_{R,0} = a_{eq}^{-3}\Omega_0$ [cf. eq. (2.28)], eq. (2.35) can be written

$$H(t) = H_0\,\Omega_0^{1/2}\,a^{-3/2}\left(1+\frac{a_{\rm eq}}{a}\right)^{1/2}\;. \tag{2.36}$$

Hence,

$$H(t) = H_0\,\Omega_0^{1/2}\begin{cases} a_{\rm eq}^{1/2}\,a^{-2} & (a \ll a_{\rm eq}) \\ a^{-3/2} & (a_{\rm eq} \ll a \ll 1) \end{cases}\;. \tag{2.37}$$

Likewise, the expression for the cosmic time reduces to

$$t(a) = \frac{2}{3H_0}\Omega_0^{-1/2}\left[a^{3/2}\left(1-2\frac{a_{\rm eq}}{a}\right)\left(1+\frac{a_{\rm eq}}{a}\right)^{1/2}+2a_{\rm eq}^{3/2}\right]\;, \tag{2.38}$$

or

$$t(a) = \frac{1}{H_0}\Omega_0^{-1/2}\begin{cases} \frac{1}{2}\,a_{\rm eq}^{-1/2}\,a^2 & (a \ll a_{\rm eq}) \\ \frac{2}{3}\,a^{3/2} & (a_{\rm eq} \ll a \ll 1) \end{cases}\;. \tag{2.39}$$

Equation (2.36) is called the Einstein-de Sitter limit of Friedmann's equation. Where not mentioned otherwise, we consider in the following only cosmic epochs at times much later than $t_{\rm eq}$, i.e., when $a \gg a_{\rm eq}$, where the Universe is dominated by dust, so that the pressure can be neglected, $p = 0$.

2.1.8. *Necessity of a Big Bang*

Starting from $a = 1$ at the present epoch and integrating Friedmann's equation (2.11) back in time shows that there are combinations of the cosmic parameters such that $a > 0$ at all times. Such models would have no Big Bang. The necessity of a Big Bang is usually inferred from the existence of the cosmic microwave background, which is most naturally explained by an early, hot phase of the Universe. Börner & Ehlers (1988) showed that two simple observational facts suffice to show that the Universe must have gone through a Big Bang, if it is properly described by the class of Friedmann-Lemaître models. Indeed, the facts that there are cosmological objects at redshifts $z > 4$, and that the cosmic density parameter of non-relativistic matter, as inferred from observed galaxies and clusters of galaxies is $\Omega_0 > 0.02$, exclude models which have $a(t) > 0$ at all times. Therefore, if we describe the Universe at large by Friedmann-Lemaître models, we must assume a Big Bang, or $a = 0$ at some time in the past.

2.1.9. *Distances*

The meaning of "distance" is no longer unique in a curved space-time. In contrast to the situation in Euclidean space, distance definitions in terms of different measurement prescriptions lead to different distances. Distance measures are therefore defined in analogy to relations between measurable quantities in Euclidean space. We define here four different distance scales, the proper distance, the comoving distance, the angular-diameter distance, and the luminosity distance.

Distance measures relate an emission event and an observation event on two separate geodesic lines which fall on a common light cone, either the forward light cone of the source or the backward light cone of the observer. They are therefore characterised by the times t_2 and t_1 of emission and observation respectively, and by the structure of the light cone. These times can uniquely be expressed by the values $a_2 = a(t_2)$ and $a_1 = a(t_1)$ of the scale factor, or by the redshifts z_2 and z_1 corresponding to a_2 and a_1. We choose the latter parameterisation because redshifts are directly observable. We also assume that the observer is at the origin of the coordinate system.

The *proper distance* $D_{\rm prop}(z_1,z_2)$ is the distance measured by the travel time of a light ray which propagates from a source at z_2 to an observer at $z_1 < z_2$. It is defined by $\mathrm{d}D_{\rm prop} = -c\mathrm{d}t$, hence $\mathrm{d}D_{\rm prop} = -c\mathrm{d}a\dot{a}^{-1} = -c\mathrm{d}a(aH)^{-1}$. The minus sign arises because, due to the choice of coordinates centred on the observer, distances increase away from the observer, while the time t and the scale factor a increase towards the observer. We get

$$\begin{aligned} D_{\rm prop}(z_1,z_2) &= \frac{c}{H_0}\int_{a(z_2)}^{a(z_1)}\Big[a^{-1}\Omega_0 \\ &+ (1-\Omega_0-\Omega_\Lambda)+a^2\Omega_\Lambda\Big]^{-1/2}\mathrm{d}a\;. \end{aligned} \tag{2.40}$$

The *comoving distance* $D_{\rm com}(z_1,z_2)$ is the distance on the spatial hyper-surface $t = t_0$ between the world-lines of a source and an observer comoving with the cosmic flow. Due to the choice of coordinates, it is the coordinate distance between a source at z_2 and an observer at z_1, $\mathrm{d}D_{\rm com} = \mathrm{d}w$. Since light rays propagate with $\mathrm{d}s = 0$, we have $c\mathrm{d}t = -a\mathrm{d}w$ from the metric, and therefore $\mathrm{d}D_{\rm com} = -a^{-1}c\mathrm{d}t = -c\mathrm{d}a(a\dot{a})^{-1} = -c\mathrm{d}a(a^2H)^{-1}$. Thus

$$\begin{aligned} D_{\rm com}(z_1,z_2) &= \frac{c}{H_0}\int_{a(z_2)}^{a(z_1)}\Big[a\Omega_0 \\ &+ a^2(1-\Omega_0-\Omega_\Lambda)+a^4\Omega_\Lambda\Big]^{-1/2}\mathrm{d}a \\ &= w(z_1,z_2)\;. \end{aligned} \tag{2.41}$$

The *angular-diameter distance* $D_{\rm ang}(z_1,z_2)$ is defined in analogy to the relation in Euclidean space between the physical cross section δA of an object at z_2 and the solid angle $\delta\omega$ that it subtends for an observer at z_1, $\delta\omega D_{\rm ang}^2 = \delta A$. Hence,

$$\frac{\delta A}{4\pi a^2(z_2)\,f_K^2[w(z_1,z_2)]} = \frac{\delta\omega}{4\pi}\;, \tag{2.42}$$

where $a(z_2)$ is the scale factor at emission time and $f_K[w(z_1,z_2)]$ is the radial coordinate distance between the observer and the source. It follows

$$D_{\rm ang}(z_1,z_2) = \left(\frac{\delta A}{\delta\omega}\right)^{1/2} = a(z_2)\,f_K[w(z_1,z_2)]\;. \tag{2.43}$$

According to the definition of the comoving distance, the angular-diameter distance therefore is

$$D_{\rm ang}(z_1,z_2) = a(z_2)\,f_K[D_{\rm com}(z_1,z_2)]\;. \tag{2.44}$$

The *luminosity distance* $D_{\rm lum}(a_1,a_2)$ is defined by the relation in Euclidean space between the luminosity L of an object at z_2 and the flux S received by an observer at z_1. It is related to the angular-diameter distance through

$$D_{\rm lum}(z_1,z_2) = \left[\frac{a(z_1)}{a(z_2)}\right]^2 D_{\rm ang}(z_1,z_2) = \frac{a(z_1)^2}{a(z_2)}\,f_K[D_{\rm com}(z_1,z_2)]\;. \tag{2.45}$$

The first equality in (2.45), which is due to Etherington (1933), is valid in arbitrary space-times. It is physically intuitive because photons are redshifted by $a(z_1)a(z_2)^{-1}$, their arrival times are delayed by another factor $a(z_1)a(z_2)^{-1}$, and the area of the observer's sphere on which the photons are distributed grows between emission and absorption in proportion to $[a(z_1)a(z_2)^{-1}]^2$.

This accounts for a total factor of $[a(z_1)a(z_2)^{-1}]^4$ in the flux, and hence for a factor of $[a(z_1)a(z_2)^{-1}]^2$ in the distance relative to the angular-diameter distance.

We plot the four distances $D_{\rm prop}$, $D_{\rm com}$, $D_{\rm ang}$, and $D_{\rm lum}$ for $z_1 = 0$ as a function of z in Fig. 5.

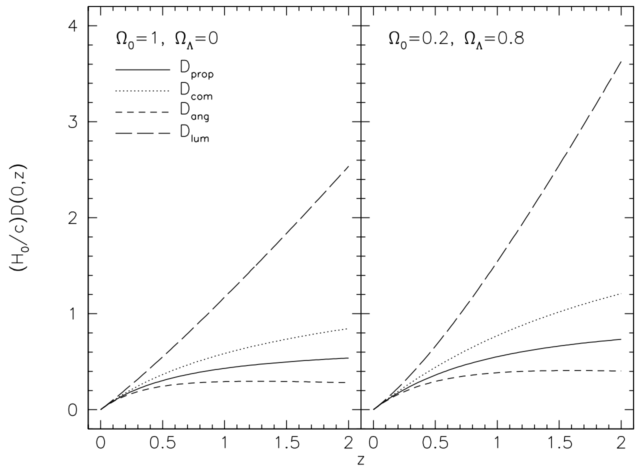

FIG. 5.—Four distance measures are plotted as a function of source redshift for two cosmological models and an observer at redshift zero. These are the proper distance $D_{\rm prop}$ (1, solid line), the comoving distance $D_{\rm com}$ (2, dotted line), the angular-diameter distance $D_{\rm ang}$ (3, short-dashed line), and the luminosity distance $D_{\rm lum}$ (4, long-dashed line).

The distances are larger for lower cosmic density and higher cosmological constant. Evidently, they differ by a large amount at high redshift. For small redshifts, $z \ll 1$, they all follow the Hubble law,

$$\text{distance} = \frac{cz}{H_0} + O(z^2)\;. \tag{2.46}$$

2.1.10. The Einstein-de Sitter Model

In order to illustrate some of the results obtained above, let us now specialise to a model universe with a critical density of dust, $\Omega_0 = 1$ and $p = 0$, and with zero cosmological constant, $\Omega_\Lambda = 0$. Friedmann's equation then reduces to $H(t) = H_0\, a^{-3/2}$, and the age of the Universe becomes $t_0 = 2(3H_0)^{-1}$. The distance measures are

$$\begin{aligned}
D_{\rm prop}(z_1,z_2) &= \frac{2c}{3H_0}\left[(1+z_1)^{-3/2} - (1+z_2)^{-3/2}\right] \qquad (2.47)\\
D_{\rm com}(z_1,z_2) &= \frac{2c}{H_0}\left[(1+z_1)^{-1/2} - (1+z_2)^{-1/2}\right]\\
D_{\rm ang}(z_1,z_2) &= \frac{2c}{H_0}\,\frac{1}{1+z_2}\left[(1+z_1)^{-1/2} - (1+z_2)^{-1/2}\right]\\
D_{\rm lum}(z_1,z_2) &= \frac{2c}{H_0}\,\frac{1+z_2}{(1+z_1)^2}\left[(1+z_1)^{-1/2} - (1+z_2)^{-1/2}\right]\;.
\end{aligned}$$

2.2. Density Perturbations

The standard model for the formation of structure in the Universe assumes that there were small fluctuations at some very early initial time, which grew by gravitational instability. Although the origin of the seed fluctuations is yet unclear, they possibly originated from quantum fluctuations in the very early Universe, which were blown up during a later inflationary phase. The fluctuations in this case are uncorrelated and the distribution of their amplitudes is Gaussian. Gravitational instability leads to a growth of the amplitudes of the relative density fluctuations. As long as the relative density contrast of the matter fluctuations is much smaller than unity, they can be considered as small perturbations of the otherwise homogeneous and isotropic background density, and linear perturbation theory suffices for their description.

The linear theory of density perturbations in an expanding universe is generally a complicated issue because it needs to be relativistic (e.g. Lifshitz 1946; Bardeen 1980). The reason is that perturbations on any length scale are comparable to or larger than the size of the horizon[2] at sufficiently early times, and then Newtonian theory ceases to be applicable. In other words, since the horizon scale is comparable to the curvature radius of spacetime, Newtonian theory fails for larger-scale perturbations due to non-zero spacetime curvature. The main features can nevertheless be understood by fairly simple reasoning. We shall not present a rigorous mathematical treatment here, but only quote the results which are relevant for our later purposes. For a detailed qualitative and quantitative discussion, we refer the reader to the excellent discussion in chapter 4 of the book by Padmanabhan (1993).

2.2.1. Horizon Size

The size of causally connected regions in the Universe is called the *horizon size*. It is given by the distance by which a photon can travel in the time t since the Big Bang. Since the appropriate time scale is provided by the inverse Hubble parameter $H^{-1}(a)$, the horizon size is $d'_{\rm H} = cH^{-1}(a)$, and the *comoving* horizon size is

$$d_{\rm H} = \frac{c}{aH(a)} = \frac{c}{H_0}\,\Omega_0^{-1/2}\,a^{1/2}\left(1 + \frac{a_{\rm eq}}{a}\right)^{-1/2}\;, \tag{2.48}$$

where we have inserted the Einstein-de Sitter limit (2.36) of Friedmann's equation. The length $cH_0^{-1} = 3h^{-1}\,$Gpc is called the *Hubble radius*. We shall see later that the horizon size at $a_{\rm eq}$ plays a very important rôle for structure formation. Inserting $a = a_{\rm eq}$ into eq. (2.48), yields

$$d_{\rm H}(a_{\rm eq}) = \frac{c}{\sqrt{2}H_0}\,\Omega_0^{-1/2}\,a_{\rm eq}^{1/2} \approx 12\,(\Omega_0 h^2)^{-1}\,{\rm Mpc}\;, \tag{2.49}$$

where $a_{\rm eq}$ from eq. (2.28) has been inserted.

2.2.2. Linear Growth of Density Perturbations

We adopt the commonly held view that the density of the Universe is dominated by weakly interacting dark matter at the relatively late times which are relevant for weak gravitational lensing, $a \gg a_{\rm eq}$. Dark-matter perturbations are characterised by the density contrast

$$\delta(\vec{x},a) = \frac{\rho(\vec{x},a) - \bar\rho(a)}{\bar\rho(a)}\;, \tag{2.50}$$

where $\bar\rho = \rho_0\, a^{-3}$ is the average cosmic density. Relativistic and non-relativistic perturbation theory shows that linear density fluctuations, i.e. perturbations with $\delta \ll 1$, grow like

$$\delta(a) \propto a^{n-2} = \begin{cases} a^2 & \text{before } a_{\rm eq} \\ a & \text{after } a_{\rm eq} \end{cases} \tag{2.51}$$

as long as the Einstein-de Sitter limit holds. For later times, $a \gg a_{\rm eq}$, when the Einstein-de Sitter limit no longer applies if $\Omega_0 \neq 1$

[2] In this context, the size of the horizon is the distance ct by which light can travel in the time t since the big bang.

or $\Omega_\Lambda \neq 0$, the linear growth of density perturbations is changed according to

$$\delta(a) = \delta_0\, a\, \frac{g'(a)}{g'(1)} \equiv \delta_0\, a\, g(a)\;, \qquad (2.52)$$

where δ_0 is the density contrast linearly extrapolated to the present epoch, and the density-dependent growth function $g'(a)$ is accurately fit by (Carroll et al. 1992)

$$\begin{aligned} g'(a;\Omega_0,\Omega_\Lambda) &= \frac{5}{2}\Omega(a)\left[\Omega^{4/7}(a) - \Omega_\Lambda(a)\right. \\ &+ \left.\left(1+\frac{\Omega(a)}{2}\right)\left(1+\frac{\Omega_\Lambda(a)}{70}\right)\right]^{-1} \end{aligned} \qquad (2.53)$$

The dependence of Ω and Ω_Λ on the scale factor a is given in eqs. (2.34). The growth function $a\,g(a;\Omega_0,\Omega_\Lambda)$ is shown in Fig. 6 for a variety of parameters Ω_0 and Ω_Λ.

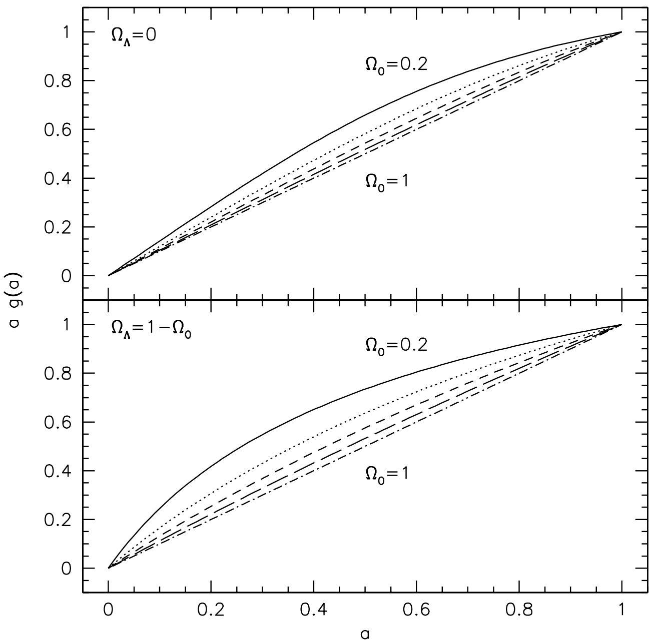

FIG. 6.—The growth function $a\,g(a) \equiv a\,g'(a)/g'(1)$ given in eqs. (2.52) and (2.53) for Ω_0 between 0.2 and 1.0 in steps of 0.2. Top panel: $\Omega_\Lambda = 0$; bottom panel: $\Omega_\Lambda = 1-\Omega_0$. The growth rate is constant for the Einstein-de Sitter model ($\Omega_0 = 1$, $\Omega_\Lambda = 0$), while it is higher for $a \ll 1$ and lower for $a \approx 1$ for low-Ω_0 models. Consequently, structure forms earlier in low- than in high-Ω_0 models.

The cosmic microwave background reveals relative temperature fluctuations of order 10^{-5} on large scales. By the Sachs-Wolfe effect (Sachs & Wolfe 1967), these temperature fluctuations reflect density fluctuations of the same order of magnitude. The cosmic microwave background originated at $a \approx 10^{-3} \gg a_{\rm eq}$, well after the Universe became matter-dominated. Equation (2.51) then implies that the density fluctuations today, expected from the temperature fluctuations at $a \approx 10^{-3}$, should only reach a level of 10^{-2}. Instead, structures (e.g. galaxies) with $\delta \gg 1$ are observed. How can this discrepancy be resolved? The cosmic microwave background displays fluctuations in the baryonic matter component only. If there is an additional matter component that only couples through weak interactions, fluctuations in that component could grow as soon as it decoupled from the cosmic plasma, well before photons decoupled from baryons to set the cosmic microwave background free. Such fluctuations could therefore easily reach the amplitudes observed today, and thereby resolve the apparent mismatch between the amplitudes of the temperature fluctuations in the cosmic microwave background and the present cosmic structures. This is one of the strongest arguments for the existence of a dark matter component in the Universe.

2.2.3. Suppression of Growth

It is convenient to decompose the density contrast δ into Fourier modes. In linear perturbation theory, individual Fourier components evolve independently. A perturbation of (comoving) wavelength λ is said to "enter the horizon" when $\lambda = d_{\rm H}(a)$. If $\lambda < d_{\rm H}(a_{\rm eq})$, the perturbation enters the horizon while radiation is still dominating the expansion. Until $a_{\rm eq}$, the expansion time-scale, $t_{\rm exp} = H^{-1}$, is determined by the radiation density $\rho_{\rm R}$, which is shorter than the collapse time-scale of the dark matter, $t_{\rm DM}$:

$$t_{\rm exp} \sim (G\rho_{\rm R})^{-1/2} < (G\rho_{\rm DM})^{-1/2} \sim t_{\rm DM}\;. \qquad (2.54)$$

In other words, the fast radiation-driven expansion prevents dark-matter perturbations from collapsing. Light can only cross regions that are smaller than the horizon size. The suppression of growth due to radiation is therefore restricted to scales smaller than the horizon, and larger-scale perturbations remain unaffected. This explains why the horizon size at $a_{\rm eq}$, $d_{\rm H}(a_{\rm eq})$, sets an important scale for structure growth.

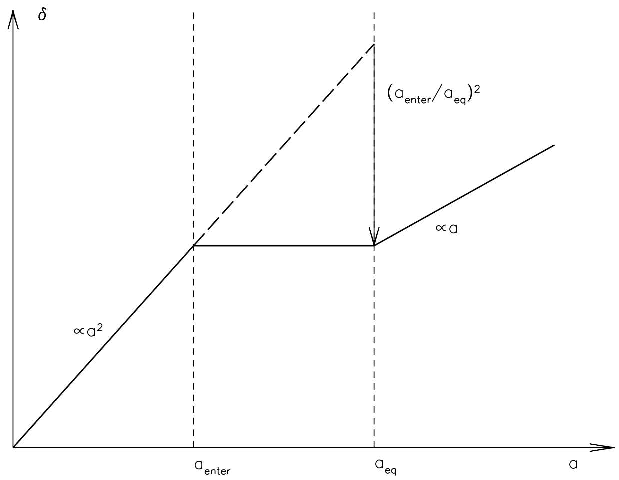

FIG. 7.—Sketch illustrating the suppression of structure growth during the radiation-dominated phase. The perturbation grows $\propto a^2$ before $a_{\rm eq}$, and $\propto a$ thereafter. If the perturbation is smaller than the horizon at $a_{\rm eq}$, it enters the horizon at $a_{\rm enter} < a_{\rm eq}$ while radiation is still dominating. The rapid radiation-driven expansion prevents the perturbation from growing further. Hence it stalls until $a_{\rm eq}$. By then, its amplitude is smaller by $f_{\rm sup} = (a_{\rm enter}/a_{\rm eq})^2$ than it would be without suppression.

Figure 7 illustrates the growth of a perturbation with $\lambda < d_{\rm H}(a_{\rm eq})$, that is small enough to enter the horizon at $a_{\rm enter} < a_{\rm eq}$. It can be read off from the figure that such perturbations are suppressed by the factor

$$f_{\rm sup} = \left(\frac{a_{\rm enter}}{a_{\rm eq}}\right)^2\;. \qquad (2.55)$$

It remains to be evaluated at what time $a_{\rm enter}$ a density perturbation with comoving wavelength λ enters the horizon. The condition is

$$\lambda = d_{\rm H}(a_{\rm enter}) = \frac{c}{a_{\rm enter}\,H(a_{\rm enter})} \; . \tag{2.56}$$

Well in the Einstein-de Sitter regime, the Hubble parameter is given by eq. (2.37). Inserting that expression into (2.56) yields

$$\lambda \propto \left\{ \begin{array}{ll} a_{\rm enter} & (a_{\rm enter} \ll a_{\rm eq}) \\ a_{\rm enter}^{1/2} & (a_{\rm eq} \ll a_{\rm enter} \ll 1) \end{array} \right. \; . \tag{2.57}$$

Let now $k = \lambda^{-1}$ be the wave number of the perturbation, and $k_0 = d_{\rm H}^{-1}(a_{\rm eq})$ the wave number corresponding to the horizon size at $a_{\rm eq}$. The suppression factor (2.55) can then be written

$$f_{\rm sup} = \left(\frac{k_0}{k}\right)^2 \; . \tag{2.58}$$

From eq. (2.49),

$$k_0 \approx 0.083\,(\Omega_0 h^2)\,{\rm Mpc}^{-1} \approx 250\,(\Omega_0 h)\,({\rm Hubble\ radii})^{-1} \; . \tag{2.59}$$

2.2.4. *Density Power Spectrum*

The assumed Gaussian density fluctuations $\delta(\vec{x})$ at the comoving position $\vec{x}$ can completely be characterised by their power spectrum $P_\delta(k)$, which can be defined by (see Sect. 2.4.)

$$\left\langle \hat\delta(\vec{k})\hat\delta^*(\vec{k}') \right\rangle = (2\pi)^3\,\delta_{\rm D}(\vec{k}-\vec{k}')\,P_\delta(k) \; , \tag{2.60}$$

where $\hat\delta(\vec{k})$ is the Fourier transform of δ, and the asterisk denotes complex conjugation. Strictly speaking, the Fourier decomposition is valid only in flat space. However, at early times space is flat in any cosmological model, and at late times the interesting scales k^{-1} of the density perturbations are much smaller than the curvature radius of the Universe. Hence, we can apply Fourier decomposition here.

Consider now the primordial perturbation spectrum at some very early time, $P_{\rm i}(k) = |\hat\delta_{\rm i}^2(k)|$. Since the density contrast grows as $\delta \propto a^{n-2}$ [eq. (2.51)], the spectrum grows as $P_\delta(k) \propto a^{2(n-2)}$. At $a_{\rm enter}$, the spectrum has therefore changed to

$$P_{\rm enter}(k) \propto a_{\rm enter}^{2(n-2)}\,P_{\rm i}(k) \propto k^{-4}\,P_{\rm i}(k) \; . \tag{2.61}$$

where eq. (2.57) was used for $k \gg k_0$.

It is commonly assumed that the total power of the density fluctuations at $a_{\rm enter}$ should be scale-invariant. This implies $k^3\,P_{\rm enter}(k) = {\rm const.}$, or $P_{\rm enter}(k) \propto k^{-3}$. Accordingly, the primordial spectrum has to scale with k as $P_{\rm i}(k) \propto k$. This *scale-invariant* spectrum is called the *Harrison-Zel'dovich* spectrum (Harrison 1970; Peebles & Yu 1970; Zel'dovich 1972). Combining that with the suppression of small-scale modes (2.58), we arrive at

$$P_\delta(k) \propto \left\{ \begin{array}{ll} k & {\rm for}\ k \ll k_0 \\ k^{-3} & {\rm for}\ k \gg k_0 \end{array} \right. \; . \tag{2.62}$$

An additional complication arises when the dark matter consists of particles moving with a velocity comparable to the speed of light. In order to keep them gravitationally bound, density perturbations then have to have a certain minimum mass, or equivalently a certain minimum size. All perturbations smaller than that size are damped away by free streaming of particles. Consequently, the density perturbation spectrum of such particles has an exponential cut-off at large k. This clarifies the distinction between *hot* and *cold* dark matter: Hot dark matter (HDM) consists of fast particles that damp away small-scale perturbations, while cold dark matter (CDM) particles are slow enough to cause no significant damping.

2.2.5. *Normalisation of the Power Spectrum*

Apart from the shape of the power spectrum, its normalisation has to be fixed. Several methods are available which usually yield different answers:

1. Normalisation by microwave-background anisotropies: The COBE satellite has measured fluctuations in the temperature of the microwave sky at the *rms* level of $\Delta T/T \sim 1.3 \times 10^{-5}$ at an angular scale of $\sim 7°$ (Banday et al. 1997). Adopting a shape for the power spectrum, these fluctuations can be translated into an amplitude for $P_\delta(k)$. Due to the large angular scale of the measurement, this kind of amplitude determination specifies the amplitude on large physical scales (small k) only. In addition, microwave-background fluctuations measure the amplitude of scalar *and* tensor perturbation modes, while the growth of density fluctuations is determined by the fluctuation amplitude of scalar modes only.

2. Normalisation by the local variance of galaxy counts, pioneered by Davis & Peebles (1983): Galaxies are supposed to be biased tracers of underlying dark-matter fluctuations (Kaiser 1984; Bardeen et al. 1986; White et al. 1987). By measuring the local variance of galaxy counts within certain volumes, and assuming an expression for the bias, the amplitude of dark-matter fluctuations can be inferred. Conventionally, the variance of galaxy counts $\sigma_{8,\rm galaxies}$ is measured within spheres of radius $8\,h^{-1}\,{\rm Mpc}$, and the result is $\sigma_{8,\rm galaxies} \approx 1$. The problem of finding the corresponding variance σ_8 of matter-density fluctuations is that the exact bias mechanism of galaxy formation is still under debate (e.g. Kauffmann et al. 1997).

3. Normalisation by the local abundance of galaxy clusters (White et al. 1993; Eke et al. 1996; Viana & Liddle 1996): Galaxy clusters form by gravitational instability from dark-matter density perturbations. Their spatial number density reflects the amplitude of appropriate dark-matter fluctuations in a very sensitive manner. It is therefore possible to determine the amplitude of the power spectrum by demanding that the local spatial number density of galaxy clusters be reproduced. Typical scales for dark-matter fluctuations collapsing to galaxy clusters are of order $10\,h^{-1}\,{\rm Mpc}$, hence cluster normalisation determines the amplitude of the power spectrum on just that scale.

Since gravitational lensing by large-scale structures is generally sensitive to scales comparable to $k_0^{-1} \sim 12\,(\Omega_0 h^2)\,{\rm Mpc}$, cluster normalisation appears to be the most appropriate normalisation method for the present purposes. The solid curve in Fig. 8 shows the CDM power spectrum, linearly and non-linearly evolved to $z = 0$ (or $a = 1$) in an Einstein-de Sitter universe with $h = 0.5$, normalised to the local cluster abundance.

2.2.6. *Non-Linear Evolution*

At late stages of the evolution and on small scales, the growth of density fluctuations begins to depart from the linear behaviour of eq. (2.52). Density fluctuations grow non-linear, and fluctuations of different size interact. Generally, the evolution of $P(k)$ then becomes complicated and needs to be evaluated numerically. However, starting from the bold *ansatz* that the two-point correlation functions in the linear and non-linear regimes are related by a general scaling relation (Hamilton et al. 1991), which turns out to hold remarkably well, analytic formulae describing the non-linear behaviour of $P(k)$ have been derived (Jain

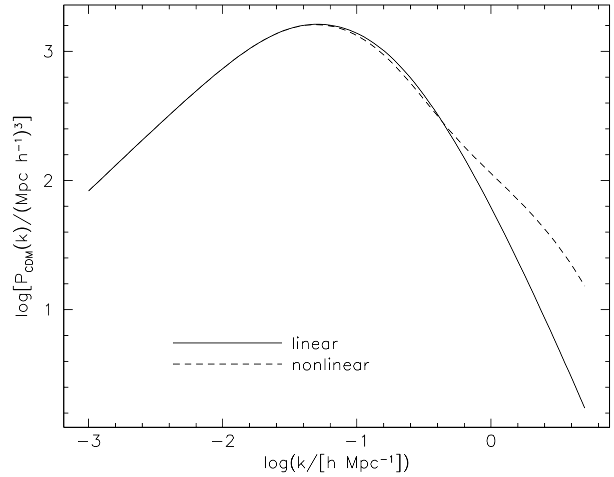

FIG. 8.—CDM power spectrum, normalised to the local abundance of galaxy clusters, for an Einstein-de Sitter universe with $h = 0.5$. Two curves are displayed. The solid curve shows the linear, the dashed curve the non-linear power spectrum. While the linear power spectrum asymptotically falls off $\propto k^{-3}$, the non-linear power spectrum, according to Peacock & Dodds (1996), illustrates the increased power on small scales due to non-linear effects, at the expense of larger-scale structures.

et al. 1995; Peacock & Dodds 1996). It will turn out in subsequent chapters that the non-linear evolution of the density fluctuations is crucial for accurately calculating weak-lensing effects by large-scale structures. As an example, we show as the dashed curve in Fig. 8 the CDM power spectrum in an Einstein-de Sitter universe with $h = 0.5$, normalised to the local cluster abundance, non-linearly evolved to $z = 0$. The non-linear effects are immediately apparent: While the spectrum remains unchanged for large scales ($k \ll k_0$), the amplitude on small scales ($k \gg k_0$) is substantially increased at the expense of scales just above the peak. It should be noted that non-linearly evolved density fluctuations are no longer fully characterised by the power spectrum only, because then non-Gaussian features develop.

2.2.7. *Poisson's Equation*

Localised density perturbations which are much smaller than the horizon and whose peculiar velocities relative to the mean motion in the Universe are much smaller than the speed of light, can be described by Newtonian gravity. Their gravitational potential obeys Poisson's equation,

$$\nabla_r^2\Phi' = 4\pi G\rho \,, \tag{2.63}$$

where $\rho = (1+\delta)\bar\rho$ is the total matter density, and Φ' is the sum of the potentials of the smooth background $\bar\Phi$ and the potential of the perturbation Φ. The gradient ∇_r operates with respect to the physical, or proper, coordinates. Since Poisson's equation is linear, we can subtract the background contribution $\nabla_r^2\bar\Phi = 4\pi G\bar\rho$. Introducing the gradient with respect to comoving coordinates $\nabla_x = a\nabla_r$, we can write eq. (2.63) in the form

$$\nabla_x^2\Phi = 4\pi G a^2\,\bar\rho\,\delta \,. \tag{2.64}$$

In the matter-dominated epoch, $\bar\rho = a^{-3}\bar\rho_0$. With the critical density (2.15), Poisson's equation can be re-written as

$$\nabla_x^2\Phi = \frac{3H_0^2}{2a}\Omega_0\delta \,. \tag{2.65}$$

2.3. *Relevant Properties of Lenses and Sources*

Individual reviews have been written on galaxies (e.g. Faber & Gallagher 1979; Binggeli et al. 1988; Giovanelli & Haynes 1991; Koo & Kron 1992; Ellis 1997), clusters of galaxies (e.g. Bahcall 1977; Rood 1981; Forman & Jones 1982; Bahcall 1988; Sarazin 1986), and active galactic nuclei (e.g. Rees 1984; Weedman 1986; Blandford et al. 1990; Hartwick & Schade 1990; Warren & Hewett 1990; Antonucci 1993; Peterson 1997). A detailed presentation of these objects is not the purpose of this review. It suffices here to summarise those properties of these objects that are relevant for understanding the following discussion. Properties and peculiarities of individual objects are not necessary to know; rather, we need to specify the objects statistically. This section will therefore focus on a statistical description, leaving subtleties aside.

2.3.1. *Galaxies*

For the purposes of this review, we need to characterise the statistical properties of galaxies as a class. Galaxies can broadly be grouped into two populations, dubbed *early-type* and *late-type* galaxies, or *ellipticals* and *spirals*, respectively. While spiral galaxies include disks structured by more or less pronounced spiral arms, and approximately spherical bulges centred on the disk centre, elliptical galaxies exhibit amorphous projected light distributions with roughly elliptical isophotes. There are, of course, more elaborate morphological classification schemes (e.g. de Vaucouleurs et al. 1991; Buta et al. 1994; Naim et al. 1995a; Naim et al. 1995b), but the broad distinction between ellipticals and spirals suffices for this review.

Outside galaxy clusters, the galaxy population consists of about $3/4$ spiral galaxies and $1/4$ elliptical galaxies, while the fraction of ellipticals increases towards cluster centres. Elliptical galaxies are typically more massive than spirals. They contain little gas, and their stellar population is older, and thus 'redder', than in spiral galaxies. In spirals, there is a substantial amount of gas in the disk, providing the material for ongoing formation of new stars. Likewise, there is little dust in ellipticals, but possibly large amounts of dust are associated with the gas in spirals.

Massive galaxies have of order 10^{11} solar masses, or 2×10^{44} g within their visible radius. Such galaxies have luminosities of order 10^{10} times the solar luminosity. The kinematics of the stars, gas and molecular clouds in galaxies, as revealed by spectroscopy, indicate that there is a relation between the characteristic velocities inside galaxies and their luminosity (Faber & Jackson 1976; Tully & Fisher 1977); brighter galaxies tend to have larger masses.

The differential luminosity distribution of galaxies can very well be described by the functional form

$$\Phi(L)\,\frac{\mathrm{d}L}{L_*} = \Phi_0\left(\frac{L}{L_*}\right)^{-\nu}\exp\left(-\frac{L}{L_*}\right)\frac{\mathrm{d}L}{L_*} \,, \tag{2.66}$$

proposed by Schechter (1976). The parameters have been measured to be

$$\nu \approx 1.1 \,, \quad L_* \approx 1.1\times10^{10}L_\odot \,, \quad \Phi_* \approx 1.5\times10^{-2}h^3\,\mathrm{Mpc}^{-3} \tag{2.67}$$

(e.g. Efstathiou et al. 1988; Marzke et al. 1994a; Marzke et al. 1994b). This distribution means that there is essentially a sharp cut-off in the galaxy population above luminosities of $\sim L_*$, and the mean separation between L_*-galaxies is of order $\sim \Phi_*^{-1/3} \approx 4\,h^{-1}$ Mpc.

The stars in elliptical galaxies have randomly oriented orbits, while by far the most stars in spirals have orbits roughly coplanar with the galactic disks. Stellar velocities are therefore characterised by a velocity dispersion σ_v in ellipticals, and by an asymptotic circular velocity v_c in spirals.[3] These characteristic velocities are related to galaxy luminosities by laws of the form

$$\frac{\sigma_v}{\sigma_{v,*}} = \left(\frac{L}{L_*}\right)^{1/\alpha} = \frac{v_c}{v_{c,*}} , \qquad (2.68)$$

where α ranges around $3-4$. For spirals, eq. (2.68) is called Tully-Fisher (Tully & Fisher 1977) relation, for ellipticals Faber-Jackson (Faber & Jackson 1976) relation. Both velocity scales $\sigma_{v,*}$ and $v_{c,*}$ are of order $220\,\mathrm{km\,s^{-1}}$. Since $v_c = \sqrt{2}\sigma_v$, ellipticals with the same luminosity are more massive than spirals.

Most relevant for weak gravitational lensing is a population of faint galaxies emitting bluer light than local galaxies, the so-called *faint blue galaxies* (Tyson 1988; see Ellis 1997 for a review). There are of order $30-50$ such galaxies per square arc minute on the sky which can be mapped with current ground-based optical telescopes, i.e. there are $\approx 20,000-40,000$ such galaxies on the area of the full moon. The picture that the sky is covered with a 'wall paper' of those faint and presumably distant blue galaxies is therefore justified. It is this fine-grained pattern on the sky that makes many weak-lensing studies possible in the first place, because it allows the detection of the coherent distortions imprinted by gravitational lensing on the images of the faint blue galaxy population.

Due to their faintness, redshifts of the faint blue galaxies are hard to measure spectroscopically. The following picture, however, seems to be reasonably secure. It has emerged from increasingly deep and detailed observations (see, e.g. Broadhurst et al. 1988; Colless et al. 1991; Colless et al. 1993; Lilly et al. 1991; Lilly 1993; Crampton et al. 1995; and also the reviews by Koo & Kron 1992 and Ellis 1997). The redshift distribution of faint galaxies has been found to agree fairly well with that expected for a non-evolving comoving number density. While the galaxy number counts in blue light are substantially above an extrapolation of the local counts down to increasingly faint magnitudes, those in the red spectral bands agree fairly well with extrapolations from local number densities. Further, while there is significant evolution of the luminosity function in the blue, in that the luminosity scale L_* of a Schechter-type fit increases with redshift, the luminosity function of the galaxies in the red shows little sign of evolution. Highly resolved images of faint blue galaxies obtained with the *Hubble Space Telescope* are now becoming available. In red light, they reveal mostly ordinary spiral galaxies, while their substantial emission in blue light is more concentrated to either spiral arms or bulges. Spectra exhibit emission lines characteristic of star formation.

These findings support the view that the galaxy evolution towards higher redshifts apparent in blue light results from enhanced star-formation activity taking place in a population of galaxies which, apart from that, may remain unchanged even out to redshifts of $z \gtrsim 1$. The redshift distribution of the faint blue galaxies is then sufficiently well described by

$$p(z)\mathrm{d}z = \frac{\beta}{z_0^3\,\Gamma(3/\beta)}\,z^2 \exp\left[-\left(\frac{z}{z_0}\right)^{\beta}\right]\mathrm{d}z . \qquad (2.69)$$

This expression is normalised to $0 \le z < \infty$ and provides a good fit to the observed redshift distribution (e.g. Smail et al. 1995b). The mean redshift $\langle z\rangle$ is proportional to z_0, and the parameter β describes how steeply the distribution falls off beyond z_0. For $\beta = 1.5$, $\langle z\rangle \approx 1.5\,z_0$. The parameter z_0 depends on the magnitude cutoff and the colour selection of the galaxy sample.

Background galaxies would be ideal tracers of distortions caused by gravitational lensing if they were intrinsically circular. Then, any measured ellipticity would directly reflect the action of the gravitational tidal field of the lenses. Unfortunately, this is not the case. To first approximation, galaxies have intrinsically elliptical shapes, but the ellipses are randomly oriented. The intrinsic ellipticities introduce noise into the inference of the tidal field from observed ellipticities, and it is important for the quantification of the noise to know the intrinsic ellipticity distribution. Let $|\varepsilon|$ be the ellipticity of a galaxy image, defined such that for an ellipse with axes a and $b < a$,

$$|\varepsilon| \equiv \frac{a-b}{a+b} . \qquad (2.70)$$

Ellipses have an orientation, hence the ellipticity has two components $\varepsilon_{1,2}$, with $|\varepsilon| = (\varepsilon_1^2+\varepsilon_2^2)^{1/2}$. It turns out empirically that a Gaussian is a good description for the ellipticity distribution,

$$p_\varepsilon(\varepsilon_1,\varepsilon_2)\mathrm{d}\varepsilon_1\mathrm{d}\varepsilon_2 = \frac{\exp(-|\varepsilon|^2/\sigma_\varepsilon^2)}{\pi\sigma_\varepsilon^2\left[1-\exp(-1/\sigma_\varepsilon^2)\right]}\mathrm{d}\varepsilon_1\mathrm{d}\varepsilon_2 , \qquad (2.71)$$

with a characteristic width of $\sigma_\varepsilon \approx 0.2$ (e.g. Miralda-Escude 1991; Tyson & Seitzer 1988; Brainerd et al. 1996). We will later (Sect. 4.2.) define galaxy ellipticities for the general situation where the isophotes are not ellipses. This completes our summary of galaxy properties as required here.

2.3.2. Groups and Clusters of Galaxies

Galaxies are not randomly distributed in the sky. Their positions are correlated, and there are areas in the sky where the galaxy density is noticeably higher or lower than average (cf. the galaxy count map in Fig. 9). There are groups consisting of a few galaxies, and there are *clusters of galaxies* in which some hundred up to a thousand galaxies appear very close together.

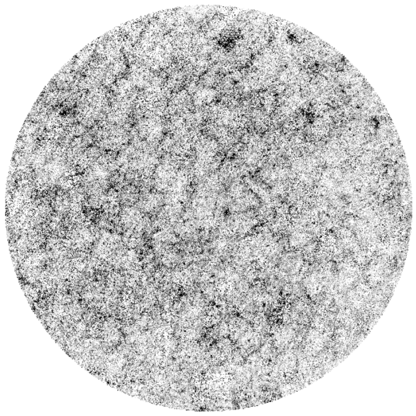

FIG. 9.—The Lick galaxy counts within 50° radius around the North Galactic pole (Seldner et al. 1977). The galaxy number density is highest at the black and lowest at the white regions on the map. The picture illustrates structure in the distribution of fairly nearby galaxies, viz. under-dense regions, long extended filaments, and clusters of galaxies.

The most prominent galaxy cluster in the sky covers a huge area centred on the Virgo constellation. Its central region has a diameter of about 7°, and its main body extends over roughly

[3] The circular velocity of stars and gas in spiral galaxies turns out to be fairly independent of radius, except close to their centre. These flat rotations curves cannot be caused by the observable matter in these galaxies, but provide strong evidence for the presence of a dark halo, with density profile $\rho \propto r^{-2}$ at large radii.

$15^\circ \times 40^\circ$. It was already noted by Sir William Herschel in the 18th century that the entire Virgo cluster covers about $1/8$th of the sky, while containing about $1/3$rd of the galaxies observable at that time.

Zwicky noted in 1933 that the galaxies in the Coma cluster and other rich clusters move so fast that the clusters required about ten 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. (1968) identifies 9134 clusters north of -3° declination. Cluster masses can exceed 10^{48} g or $5 \times 10^{14} M_\odot$, and they have typical radii of $\approx 5 \times 10^{24}$ cm or ≈ 1.5 Mpc.

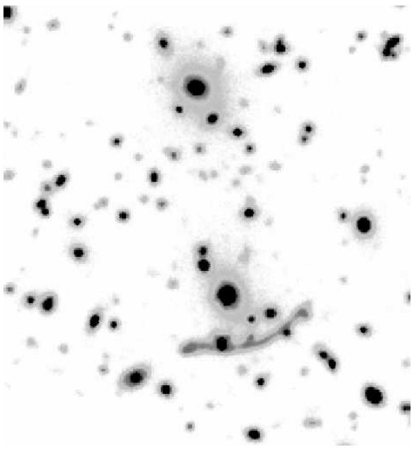

FIG. 10.—The galaxy cluster Abell 370, in which the first gravitationally lensed arc was detected (Lynds & Petrosian 1986; Soucail et al. 1987a, 1987b). Most of the bright galaxies seen are cluster members at $z = 0.37$, whereas the arc, i.e. the highly elongated feature, is the image of a galaxy at redshift $z = 0.724$ (Soucail et al. 1988). (courtesy of J.-P. Kneib)

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* (free-free radiation) from a hot, dilute plasma with temperatures in the range $(10^7 - 10^8)$ K and densities of $\sim 10^{-3}$ particles per cm^3. Based on the assumption that this intra-cluster gas is in hydrostatic equilibrium with a spherically symmetric 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 employing the virial theorem. The mass of the intra-cluster 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, 1987b; see Fig. 10). Their light is typically bluer than that from the cluster galaxies, and their length is comparable to the size of the central cluster region. Paczyński (1987) suggested that these *arcs* are images of galaxies in the background of the clusters which are strongly distorted by the gravitational tidal field close to the cluster centres. This explanation was generally accepted when spectroscopy revealed that the sources of the arcs 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 centre, images of background galaxies are only weakly deformed, and they are referred to as *arclets* (Fort et al. 1988; Tyson et al. 1990). The high number density of faint arclets allows one to measure the coherent distortion caused by the tidal gravitational field of the cluster out to fairly large radii. One of the main applications of weak gravitational lensing is to reconstruct the (projected) mass distribution of galaxy clusters from their measurable tidal fields. Consequently, the corresponding theory constitutes one of the largest sections of this review.

Such strong and weak 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 in their own right, 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 et al. 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 cosmic history than if the matter density is low (first noted by Richstone et al. 1992). This is due to the behaviour of the growth factor shown in Fig. 6, combined with the Gaussian nature of the initial density fluctuations. Consequently, the compactness and the morphology of clusters reflect the cosmic matter density, and this has various observable implications. One method to normalise the density-perturbation power spectrum fixes its overall amplitude such that the local spatial number density of galaxy clusters is reproduced. This method, called *cluster normalisation* and pioneered by White et al. (1993), will frequently be used in this review.

In summary, clusters are not only regions of higher galaxy number density in the sky, but they are gravitationally bound bodies whose member galaxies contribute only a small fraction of their mass. About 80% of their mass is dark, and roughly 10% is in the form of the diffuse, X–ray emitting gas spread throughout the cluster. Mass estimates inferred from galaxy kinematics, X–ray emission, and gravitational-lensing effects generally agree to within about a factor of two, typically arriving at masses of order 5×10^{14} solar masses, or 10^{48} g. Typical sizes of galaxy clusters are of order several megaparsecs, or 5×10^{24} cm. In addition, there are smaller objects, called *galaxy groups*, which contain fewer galaxies and have typical masses of order 10^{13} solar masses.

2.3.3. Active Galactic Nuclei

The term 'active galactic nuclei' (AGNs) is applied to galaxies which show signs of non-stellar radiation in their centres. Whereas the emission from 'normal' galaxies like our own is completely dominated by radiation from stars and their remnants, the emission from AGNs is a combination of stellar light and non-thermal emission from their nuclei. In fact, the most prominent class of AGNs, the quasi-stellar radio sources, or quasars, have their names derived from the fact that their optical appearance is point-like. The nuclear emission almost completely outshines the extended stellar light of its host galaxy.

AGNs do not form a homogeneous class of objects. Instead, they are grouped into several types. The main classes are: quasars, quasi-stellar objects (QSOs), Seyfert galaxies, BL Lacertae objects (BL Lacs), and radio galaxies. What unifies them is the non-thermal emission from their nucleus, which manifests itself in various ways: (1) radio emission which, owing to its spectrum and polarisation, is interpreted as synchrotron radiation from a power-law distribution of relativistic electrons; (2) strong ultraviolet and optical emission lines from highly ionised species, which in some cases can be extremely broad, corresponding to Doppler velocities up to $\sim 20,000\,\mathrm{km\,s^{-1}}$, thus indicating the presence of semi-relativistic velocities in the emission region; (3) a flat ultraviolet-to-optical continuum spectrum, often accompanied by polarisation of the optical light, which cannot naturally be explained by a superposition of stellar (Planck) spectra; (4) strong X–ray emission with a hard power-law spectrum, which can be interpreted as inverse Compton radiation by a population of relativistic electrons with a power-law energy distribution; (5) strong gamma-ray emission; (6) variability at all wavelengths, from the radio to the gamma-ray regime. Not all these phenomena occur at the same level in all the classes of AGNs. QSOs, for example, can roughly be grouped into radio-quiet QSOs and quasars, the latter emitting strongly at radio wavelengths.

Since substantial variability cannot occur on timescales shorter than the light-travel time across the emitting region, the variability provides a rigorous constraint on the compactness of the region emitting the bulk of the nuclear radiation. In fact, this causality argument based on light-travel time can mildly be violated if relativistic velocities are present in the emitting region. Direct evidence for this comes from the observation of the so-called superluminal motion, where radio-source components exhibit apparent velocities in excess of c (e.g. Zensus & Pearson 1987). This can be understood as a projection effect, combining velocities close to (but of course smaller than) the velocity of light with a velocity direction close to the line-of-sight to the observer. Observations of superluminal motion indicate that bulk velocities of the radio-emitting plasma components can have Lorentz factors of order 10, i.e., they move at $\sim 0.99\,c$.

The standard picture for the origin of this nuclear activity is that a supermassive black hole (or order $10^8 M_\odot$), situated in the centre of the host galaxy, accretes gas from the host. In this process, gravitational binding energy is released, part of which can be transformed into radiation. The appearance of an AGN then depends on the black-hole mass and angular momentum, the accretion rate, the efficiency of the transformation of binding energy into radiation, and on the orientation relative to the line-of-sight. The understanding of the physical mechanisms in AGNs, and how they are related to their phenomenology, is still rather incomplete. We refer the reader to the books and articles by Begelman et al. (1984), Weedman (1986), Blandford et al. (1990), Peterson (1997), and Krolik (1999), and references therein, for an overview of the phenomena in AGNs, and of our current ideas on their interpretation. For the current review, we only make use of one particular property of AGNs:

QSOs can be extremely luminous. Their optical luminosity can reach a factor of thousand or more times the luminosity of normal galaxies. Therefore, their nuclear activity completely outshines that of the host galaxy, and the nuclear sources appear point-like on optical images. Furthermore, the high luminosity implies that QSOs can be seen to very large distances, and in fact, until a few years ago QSOs held the redshift record. In addition, the comoving number density of QSOs evolves rapidly with redshift. It was larger than today by a factor of ~ 100 at redshifts between 2 and 3. Taken together, these two facts imply that a flux-limited sample of QSOs has a very broad redshift distribution, in particular, very distant objects are abundant in such a sample.

However, it is quite difficult to obtain a 'complete' flux-limited sample of QSOs. Of all point-like objects at optical wavelengths, QSOs constitute only a tiny fraction, most being stars. Hence, morphology alone does not suffice to obtain a candidate QSO sample which can be verified spectroscopically. However, QSOs are found to have very blue optical colours, by which they can efficiently be selected. Colour selection typically yields equal numbers of white dwarfs and QSOs with redshifts below ~ 2.3. For higher-redshift QSOs, the strong Lyα emission line moves from the U-band filter into the B-band, yielding redder U$-$B colours. For these higher-redshift QSOs, multi-colour or emission-line selection criteria must be used (cf. Fan et al. 1999). In contrast to optical selection, AGNs are quite efficiently selected in radio surveys. The majority of sources selected at centimeter wavelengths are AGNs. A flux-limited sample of radio-selected AGNs also has a very broad redshift distribution. The large fraction of distant objects in these samples make AGNs particularly promising sources for the gravitational lensing effect, as the probability of finding an intervening mass concentration close to the line-of-sight increases with the source distance. In fact, most of the known multiple-image gravitational lens systems have AGN sources.

In addition to their high redshifts, the number counts of AGNs are important for lensing. For bright QSOs with apparent B-band magnitudes $B \lesssim 19$, the differential source counts can be approximated by a power law, $n(S) \propto S^{-(\alpha+1)}$, where $n(S)\,\mathrm{d}S$ is the number density of QSOs per unit solid angle with flux within $\mathrm{d}S$ of S, and $\alpha \approx 2.6$. At fainter magnitudes, the differential source counts can also be approximated by a power law in flux, but with a much flatter index of $\alpha \sim 0.5$. The source counts at radio wavelengths are also quite steep for the highest fluxes, and flatten as the flux decreases. The steepness of the source counts will be the decisive property of AGNs for the magnification bias, which will be discussed in Sect. 6..

2.4. *Correlation Functions, Power Spectra, and their Projections*

2.4.1. *Definitions; Homogeneous and Isotropic Random Fields*

In this subsection, we define the correlation function and the power spectrum of a random field, which will be used extensively in later sections. One example already occurred above, namely the power spectrum P_δ of the density fluctuation field δ.

Consider a random field $g(\vec{x})$ whose expectation value is zero everywhere. This means that an average over many realisations of the random field should vanish, $\langle g(\vec{x})\rangle = 0$, for all $\vec{x}$. This is not an important restriction, for if that was not the case, we could consider the field $g(\vec{x}) - \langle g(\vec{x})\rangle$ instead, which would have the desired property. Spatial positions $\vec{x}$ have n dimensions, and the field can be either real or complex.

A random field $g(\vec{x})$ is called *homogeneous* if it cannot statistically be distinguished from the field $g(\vec{x}+\vec{y})$, where $\vec{y}$ is an arbitrary translation vector. Similarly, a random field $g(\vec{x})$ is called *isotropic* if it has the same statistical properties as the random field $g(\mathcal{R}\vec{x})$, where $\mathcal{R}$ is an arbitrary rotation matrix in n dimensions. Restricting our attention to homogeneous and isotropic random fields, we note that the *two-point correlation function*

$$\langle g(\vec{x})\, g^*(\vec{y})\rangle = C_{gg}(|\vec{x}-\vec{y}|) \tag{2.72}$$

can only depend on the absolute value of the difference vector between the two points $\vec{x}$ and $\vec{y}$. Note that C_{gg} is real, even if g is complex. This can be seen by taking the complex conjugate of (2.72), which is equivalent to interchanging $\vec{x}$ and $\vec{y}$, leaving the right-hand-side unaffected.

We define the Fourier-transform pair of g as

$$\hat{g}(\vec{k}) = \int_{\mathbb{R}^n} \mathrm{d}^n x\, g(\vec{x})\, \mathrm{e}^{\mathrm{i}\vec{x}\cdot\vec{k}}\;; \quad g(\vec{x}) = \int_{\mathbb{R}^n} \frac{\mathrm{d}^n k}{(2\pi)^n} \hat{g}(\vec{k})\, \mathrm{e}^{-\mathrm{i}\vec{x}\cdot\vec{k}}\;. \tag{2.73}$$

We now calculate the correlation function in Fourier space,

$$\langle \hat{g}(\vec{k})\hat{g}^*(\vec{k}')\rangle = \int_{\mathbb{R}^n} \mathrm{d}^n x\, \mathrm{e}^{\mathrm{i}\vec{x}\cdot\vec{k}} \int_{\mathbb{R}^n} \mathrm{d}^n x'\, \mathrm{e}^{-\mathrm{i}\vec{x}'\cdot\vec{k}'} \langle g(\vec{x})\, g^*(\vec{x}')\rangle\;. \tag{2.74}$$

Using (2.72) and substituting $\vec{x}' = \vec{x}+\vec{y}$, this becomes

$$\begin{aligned} \langle \hat{g}(\vec{k})\hat{g}^*(\vec{k}')\rangle &= \int_{\mathbb{R}^n} \mathrm{d}^n x\, \mathrm{e}^{\mathrm{i}\vec{x}\cdot\vec{k}} \int_{\mathbb{R}^n} \mathrm{d}^n y\, \mathrm{e}^{-\mathrm{i}(\vec{x}+\vec{y})\cdot\vec{k}'} C_{gg}(|\vec{y}|) \\ &= (2\pi)^n \delta_{\mathrm{D}}(\vec{k}-\vec{k}') \int_{\mathbb{R}^n} \mathrm{d}^n y\, \mathrm{e}^{-\mathrm{i}\vec{y}\cdot\vec{k}} C_{gg}(|\vec{y}|) \\ &\equiv (2\pi)^n \delta_{\mathrm{D}}(\vec{k}-\vec{k}')\, P_g(|\vec{k}|)\;. \end{aligned} \tag{2.75}$$

In the final step, we defined the *power spectrum* of the homogeneous and isotropic random field g,

$$P_g(|\vec{k}|) = \int_{\mathbb{R}^n} \mathrm{d}^n y\, \mathrm{e}^{-\mathrm{i}\vec{y}\cdot\vec{k}} C_{gg}(|\vec{y}|)\;, \tag{2.76}$$

which is the Fourier transform of the two-point correlation function. Isotropy of the random field implies that P_g can only depend on the modulus of $\vec{k}$.

Gaussian random fields are characterised by the property that the probability distribution of any linear combination of the random field $g(\vec{x})$ is Gaussian. More generally, the joint probability distribution of a number M of linear combinations of the random variable $g(\vec{x}_i)$ is a multivariate Gaussian. This is equivalent to requiring that the Fourier components $\hat{g}(\vec{k})$ are mutually statistically independent, and that the probability densities for the $\hat{g}(\vec{k})$ are Gaussian with dispersion $P_g(|\vec{k}|)$. Thus, a Gaussian random field is fully characterised by its power spectrum.

2.4.2. *Projections; Limber's Equation*

We now derive a relation between the power spectrum (or the correlation function) of a homogeneous isotropic random field in three dimensions, and its projection onto two dimensions. Specifically, for the three-dimensional field, we consider the density contrast $\delta[f_K(w)\vec{\theta}, w]$, where $\vec{\theta}$ is a two-dimensional vector, which could be an angular position on the sky. Hence, $f_K(w)\vec{\theta}$ and w form a local comoving isotropic Cartesian coordinate system. We define two different projections of δ along the backward-directed light cone of the observer at $w=0$, $t=t_0$,

$$g_i(\vec{\theta}) = \int \mathrm{d}w\, q_i(w)\, \delta[f_K(w)\vec{\theta}, w]\;, \tag{2.77}$$

for $i=1,2$. The $q_i(w)$ are weight functions, and the integral extends from $w=0$ to the horizon $w=w_{\mathrm{H}}$. Since δ is a homogeneous and isotropic random field, so is its projection. Consider now the correlation function

$$\begin{aligned} C_{12} &= \langle g_1(\vec{\theta})g_2(\vec{\theta}')\rangle \\ &= \int \mathrm{d}w\, q_1(w) \int \mathrm{d}w'\, q_2(w') \\ &\times\; \langle \delta[f_K(w)\vec{\theta}, w]\, \delta[f_K(w')\vec{\theta}', w']\rangle\;. \end{aligned} \tag{2.78}$$

We assume that there is no power in the density fluctuations on scales larger than a coherence scale L_{coh}. This is justified because the power spectrum P_δ declines $\propto k$ as $k \to 0$; see (2.62). This implies that the correlation function on the right-hand side of eq. (2.78) vanishes for $w_{\mathrm{H}} \gg |w-w'| \gtrsim L_{\mathrm{coh}}$. Although δ evolves cosmologically, it can be considered constant over a time scale on which light travels across a comoving distance L_{coh}. We note that the second argument of δ simultaneously denotes the third local spatial dimension and the cosmological epoch, related through the light-cone condition $|c\mathrm{d}t| = a\mathrm{d}w$. Furthermore, we assume that the weight functions $q_i(w)$ do not vary appreciably over a scale $\Delta w \le L_{\mathrm{coh}}$. Consequently, $|w-w'| \lesssim L_{\mathrm{coh}}$ over the scale where $C_{\delta\delta}$ is non-zero, and we can set $f_K(w') \approx f_K(w)$ and $q_2(w') = q_2(w)$ to obtain

$$\begin{aligned} C_{12}(\theta) &= \int \mathrm{d}w\, q_1(w)q_2(w) \int \mathrm{d}(\Delta w) \\ &\times\; C_{\delta\delta}\left(\sqrt{f_K^2(w)\theta^2 + (\Delta w)^2}, w\right)\;. \end{aligned} \tag{2.79}$$

The second argument of $C_{\delta\delta}$ now denotes the dependence of the correlation function on cosmic epoch. Equation (2.79) is one form of Limber's (1953) equation, which relates the two-point correlation of the *projected* field to that of the *three-dimensional* field.

Another very useful form of this equation relates the projected two-point correlation function to the power spectrum of the three-dimensional field. The easiest way to derive this relation is by replacing the δ's in (2.78) by their Fourier transforms, where upon

$$\begin{aligned} C_{12} &= \int \mathrm{d}w\, q_1(w) \int \mathrm{d}w'\, q_2(w') \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \\ &\times\; \langle \hat{\delta}(\vec{k}, w)\, \hat{\delta}^*(\vec{k}', w')\rangle \\ &\times\; \mathrm{e}^{-\mathrm{i}f_K(w)\vec{k}_\perp\cdot\vec{\theta}}\, \mathrm{e}^{\mathrm{i}f_K(w')\vec{k}'_\perp\cdot\vec{\theta}'}\, \mathrm{e}^{-\mathrm{i}k_3 w}\mathrm{e}^{\mathrm{i}k'_3 w'}\;. \end{aligned} \tag{2.80}$$

$\vec{k}_\perp$ is the two-dimensional wave vector perpendicular to the line-of-sight. The correlator can be replaced by the power spectrum P_δ using (2.75). This introduces a Dirac delta function $\delta_{\mathrm{D}}(\vec{k}-\vec{k}')$, which allows us to carry out the $\vec{k}'$-integration. Under the same assumptions on the spatial variation of $q_i(w)$ and $f_K(w)$ as before, we find

$$\begin{aligned} C_{12} &= \int \mathrm{d}w\, q_1(w)q_2(w) \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P_\delta(|\vec{k}|, w) \\ &\times\; \mathrm{e}^{-\mathrm{i}f_K(w)\vec{k}_\perp\cdot(\vec{\theta}-\vec{\theta}')}\, \mathrm{e}^{-\mathrm{i}k_3 w} \int \mathrm{d}w'\, \mathrm{e}^{\mathrm{i}k_3 w'}\;. \end{aligned} \tag{2.81}$$

The final integral yields $2\pi\delta_{\mathrm{D}}(k_3)$, indicating that only such modes contribute to the projected correlation function whose wave-vectors lie in the plane of the sky (Blandford et al. 1991). Finally, carrying out the trivial k_3-integration yields

$$\begin{aligned} C_{12}(\theta) &= \int \mathrm{d}w\, q_1(w)q_2(w) \\ &\times\; \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} P_\delta(|\vec{k}_\perp|, w)\, \mathrm{e}^{-\mathrm{i}f_K(w)\vec{k}_\perp\cdot\vec{\theta}} \qquad (2.82) \\ &= \int \mathrm{d}w\, q_1(w)q_2(w) \\ &\times\; \int \frac{k\mathrm{d}k}{2\pi} P(k,w)\, \mathrm{J}_0[f_K(w)\,\theta k]\;. \qquad (2.83) \end{aligned}$$

The definition (2.73) of the Fourier transform, and the relation (2.76) between power spectrum and correlation function allow us to write the (cross) power spectrum $P_{12}(l)$ as

$$\begin{aligned} P_{12}(l) &= \int \mathrm{d}^2\theta\, C_{12}(\theta) \mathrm{e}^{\mathrm{i}\vec{l}\cdot\vec{\theta}} \\ &= \int \mathrm{d}w\, q_1(w) q_2(w) \\ &\times \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} P_\delta(|\vec{k}_\perp|, w)\, (2\pi)^2\, \delta_{\mathrm{D}}[\vec{l} - f_k(w)\vec{k}_\perp] \\ &= \int \mathrm{d}w\, \frac{q_1(w) q_2(w)}{f_K^2(w)} P_\delta\left(\frac{l}{f_K(w)}, w\right) , \qquad (2.84) \end{aligned}$$

which is Limber's equation in Fourier space (Kaiser 1992, 1998). We shall make extensive use of these relations in later sections.

3. GRAVITATIONAL LIGHT DEFLECTION

In this section, we summarise the theoretical basis for the description of light deflection by gravitational fields. Granted the validity of Einstein's Theory of General Relativity, light propagates on the null geodesics of the space-time metric. However, most astrophysically relevant situations permit a much simpler approximate description of light rays, which is called gravitational lens theory; we first describe this theory in Sect. 3.1.. It is sufficient for the treatment of lensing by galaxy clusters in Sect. 5., where the deflecting mass is localised in a region small compared to the distance between source and deflector, and between deflector and observer. In contrast, mass distributions on a cosmic scale cause small light deflections all along the path from the source to the observer. The magnification and shear effects resulting therefrom require a more general description, which we shall develop in Sect. 3.2.. In particular, we outline how the gravitational lens approximation derives from this more general description.

3.1. *Gravitational Lens Theory*

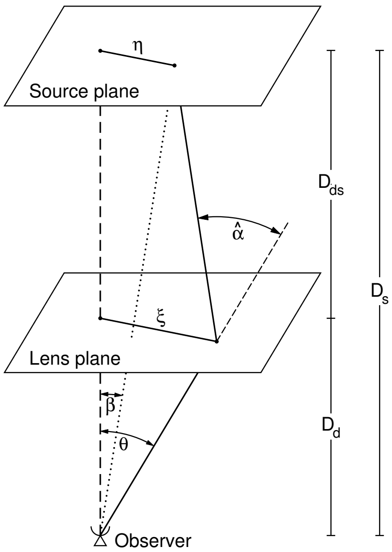

FIG. 11.—Sketch of a typical gravitational lens system.

A typical situation considered in gravitational lensing is sketched in Fig. 11, where a mass concentration at redshift z_{d} (or angular diameter distance D_{d}) deflects the light rays from a source at redshift z_{s} (or angular diameter distance D_{s}). If there are no other deflectors close to the line-of-sight, and if the extent of the deflecting mass along the line-of-sight is very much smaller than both D_{d} and the angular diameter distance D_{ds} from the deflector to the source,[4] the actual light rays which are smoothly curved in the neighbourhood of the deflector can be replaced by two straight rays with a kink near the deflector. The magnitude and direction of this kink is described by the *deflection angle* $\hat{\vec{\alpha}}$, which depends on the mass distribution of the deflector and the impact vector of the light ray.

3.1.1. *The Deflection Angle*

Consider first the deflection by a point mass M. If the light ray does not propagate through the strong gravitational field close to the horizon, that is, if its impact parameter ξ is much larger than the Schwarzschild radius of the lens, $\xi \gg R_{\mathrm{S}} \equiv 2GMc^{-2}$, then General Relativity predicts that the deflection angle $\hat{\alpha}$ is

$$\hat{\alpha} = \frac{4GM}{c^2\xi} . \qquad (3.1)$$

This is just twice the value obtained in Newtonian gravity (see the historical remarks in Schneider et al. 1992). According to the condition $\xi \gg R_{\mathrm{S}}$, the deflection angle is small, $\hat{\alpha} \ll 1$.

The field equations of General Relativity can be linearised if the gravitational field is weak. The deflection angle of an ensemble of point masses is then the (vectorial) sum of the deflections due to individual lenses. Consider now a three-dimensional mass distribution with volume density $\rho(\vec{r})$. We can divide it into cells of size $\mathrm{d}V$ and mass $\mathrm{d}m = \rho(\vec{r})\,\mathrm{d}V$. Let a light ray pass this mass distribution, and describe its spatial trajectory by $(\xi_1(\lambda), \xi_2(\lambda), r_3(\lambda))$, where the coordinates are chosen such that the incoming light ray (i.e. far from the deflecting mass distribution) propagates along r_3. The actual light ray is deflected, but if the deflection angle is small, it can be approximated as a straight line in the neighbourhood of the deflecting mass. This corresponds to the Born approximation in atomic and nuclear physics. Then, $\vec{\xi}(\lambda) \equiv \vec{\xi}$, independent of the affine parameter λ. Note that $\vec{\xi} = (\xi_1, \xi_2)$ is a two-dimensional vector. The impact vector of the light ray relative to the mass element $\mathrm{d}m$ at $\vec{r} = (\xi_1', \xi_2', r_3')$ is then $\vec{\xi} - \vec{\xi}'$, independent of r_3', and the total deflection angle is

$$\begin{aligned} \hat{\vec{\alpha}}(\vec{\xi}) &= \frac{4G}{c^2} \sum \mathrm{d}m(\xi_1', \xi_2', r_3') \frac{\vec{\xi} - \vec{\xi}'}{|\vec{\xi} - \vec{\xi}'|^2} \\ &= \frac{4G}{c^2} \int \mathrm{d}^2\xi' \int \mathrm{d}r_3'\, \rho(\xi_1', \xi_2', r_3') \frac{\vec{\xi} - \vec{\xi}'}{|\vec{\xi} - \vec{\xi}'|^2} , \qquad (3.2) \end{aligned}$$

which is also a two-dimensional vector. Since the last factor in eq. (3.2) is independent of r_3', the r_3'-integration can be carried out by defining the *surface mass density*

$$\Sigma(\vec{\xi}) \equiv \int \mathrm{d}r_3\, \rho(\xi_1, \xi_2, r_3) , \qquad (3.3)$$

which is the mass density projected onto a plane perpendicular to the incoming light ray. Then, the deflection angle finally becomes

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi'\, \Sigma(\vec{\xi}') \frac{\vec{\xi} - \vec{\xi}'}{|\vec{\xi} - \vec{\xi}'|^2} . \qquad (3.4)$$

This expression is valid as long as the deviation of the actual light ray from a straight (undeflected) line within the mass distribution is small compared to the scale on which the mass distribution

[4]This condition is very well satisfied in most astrophysical situations. A cluster of galaxies, for instance, has a typical size of a few Mpc, whereas the distances D_{d}, D_{s}, and D_{ds} are fair fractions of the Hubble length $cH_0^{-1} = 3\,h^{-1} \times 10^3$ Mpc.

changes significantly. This condition is satisfied in virtually all astrophysically relevant situations (i.e. lensing by galaxies and clusters of galaxies), unless the deflecting mass extends all the way from the source to the observer (a case which will be dealt with in Sect. 6.). It should also be noted that in a lensing situation such as displayed in Fig. 11, the incoming light rays are not mutually parallel, but fall within a beam with opening angle approximately equal to the angle which the mass distribution subtends on the sky. This angle, however, is typically *very* small (in the case of cluster lensing, the relevant angular scales are of order 1 arc min $\approx 2.9 \times 10^{-4}$).

3.1.2. The Lens Equation

We now require an equation which relates the true position of the source to its observed position on the sky. As sketched in Fig. 11, the source and lens planes are defined as planes perpendicular to a straight line (the optical axis) from the observer to the lens at the distance of the source and of the lens, respectively. The exact definition of the optical axis does not matter because of the smallness of angles involved in a typical lens situation, and the distance to the lens is well defined for a geometrically-thin matter distribution. Let $\vec{\eta}$ denote the two-dimensional position of the source on the source plane. Recalling the definition of the angular-diameter distance, we can read off Fig. 11

$$\vec{\eta} = \frac{D_\mathrm{s}}{D_\mathrm{d}}\vec{\xi} - D_\mathrm{ds}\hat{\vec{\alpha}}(\vec{\xi}) \; . \tag{3.5}$$

Introducing angular coordinates by $\vec{\eta} = D_\mathrm{s}\vec{\beta}$ and $\vec{\xi} = D_\mathrm{d}\vec{\theta}$, we can transform eq. (3.5) to

$$\vec{\beta} = \vec{\theta} - \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\vec{\alpha}}(D_\mathrm{d}\vec{\theta}) \equiv \vec{\theta} - \vec{\alpha}(\vec{\theta}) \; , \tag{3.6}$$

where we defined the scaled deflection angle $\vec{\alpha}(\vec{\theta})$ in the last step. The interpretation of the lens equation (3.6) is that a source with true position $\vec{\beta}$ can be seen by an observer at angular positions $\vec{\theta}$ satisfying (3.6). If (3.6) has more than one solution for fixed $\vec{\beta}$, a source at $\vec{\beta}$ has images at several positions on the sky, i.e. the lens produces multiple images. For this to happen, the lens must be 'strong'. This can be quantified by the dimension-less surface mass density

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_\mathrm{d}\vec{\theta})}{\Sigma_\mathrm{cr}} \quad \text{with} \quad \Sigma_\mathrm{cr} = \frac{c^2}{4\pi G}\frac{D_\mathrm{s}}{D_\mathrm{d}D_\mathrm{ds}} \; , \tag{3.7}$$

where Σ_cr is called the critical surface mass density (which depends on the redshifts of source and lens). A mass distribution which has $\kappa \geq 1$ somewhere, i.e. $\Sigma \geq \Sigma_\mathrm{cr}$, produces multiple images for some source positions $\vec{\beta}$ (see Schneider et al. 1992, Sect. 5.4.3). Hence, Σ_cr is a characteristic value for the surface mass density which distinguishes between 'weak' and 'strong' lenses. Note that $\kappa \geq 1$ is sufficient but not necessary for producing multiple images. In terms of κ, the scaled deflection angle reads

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta'\,\kappa(\vec{\theta}')\,\frac{\vec{\theta}-\vec{\theta}'}{|\vec{\theta}-\vec{\theta}'|^2} \; . \tag{3.8}$$

Equation (3.8) implies that the deflection angle can be written as the gradient of the *deflection potential*,

$$\psi(\vec{\theta}) = \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta'\,\kappa(\vec{\theta}')\,\ln|\vec{\theta}-\vec{\theta}'| \; , \tag{3.9}$$

as $\vec{\alpha} = \nabla\psi$. The potential $\psi(\vec{\theta})$ is the two-dimensional analogue of the Newtonian gravitational potential and satisfies the Poisson equation $\nabla^2\psi(\vec{\theta}) = 2\kappa(\vec{\theta})$.

3.1.3. Magnification and Distortion

The solutions $\vec{\theta}$ of the lens equation yield the angular positions of the images of a source at $\vec{\beta}$. The shapes of the images will differ from the shape of the source because light bundles are deflected differentially. The most visible consequence of this distortion is the occurrence of giant luminous arcs in galaxy clusters. In general, the shape of the images must be determined by solving the lens equation for all points within an extended source. Liouville's theorem and the absence of emission and absorption of photons in gravitational light deflection imply that lensing conserves surface brightness (or specific intensity). Hence, if $I^{(s)}(\vec{\beta})$ is the surface brightness distribution in the source plane, the observed surface brightness distribution in the lens plane is

$$I(\vec{\theta}) = I^{(s)}[\vec{\beta}(\vec{\theta})] \; . \tag{3.10}$$

If a source is much smaller than the angular scale on which the lens properties change, the lens mapping can locally be linearised. The distortion of images is then described by the Jacobian matrix

$$\mathcal{A}(\vec{\theta}) = \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = \left(\delta_{ij} - \frac{\partial^2\psi(\vec{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{3.11}$$

where we have introduced the components of the shear $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2 = |\gamma|\mathrm{e}^{2\mathrm{i}\varphi}$,

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) \; , \quad \gamma_2 = \psi_{,12} \; , \tag{3.12}$$

and κ is related to ψ through Poisson's equation. Hence, if $\vec{\theta}_0$ is a point within an image, corresponding to the point $\vec{\beta}_0 = \vec{\beta}(\vec{\theta}_0)$ within the source, we find from (3.10) using the locally linearised lens equation

$$I(\vec{\theta}) = I^{(s)}\left[\vec{\beta}_0 + \mathcal{A}(\vec{\theta}_0)\cdot(\vec{\theta}-\vec{\theta}_0)\right] \; . \tag{3.13}$$

According to this equation, the images of a circular source are ellipses. The ratios of the semi-axes of such an ellipse to the radius of the source are given by the inverse of the eigenvalues of $\mathcal{A}(\vec{\theta}_0)$, which are $1-\kappa \pm |\gamma|$, and the ratio of the solid angles subtended by an image and the unlensed source is the inverse of the determinant of $\mathcal{A}$. The fluxes observed from the image and from the unlensed source are given as integrals over the brightness distributions $I(\vec{\theta})$ and $I^{(s)}(\vec{\beta})$, respectively, and their ratio is the *magnification* $\mu(\vec{\theta}_0)$. From (3.13), we find

$$\mu = \frac{1}{\det\mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \; . \tag{3.14}$$

The images are thus distorted in shape and size. The shape distortion is due to the tidal gravitational field, described by the shear γ, whereas the magnification is caused by both isotropic focusing caused by the local matter density κ and anisotropic focusing caused by shear.

Since the shear is defined by the trace-free part of the symmetric Jacobian matrix $\mathcal{A}$, it has two independent components. There exists a one-to-one mapping from symmetric, trace-free 2×2 matrices onto complex numbers, and we shall extensively use complex notation. Note that the shear transforms as $\mathrm{e}^{2\mathrm{i}\varphi}$ under rotations of the coordinate frame, and is therefore not a vector. Equations (3.9) and (3.12) imply that the complex shear can be written

$$\gamma(\vec\theta) = \frac{1}{\pi}\int_{\mathbb{R}^2} d^2\theta'\,\mathcal{D}(\vec\theta-\vec\theta')\,\kappa(\vec\theta')\,,$$
$$\text{with}\quad \mathcal{D}(\vec\theta) \equiv \frac{\theta_2^2-\theta_1^2-2i\theta_1\theta_2}{|\vec\theta|^4} = \frac{-1}{(\theta_1-i\theta_2)^2} \quad (3.15)$$

3.1.4. *Critical Curves and Caustics*

Points in the lens plane where the Jacobian $\mathcal{A}$ is singular, i.e. where $\det\mathcal{A}=0$, form closed curves, the *critical curves*. Their image curves in the source plane are called *caustics*. Equation (3.14) predicts that sources on caustics are infinitely magnified; however, infinite magnification does not occur in reality, for two reasons. First, each astrophysical source is extended, and its magnification (given by the surface brightness-weighted point-source magnification across its solid angle) remains finite. Second, even point sources would be magnified by a finite value since for them, the geometrical-optics approximation fails near critical curves, and a wave-optics description leads to a finite magnification (e.g. Ohanian 1983; Schneider et al. 1992, Chap. 7). For the purposes of this review, the first effect always dominates. Nevertheless, images near critical curves can be magnified and distorted substantially, as is demonstrated by the giant luminous arcs which are formed from source galaxies close to caustics. (Point) sources which move across a caustic have their number of images changed by ±2, and the two additional images appear or disappear at the corresponding critical curve in the lens plane. Hence, only sources inside a caustic are multiply imaged.

3.1.5. *An Illustrative Example: Isothermal Spheres*

The rotation curves of spiral galaxies are observed to be approximately flat out to the largest radii where they can be measured. If the mass distribution in a spiral galaxy followed the light distribution, the rotation curves would have to decrease at large radii in roughly Keplerian fashion. Flat rotation curves thus provide the clearest evidence for dark matter on galactic scales. They can be understood if galactic disks are embedded in a dark halo with density profile $\rho\propto r^{-2}$ for large r. The projected mass density then behaves like θ^{-1}. Such density profiles are obtained by assuming that the velocity dispersion of the dark matter particles is spatially constant. They are therefore also called isothermal profiles. We shall describe some simple properties of a gravitational lens with an isothermal mass profile, which shall later serve as a reference.

The projected surface mass density of a *singular isothermal sphere* is

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi}\,, \quad (3.16)$$

where σ_v is the line-of-sight velocity dispersion of the 'particles' (e.g. stars in galaxies, or galaxies in clusters of galaxies) in the gravitational potential of the mass distribution, assuming that they are in virial equilibrium. The corresponding dimensionless surface mass density is

$$\kappa(\theta) = \frac{\theta_E}{2\theta}\,,\quad\text{where}\quad \theta_E = 4\pi\left(\frac{\sigma_v}{c}\right)^2\frac{D_{ds}}{D_s} \quad (3.17)$$

is called the *Einstein deflection angle*. As can easily be verified from (3.8), the magnitude of the scaled deflection angle is constant for this mass profile, $|\vec\alpha| = \theta_E$, and the deflection potential is $\psi = \theta_E|\vec\theta|$. From that, the shear is obtained using (3.12)[5],

$$\gamma(\vec\theta) = -\frac{\theta_E}{2|\vec\theta|}e^{2i\varphi}\,, \quad (3.18)$$

and the magnification is

$$\mu(\vec\theta) = \frac{|\vec\theta|}{|\vec\theta|-\theta_E}\,. \quad (3.19)$$

This shows that $|\vec\theta| = \theta_E$ defines a critical curve, which is called the *Einstein circle*. The corresponding caustic, obtained by mapping the Einstein circle back into the source plane under the lens equation, degenerates to a single point at $\vec\beta=\vec0$. Such degenerate caustics require highly symmetric lenses. Any perturbation of the mass distribution breaks the degeneracy and expands the singular caustic point into a caustic curve (see Chapter 6 in Schneider et al. 1992 for a detailed treatment of critical curves and caustics). The lens (3.17) produces two images with angular separation $2\theta_E$ for a source with $|\vec\beta|<1$, and one image otherwise.

The mass distribution (3.17) has two unsatisfactory properties. The surface mass density diverges for $|\vec\theta|\to0$, and the total mass of the lens is infinite. Clearly, both of these properties will not match real mass distributions. Despite this fact, the singular isothermal sphere fits many of the observed lens systems fairly well. In order to construct a somewhat more realistic lens model, one can cut off the distribution at small and large distances, e.g. by

$$\kappa(\vec\theta) = \frac{\theta_E}{2\sqrt{|\vec\theta|^2+\theta_c^2}} - \frac{\theta_E}{2\sqrt{|\vec\theta|^2+\theta_t^2}}\,, \quad (3.20)$$

which has a core radius θ_c, and a truncation radius θ_t. For $\theta_c \ll |\vec\theta| \ll \theta_t$, this mass distribution behaves like θ^{-1}. This lens can produce three images, but only if $\theta_c\theta_t\,(\theta_c+\theta_t)^{-1} < \theta_E/2$. One of the three images occurs near the centre of the lens and is strongly de-magnified if $\theta_c \ll \theta_E$. In most of the multiple-image QSO lens systems, there is no indication for a third central image, imposing strict upper bounds on θ_c, whereas for some arc systems in clusters, a finite core size is required when a lens model like (3.20) is assumed.

3.2. *Light Propagation in Arbitrary Spacetimes*

We now turn to a more rigorous description of the propagation of light rays, based on the theory of geometrical optics in General Relativity. We then specialise the resulting propagation equations to the case of weak gravitational fields and metric perturbations to the background of an expanding universe. These equations contain the gravitational lens equation discussed previously as a special case. We shall keep the discussion brief and follow closely the work of Schneider et al. (1992, Chaps. 3 & 4), and Seitz et al. (1994), where further references can be found.

3.2.1. *Propagation of Light Bundles*

In Sect. 3.1.2., we have derived the lens equation (3.5) in a heuristic way. A rigorous derivation in an arbitrary spacetime

[5]For axially-symmetric projected mass profiles, the magnitude of the shear can be calculated from $|\gamma|(\theta) = \bar\kappa(\theta)-\kappa(\theta)$, where $\bar\kappa(\theta)$ is the mean surface mass density inside a circle of radius θ from the lens centre. Accordingly, the magnitude of the deflection angle is $|\vec\alpha| = \theta\bar\kappa(\theta)$.

must account for the fact that distance vectors between null geodesics are four-vectors. Nevertheless, by choosing an appropriate coordinate system, the separation transverse to the line-of-sight between two neighbouring light rays can effectively be described by a two-dimensional vector $\vec{\xi}$. We outline this operation in the following two paragraphs.

We first consider the propagation of infinitesimally thin light beams in an arbitrary space-time, characterised by the metric tensor $g_{\mu\nu}$. The propagation of a fiducial ray γ_0 of the bundle is determined by the geodesic equation (e.g. Misner et al. 1973; Weinberg 1972). We are interested here in the evolution of the shape of the bundle as a function of the affine parameter along the fiducial ray. Consider an observer O with four-velocity $U_{\rm o}^{\mu}$, satisfying $U_{\rm o}^{\mu} U_{{\rm o}\mu} = 1$. The physical wave vector k^{μ} of a photon depends on the light frequency. We define $\tilde{k}^{\mu} \equiv -c^{-1}\omega_{\rm o} k^{\mu}$ as a past-directed dimensionless wave vector which is independent of the frequency $\omega_{\rm o}$ measured by the observer. We choose an affine parameter λ of the rays passing through O such that (1) $\lambda = 0$ at the observer, (2) λ increases along the backward light cone of O, and (3) $U_{\rm o}^{\mu}\tilde{k}_{\mu} = -1$ at O. Then, with the definition of $\tilde{k}^{\mu}$, it follows that $\tilde{k}^{\mu} = {\rm d}x^{\mu}/{\rm d}\lambda$, and that λ measures the proper distance along light rays for events close to O.

Let $\gamma^{\mu}(\vec{\theta},\lambda)$ characterise the rays of a light beam with vertex at O, such that $\vec{\theta}$ is the angle between a ray and the fiducial ray with $\gamma_0^{\mu}(\lambda) \equiv \gamma^{\mu}(\vec{0},\lambda)$. Further, let $Y^{\mu}(\vec{\theta},\lambda) = \gamma^{\mu}(\vec{\theta},\lambda) - \gamma^{\mu}(\vec{0},\lambda) = [\partial\gamma^{\mu}(\vec{\theta},\lambda)/\partial\theta_k]\theta_k$ denote the vector connecting the ray characterised by $\vec{\theta}$ with the fiducial ray at the same affine parameter λ, where we assumed sufficiently small $|\vec{\theta}|$ so that Y^{μ} can be linearised in $\vec{\theta}$. We can then decompose Y^{μ} as follows. At O, the vectors $U_{\rm o}^{\mu}$ and $\tilde{k}^{\mu}$ define a two-dimensional plane perpendicular to both $U_{\rm o}^{\mu}$ and $\tilde{k}^{\mu}$. This plane is tangent to the sphere of directions seen by the observer. Now choose orthonormal unit vectors E_1 and E_2 to span that plane. Hence, $E_1^{\mu}E_{2\mu} = 0$, $E_k^{\mu}E_{k\mu} = -1$, $E_k^{\mu}\tilde{k}_{\mu} = E_k^{\mu}U_{{\rm o}\mu} = 0$, for $k = 1,2$. Transporting the four vectors $\tilde{k}^{\mu}$, $U_{\rm o}^{\mu}$, E_1^{μ}, and E_2^{μ} parallel along the fiducial ray defines a *vierbein* at each event along the fiducial ray. The deviation vector can then be decomposed into

$$Y^{\mu}(\vec{\theta},\lambda) = -\xi_1(\vec{\theta},\lambda)\,E_1^{\mu} - \xi_2(\vec{\theta},\lambda)\,E_2^{\mu} - \xi_0(\vec{\theta},\lambda)\,\tilde{k}^{\mu}\;. \quad (3.21)$$

Thus, the two-dimensional vector $\vec{\xi}(\vec{\theta},\lambda)$ with components $\xi_{1,2}(\vec{\theta},\lambda)$ describes the transverse separation of two light rays at affine parameter λ, whereas ξ_0 allows for a deviation component along the beam direction. Due to the linearisation introduced above, $\vec{\xi}$ depends linearly on $\vec{\theta}$, and the choice of λ assures that ${\rm d}\vec{\xi}/{\rm d}\lambda(\lambda = 0) = \vec{\theta}$. Hence, we can write the linear propagation equation

$$\vec{\xi}(\lambda) = \mathcal{D}(\lambda)\,\vec{\theta}\;. \quad (3.22)$$

The 2×2 matrix $\mathcal{D}$ satisfies the Jacobi differential equation

$$\frac{{\rm d}^2\mathcal{D}(\lambda)}{{\rm d}\lambda^2} = \mathcal{T}(\lambda)\,\mathcal{D}(\lambda)\;, \quad (3.23)$$

with initial conditions

$$\mathcal{D}(0) = \mathcal{O} \quad \text{and} \quad \frac{{\rm d}\mathcal{D}}{{\rm d}\lambda}(0) = \mathcal{I}\;. \quad (3.24)$$

The *optical tidal matrix* $\mathcal{T}(\lambda)$ is symmetric,

$$\mathcal{T}(\lambda) = \begin{pmatrix} \mathcal{R}(\lambda) + \Re[\mathcal{F}(\lambda)] & \Im[\mathcal{F}(\lambda)] \\ \Im[\mathcal{F}(\lambda)] & \mathcal{R}(\lambda) - \Re[\mathcal{F}(\lambda)] \end{pmatrix}\;, \quad (3.25)$$

and its components depend on the curvature of the metric. $\Re(z)$ and $\Im(z)$ denote the real and imaginary parts of the complex number z. Specifically,

$$\mathcal{R}(\lambda) = -\frac{1}{2}R_{\mu\nu}(\lambda)\tilde{k}^{\mu}(\lambda)\tilde{k}^{\nu}(\lambda)\;, \quad (3.26)$$

where $R_{\mu\nu}(\lambda)$ is the Ricci tensor at $\gamma_0^{\mu}(\lambda)$. The complex quantity $\mathcal{F}(\lambda)$ is more complicated and depends on the Weyl curvature tensor at $\gamma_0^{\mu}\lambda)$. The *source of convergence* $\mathcal{R}(\lambda)$ leads to an isotropic focusing of light bundles, in that a circular light beam continues to have a circular cross section. In contrast, a non-zero *source of shear* $\mathcal{F}(\lambda)$ causes an anisotropic focusing, changing the shape of the light bundle. For a similar set of equations, see, e.g. Blandford et al. (1991) and Peebles (1993).

To summarise this subsection, the transverse separation vector $\vec{\xi}$ of two infinitesimally close light rays, enclosing an angle $\vec{\theta}$ at the observer, depends linearly on $\vec{\theta}$. The matrix which describes this linear mapping is obtained from the Jacobi differential equation (3.23). The optical tidal matrix $\mathcal{T}$ can be calculated from the metric. This exact result from General Relativity is of course not easily applied to practical calculations in general space-times, as one first has to calculate the null geodesic $\gamma_0^{\mu}(\lambda)$, and from that the components of the tidal matrix have to be determined. However, as we shall show next, the equations attain rather simple forms in the case of weak gravitational fields.

3.2.2. *Specialisation to Weak Gravitational Fields*

We shall now specialise the transport equation (3.23) to the situation of a homogeneous and isotropic universe, and to weak gravitational fields. In a metric of the Robertson-Walker type (2.2, page 4), the source of shear $\mathcal{F}$ must vanish identically because of isotropy; otherwise preferred directions would exist. Initially circular light bundles therefore remain circular. Hence, the optical tidal matrix $\mathcal{T}$ is proportional to the unit matrix, $\mathcal{T}(\lambda) = \mathcal{R}(\lambda)\,\mathcal{I}$, and the solution of (3.23) must be of the form $\mathcal{D}(\lambda) = D(\lambda)\,\mathcal{I}$. According to (3.22), the function $D(\lambda)$ is the angular-diameter distance as a function of the affine parameter. As we shall demonstrate next, this function indeed agrees with the angular diameter distance as defined in (2.43, page 7).

To do so, we first have to find $\mathcal{R}(\lambda)$. The Ricci tensor deviates from the Einstein tensor by two terms proportional to the metric tensor $g_{\mu\nu}$, one involving the Ricci scalar, the other containing the cosmological constant. These two terms do not contribute to (3.26), since $\tilde{k}^{\mu}$ is a null vector. We can thus replace the Ricci tensor in (3.26) by the energy-momentum tensor according to Einstein's field equation. Since $k^0 = c^{-1}\omega = (1+z)c^{-1}\omega_{\rm o}$, we have $\tilde{k}^0 = -(1+z)$, and the spatial components of $\tilde{k}^{\mu}$ are described by a direction and the constraint that $\tilde{k}^{\mu}$ is a null vector. Then, using the energy-momentum tensor of a perfect fluid with density ρ and pressure p, (3.26) becomes

$$\mathcal{R}(\lambda) = -\frac{4\pi G}{c^2}\left(\rho + \frac{p}{c^2}\right)(1+z)^2\;. \quad (3.27)$$

Specialising to a universe filled with dust, i.e. $p = 0$, we find from (2.16, page 5) and (2.19, page 5)

$$\mathcal{R}(\lambda) = -\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_0\,(1+z)^5\;. \quad (3.28)$$

The transport equation (3.23) then transforms to

$$\frac{{\rm d}^2 D}{{\rm d}\lambda^2} = -\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_0\,(1+z)^5 D\;. \quad (3.29)$$

In order to show that the solution of (3.29) with initial conditions $D = 0$ and $\mathrm{d}D = \mathrm{d}\lambda$ at $\lambda = 0$ is equivalent to (2.43, page 7), we proceed as follows. First we note that (2.43) for $z_1 = 0$ can be written as an initial-value problem,

$$\frac{\mathrm{d}^2}{\mathrm{d}w^2}\left(\frac{D_{\rm ang}}{a}\right) = -K\left(\frac{D_{\rm ang}}{a}\right) , \tag{3.30}$$

with $D_{\rm ang}(0) = 0$ and $\mathrm{d}D_{\rm ang} = \mathrm{d}w$ at $w = 0$, because of the properties of the function f_K; cf. (2.4, page 4). Next, we need a relation between λ and w. The null component of the photon geodesic is $x^0 = c(t_0 - t)$. Then, from $\mathrm{d}x^\mu = \tilde{k}^\mu \mathrm{d}\lambda$, we obtain $\mathrm{d}\lambda = -ac\mathrm{d}t$. Using $\mathrm{d}t = \dot{a}^{-1}\mathrm{d}a$, we find

$$\mathrm{d}a = -\frac{\dot{a}}{ca}\mathrm{d}\lambda\,, \quad \text{or} \quad \mathrm{d}z = \frac{\dot{a}}{ca^3}\mathrm{d}\lambda\,. \tag{3.31}$$

Since $c\mathrm{d}t = -a\mathrm{d}w$ for null rays, we have $\dot{a}^{-1}\mathrm{d}a = \mathrm{d}t = -ac^{-1}\mathrm{d}w$, which can be combined with (3.31) to yield

$$\mathrm{d}\lambda = a^2\,\mathrm{d}w\,. \tag{3.32}$$

We can now calculate the analogous expression of (3.30) for D,

$$\frac{\mathrm{d}^2}{\mathrm{d}w^2}\left(\frac{D}{a}\right) = a^2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[a^2\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{D}{a}\right)\right] = a^3 D'' - a^2 a'' D\,, \tag{3.33}$$

where a prime denotes differentiation with respect to λ. From (3.31), $a' = -(ac)^{-1}\dot{a}$, and

$$a'' = \frac{1}{2}\frac{\mathrm{d}(a')^2}{\mathrm{d}a} = \frac{1}{2c^2}\frac{\mathrm{d}}{\mathrm{d}a}\left(\frac{\dot{a}^2}{a^2}\right) = \frac{1}{2c^2}\frac{\mathrm{d}H^2}{\mathrm{d}a}\,, \tag{3.34}$$

with H given in (2.31, page 6). Substituting (3.29) into the first term on the right-hand side of (3.33), and (3.34) into the second term, we immediately see that D satisfies the differential equation (3.30). Since D has the same initial conditions as $D_{\rm ang}$, they indeed agree.

For computational convenience, we can also transform (3.29) into a differential equation for $D(z)$. Using (3.31) and (2.31), one finds

$$\begin{aligned} 0 \;=\;& (1+z)\left[(1+\Omega_0 z) - \Omega_\Lambda\left(1 - \frac{1}{(1+z)^2}\right)\right]\frac{\mathrm{d}^2D}{\mathrm{d}z^2} \\ +\;& \left[\frac{7}{2}\Omega_0 z + \frac{\Omega_0}{2} + 3 - \Omega_\Lambda\left(3 - \frac{2}{(1+z)^2}\right)\right]\frac{\mathrm{d}D}{\mathrm{d}z} \\ +\;& \frac{3}{2}\Omega_0 D\,. \end{aligned} \tag{3.35}$$

We next turn to the case of a weak isolated mass inhomogeneity with a spatial extent small compared to the Hubble distance cH_0^{-1}, like galaxies or clusters of galaxies. In that case, the metric can locally be approximated by the post-Minkowskian line element

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right)\mathrm{d}x^2\,, \tag{3.36}$$

where $\mathrm{d}x^2$ is the line element of Euclidean three-space, and Φ is the Newtonian gravitational potential which is assumed to be weak, $\Phi \ll c^2$. Calculating the curvature tensor of the metric (3.36), and using Poisson's equation for Φ, we find that for a light ray which propagates into the three-direction, the sources of convergence and shear are

$$\mathcal{R} = -\frac{4\pi G}{c^2}\rho\,, \quad \text{and} \quad \mathcal{F} = -\frac{1}{c^2}\left(\Phi_{,11} - \Phi_{,22} + 2\mathrm{i}\Phi_{,12}\right)\,. \tag{3.37}$$

Now the question is raised as to how an isolated inhomogeneity can be combined with the background model of an expanding universe. There is no exact solution of Einstein's field equations which describes a universe with density fluctuations, with the exception of a few very special cases such as the Swiss-Cheese model (Einstein & Strauss 1945). We therefore have to resort to approximation methods which start from identifying 'small' parameters of the problem, and expanding the relevant quantities into a Taylor series in these parameters. If the length scales of density inhomogeneities are much smaller than the Hubble length cH_0^{-1}, the associated Newtonian gravitational potential $\Phi \ll c^2$ (note that this does not imply that the relative density fluctuations are small!), and the peculiar velocities $v \ll c$, then an approximate metric is

$$\begin{aligned} \mathrm{d}s^2 \;=\;& a^2(\tau)\left[\left(1 + \frac{2\Phi}{c^2}\right)c^2\mathrm{d}\tau^2\right. \\ -\;& \left.\left(1 - \frac{2\Phi}{c^2}\right)\left(\mathrm{d}w^2 + f_K^2(w)\mathrm{d}\omega^2\right)\right]\,, \end{aligned} \tag{3.38}$$

where $\mathrm{d}\tau = a^{-1}\mathrm{d}t$ is the conformal time element, and Φ satisfies Poisson's equation with source $\Delta\rho$, the density enhancement or reduction relative to the mean cosmic density (Futamase 1989; Futamase & Sasaki 1989; Jacobs et al. 1993).

In the case of weak metric perturbations, the sources of convergence and shear of the background metric and the perturbations can be added. Recalling that both $\mathcal{R}$ and $\mathcal{F}$ are quadratic in $\tilde{k}^\mu \propto (1+z)$, so that the expressions in (3.37) have to be multiplied by $(1+z)^2$, we find for the optical tidal matrix

$$\begin{aligned} \mathcal{T}_{ij}(\lambda) \;=\;& -\frac{3}{2}\left(\frac{H_0}{c}\right)^2\Omega_0\,(1+z)^5\,\delta_{ij} \\ -\;& \frac{(1+z)^2}{c^2}\left(2\Phi_{,ij} + \delta_{ij}\Phi_{,33}\right)\,, \end{aligned} \tag{3.39}$$

where we have assumed that the local Cartesian coordinates are chosen such that the light ray propagates in x_3-direction. The same result is obtained from the metric (3.38).

The lens equation as discussed in Sect. 3.1. can now be derived from the previous relations. To do so, one has to assume a geometrically thin matter distribution, i.e. one approximates the density perturbation $\Delta\rho$ by a distribution which is infinitely thin in the direction of photon propagation. It is then characterised by its surface mass density $\Sigma(\vec{\xi})$. The corresponding Newtonian potential Φ can then be inserted into (3.39). The integration over $\Phi_{,33}$ along the light ray vanishes, and (3.23) can be employed to calculate the change of $\mathrm{d}\mathcal{D}/\mathrm{d}\lambda$ across the thin matter sheet (the lens plane), whereas the components of $\mathcal{D}$ far from the lens plane are given by a linear combination of solutions of the transport equation (3.29). Continuity and the change of derivative at $\lambda_{\rm d}$, corresponding to the lens redshift $z_{\rm d}$, then uniquely fix the solution. If $\mathcal{D}(\vec{\theta},\lambda_{\rm s})$ denotes the solution at redshift $z_{\rm s}$, then $\mathcal{D}(\vec{\theta},\lambda_{\rm s}) = \partial\vec{\eta}/\partial\vec{\theta}$ in the notation of Sect. 3.1.. Line integration of this relation then leads to the lens equation (3.2). See Seitz et al. (1994) for details, and Pyne & Birkinshaw (1996) for an alternative derivation.

4. Principles of Weak Gravitational Lensing

4.1. Introduction

If the faint, and presumably distant, galaxy population is observed through the gravitational field of a deflector, the appearance of the galaxies is changed. The tidal component of the gravitational field distorts the *shapes* of galaxy images, and the magnification associated with gravitational light deflection changes their apparent *brightness*. If all galaxies were intrinsically circular, any galaxy image would immediately provide information on the local tidal gravitational field. With galaxies being intrinsically elliptical, the extraction of significant information from individual images is impossible, except for *giant luminous arcs* (see Fig. 10, page 13, for an example) whose distortion is so extreme that it can easily be determined.

However, assuming that the galaxies are intrinsically randomly oriented[6], the strength of the tidal gravitational field can be inferred from a sample of galaxy images, provided its net ellipticity surmounts the Poisson noise caused by the finite number of galaxy images in the sample and by the intrinsic ellipticity distribution.

Since lensing conserves surface brightness, magnification increases the size of galaxy images at a fixed surface-brightness level. The resulting flux enhancement enables galaxies to be seen down to fainter intrinsic magnitudes, and consequently the local number density of galaxy images above a certain flux threshold can be altered by lensing.

In this section, we introduce the principles of weak gravitational lensing. In Sect. 4.2., we present the laws of the transformation between source and image ellipticities and sizes, and in particular we introduce a convenient definition of the *ellipticity* of irregularly-shaped objects. Sect. 4.3. focuses on the determination of the local tidal gravitational field from an ensemble of galaxy images. We derive practical estimators for the shear and compare their relative merits. The effects of magnification on the observed galaxy images are discussed in Sect. 4.4.. We derive an estimate for the detectability of a deflector from its weak-lensing imprint on galaxy-image ellipticities in Sect. 4.5., and the final subsection 4.6. is concerned with practical aspects of the measurement of galaxy ellipticities.

4.2. Galaxy Shapes and Sizes, and their Transformation

If a galaxy had elliptical isophotes, its shape and size could simply be defined in terms of axis ratio and area enclosed by a boundary isophote. However, the shapes of faint galaxies can be quite irregular and not well approximated by ellipses. In addition, observed galaxy images are given in terms of pixel brightness on CCDs. We therefore require a definition of size and shape which accounts for the irregularity of images, and which is well adapted to observational data.

Let $I(\vec\theta)$ be the surface brightness of a galaxy image at angular position $\vec\theta$. We first assume that the galaxy image is isolated, so that I can be measured to large angular separations from the centre $\bar{\vec\theta}$ of the image,

$$\bar{\vec\theta} \equiv \frac{\int d^2\theta\, q_I[I(\vec\theta)]\,\vec\theta}{\int d^2\theta\, q_I[I(\vec\theta)]}\;, \tag{4.1}$$

where $q_I(I)$ is a suitably chosen weight function. For instance, if $q_I(I) = \mathrm{H}(I - I_{\rm th})$ is the Heaviside step function, $\bar{\vec\theta}$ is the centre of the area enclosed by a limiting isophote $I = I_{\rm th}$. Alternatively, if $q_I(I) = I$, $\bar{\vec\theta}$ is the centre of light. As a third example, if $q_I(I) = I\,\mathrm{H}(I - I_{\rm th})$, $\bar{\vec\theta}$ is the centre of light within the limiting isophote $I = I_{\rm th}$. Having chosen $q_I(I)$, we define the tensor of second brightness moments,

$$Q_{ij} = \frac{\int d^2\theta\, q_I[I(\vec\theta)]\,(\theta_i - \bar\theta_i)\,(\theta_j - \bar\theta_j)}{\int d^2\theta\, q_I[I(\vec\theta)]}\;, \quad i,j \in \{1,2\}\;, \tag{4.2}$$

(e.g. Blandford et al. 1991). In writing (4.1) and (4.2), we implicitly assumed that $q_I(I)$ is chosen such that the integrals converge. We can now define the *size* of an image in terms of the two invariants of the symmetric tensor Q. For example, we can define the size by

$$\omega = \left(Q_{11}Q_{22} - Q_{12}^2\right)^{1/2}\;, \tag{4.3}$$

so that it is proportional to the solid angle enclosed by the limiting isophote if $q(I)$ is a step function. We quantify the *shape* of the image by the *complex ellipticity*

$$\chi \equiv \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22}}\;. \tag{4.4}$$

If the image has elliptical isophotes with axis ratio $r \le 1$, then $\chi = (1 - r^2)(1 + r^2)^{-1}\exp(2\mathrm{i}\vartheta)$, where the phase of χ is twice the position angle ϑ of the major axis. This definition assures that the complex ellipticity is unchanged if the galaxy image is rotated by π, for this rotation leaves an ellipse unchanged.

If we define the centre of the source $\bar{\vec\beta}$ and the tensor of second brightness moments $Q_{ij}^{(s)}$ of the source in complete analogy to that of the image, i.e. with $I(\vec\theta)$ replaced by $I^{(s)}(\vec\beta)$ in eqs. (4.1) and (4.2), and employ the conservation of surface brightness (3.10, page 17) and the linearised lens equation (3.13, page 17), we find that the tensors of second brightness moments of source and image are related through

$$Q^{(s)} = \mathcal{A}\,Q\,\mathcal{A}^T = \mathcal{A}\,Q\,\mathcal{A}\;, \tag{4.5}$$

where $\mathcal{A} \equiv \mathcal{A}(\bar{\vec\theta})$ is the Jacobian matrix of the lens equation at position $\bar{\vec\theta}$. Defining further the complex ellipticity of the source $\chi^{(s)}$ in analogy to (4.4) in terms of $Q^{(s)}$, ellipticities transform according to

$$\chi^{(s)} = \frac{\chi - 2g + g^2\chi^*}{1 + |g|^2 - 2\Re(g\chi^*)} \tag{4.6}$$

(Schneider & Seitz 1995; similar transformation formulae were previously derived by Kochanek 1990 and Miralda-Escude 1991), where the asterisk denotes complex conjugation, and g is the *reduced shear*

$$g(\vec\theta) \equiv \frac{\gamma(\vec\theta)}{1 - \kappa(\vec\theta)}\;. \tag{4.7}$$

The inverse transformation is obtained by interchanging χ and $\chi^{(s)}$ and replacing g by $-g$ in (4.6). Equation (4.6) shows that the transformation of image ellipticities depends only on the reduced shear, and not on the shear and the surface mass density individually. Hence, the reduced shear or functions thereof are the only quantities accessible through measurements of image ellipticities. This can also immediately be seen by writing $\mathcal{A}$ as

$$\mathcal{A} = (1 - \kappa)\begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix}\;. \tag{4.8}$$

[6]This assumption is not seriously challenged. Whereas galaxies in a cluster may have non-random orientations relative to the cluster centre, or pairs of galaxies may be aligned due to mutual tidal interaction, the faint galaxies used for lensing studies are distributed over a large volume enclosed by a narrow cone with opening angle selected by the angular resolution of the mass reconstruction (see below) and length comparable to the Hubble radius, since the redshift distribution of faint galaxies is fairly broad. Thus, the faint galaxies typically have large spatial separations, which is also reflected by their weak two-point angular auto-correlation (Brainerd et al. 1995; Villumsen et al. 1997).

The pre-factor $(1-\kappa)$ only affects the size, but not the shape of the images. From (4.5) and (4.3), we immediately see that the sizes of source and image are related through

$$\omega = \mu(\vec{\theta})\,\omega^{(s)} \; . \tag{4.9}$$

We point out that the dimension-less surface mass density κ, and therefore also the shear γ, depend not only on the redshift of the lens, but also on the redshift of the sources, because the critical surface mass density (3.7, page 17) involves the source redshift. More precisely, for fixed lens redshift z_d, the lens strength is proportional to the distance ratio $D_\mathrm{ds}/D_\mathrm{s}$. This implies that the transformation (4.6) generally also depends on source redshift. We shall return to these redshift effects in Sect. 4.3., and assume for now that the lens redshift z_d is sufficiently small so that the ratio $D_\mathrm{ds}/D_\mathrm{s}$ is approximately the same for all faint galaxy images.

Instead of χ, we can define different ellipticity parameters (see Bonnet & Mellier 1995). One of these definitions turns out to be quite useful, namely

$$\varepsilon \equiv \frac{Q_{11}-Q_{22}+2\mathrm{i}Q_{12}}{Q_{11}+Q_{22}+2(Q_{11}Q_{22}-Q_{12}^2)^{1/2}} \; , \tag{4.10}$$

which we shall also call *complex ellipticity*. (Since we shall use the notation χ and ε consistently throughout this article, there should be no confusion from using the same name for two different quantities.) ε has the same phase as χ, and for elliptical isophotes with axis ratio $r \le 1$, $|\varepsilon| = (1-r)(1+r)^{-1}$. ε and χ are related through

$$\varepsilon = \frac{\chi}{1+(1-|\chi|^2)^{1/2}} \; , \quad \chi = \frac{2\varepsilon}{1+|\varepsilon|^2} \; . \tag{4.11}$$

The transformation between source and image ellipticity in terms of ε is given by

$$\varepsilon^{(s)} = \begin{cases} \dfrac{\varepsilon - g}{1-g^*\varepsilon} & \text{for} \quad |g| \le 1 \\[2ex] \dfrac{1-g\varepsilon^*}{\varepsilon^*-g^*} & \text{for} \quad |g| > 1 \end{cases} \tag{4.12}$$

(Seitz & Schneider 1997), and the inverse transformation is obtained by interchanging ε and $\varepsilon^{(s)}$ and replacing g by $-g$ in (4.12). Although the transformation of ε appears more complicated because of the case distinction, we shall see in the next subsection that it is often useful to work in terms of ε rather than χ; cf. eq. (4.17) below.

We note in passing that the possible polarisation of light of faint galaxies (Audit & Simmons 1999) or faint radio sources (Surpi & Harari 1999) may offer a different channel to detect shear. The orientation of the polarisation is unchanged in weak-field light deflection (e.g., Schneider et al. 1992, Faraoni 1993). Gravitational shear will turn the geometrical image, but not the polarisation of a galaxy. If the orientation of a galaxy is intrinsically strongly correlated with the direction of the polarisation of its light, then a mismatch of the observed directions provides information on the lensing distortion. However, the polarisation properties of faint galaxies are mostly unknown, and it is unclear whether such an intrinsic polarisation-orientation correlation exists.

For the case of weak lensing, which we define for the purpose of this section by $\kappa \ll 1$, $|\gamma| \ll 1$, and thus $|g| \ll 1$, (4.12) becomes $\varepsilon \approx \varepsilon^{(s)} + g$, provided $|\varepsilon| \approx |\varepsilon^{(s)}| \lesssim 1/2$. Likewise, eq. (4.6) simplifies to $\chi \approx \chi^{(s)} + 2g$ in this case.

4.3. *Local Determination of the Distortion*

As mentioned earlier, the observed ellipticity of a single galaxy image provides only little information about the local tidal gravitational field of the deflector, for the intrinsic ellipticity of the source is unknown. However, based on the assumption that the sources are randomly oriented, information on the local tidal field can be inferred from a local ensemble of images. Consider for example galaxy images at positions $\vec{\theta}_i$ close enough to a fiducial point $\vec{\theta}$ so that the local lens properties κ and γ do not change appreciably over the region encompassing these galaxies. The expectation value of their corresponding source ellipticities is assumed to vanish,

$$\mathrm{E}(\chi^{(s)}) = 0 = \mathrm{E}(\varepsilon^{(s)}) \; . \tag{4.13}$$

4.3.1. *All Sources at the Same Redshift*

We first consider the case that all sources are at the same redshift. Then, as mentioned following eq. (3.13, page 17), the ellipticity of a circular source determines the ratio of the local eigenvalues of the Jacobian matrix $\mathcal{A}$. This also holds for the net image ellipticity of an ensemble of sources with vanishing net ellipticity. From (3.11, page 17), we find for the ratio of the eigenvalues of $\mathcal{A}$ in terms of the reduced shear g

$$r = \frac{1 \mp |g|}{1 \pm |g|} \; . \tag{4.14}$$

Interestingly, if we replace g by $1/g^*$, r switches sign, but $|r|$ and the phase of ε remain unchanged. The sign of r cannot be determined observationally, and hence measurements cannot distinguish between g and $1/g^*$. This is called *local degeneracy*. Writing $\det\mathcal{A} = (1-\kappa)^2(1-|g|^2)$, we see that the degeneracy between g and $1/g^*$ means that we cannot distinguish between observed images inside a critical curve (so that $\det\mathcal{A} < 0$ and $|g| > 1$) or outside. Therefore, only functions of g which are invariant under $g \to 1/g^*$ are accessible to (local) measurements, as for instance the *complex distortion*

$$\delta \equiv \frac{2g}{1+|g|^2} \; . \tag{4.15}$$

Replacing the expectation value in (4.13) by the average over a local ensemble of image ellipticities, $\langle\chi^{(s)}\rangle \approx \mathrm{E}(\chi^{(s)}) = 0$, Schneider & Seitz (1995) showed that $\langle\chi^{(s)}\rangle = 0$ is equivalent to

$$\sum_i u_i \, \frac{\chi_i - \delta}{1-\Re(\delta\chi_i^*)} = 0 \; , \tag{4.16}$$

where the u_i are weight factors depending on $|\vec{\theta}_i - \vec{\theta}|$ which can give larger weight to galaxies closer to the fiducial point. Additionally, the u_i can be chosen such as to account for measurement uncertainties in the image ellipticities by giving less weight to images with larger measurement error. Equation (4.16) has a unique solution δ, so that the distortion can locally be determined. It is readily solved by a quickly converging iteration starting from $\delta = \langle\chi\rangle$.

The δ obtained from (4.16) is an unbiased estimate of the distortion. Its dispersion about the true value depends on the dispersion σ_χ of the intrinsic ellipticity distribution, and on the number of galaxy images. A fairly accurate estimate of the *rms* error of δ is $\sigma_\delta \approx \sigma_\chi N^{-1/2}$, where N is the effective number of galaxies used for the local average, $N = \left(\sum u_i\right)^2 \left(\sum u_i^2\right)^{-1}$. This overestimates the error for large values of $|\delta|$ (Schneider & Seitz 1995).

It is important to note that the expectation value of χ is *not* δ, but differs from it by a factor which depends both on $|\delta|$ and the intrinsic ellipticity distribution of the sources. In contrast to that, it follows from (4.13) and (4.12) that the expectation value of the complex ellipticity ϵ of the images *is* the reduced shear or its inverse, $\mathrm{E}(\epsilon) = g$ if $|g| < 1$ and $\mathrm{E}(\epsilon) = 1/g^*$ if $|g| > 1$ (Schramm & Kayser 1995; Seitz & Schneider 1997). Hence,

$$\langle\epsilon\rangle = \frac{\sum_i u_i \epsilon_i}{\sum_i u_i} \tag{4.17}$$

is an unbiased local estimate for g or $1/g^*$. The ellipticity parameter ϵ is useful exactly because of this property. If one deals with sub-critical lenses (i.e. lenses which are not dense enough to have critical curves, so that $\det\mathcal{A}(\vec\theta) > 0$ everywhere), or with the region outside the critical curves in critical lenses, the degeneracy between g and $1/g^*$ does not occur, and $\langle\epsilon\rangle$ is a convenient estimate for the local reduced shear. The *rms* error of this estimate is approximately $\sigma_g \approx \sigma_\epsilon\,(1-|g|^2)N^{-1/2}$ (Schneider et al. 2000), where σ_ϵ is the dispersion of the intrinsic source ellipticity $\epsilon^{(s)}$. As we shall see in a moment, ϵ is the more convenient ellipticity parameter when the sources are distributed in redshift.

The estimates for δ and g discussed above can be derived without knowing the intrinsic ellipticity distribution. If, however, the intrinsic ellipticity distribution is known (e.g. from deep *Hubble Space Telescope* images), we can exploit this additional information and determine δ (or g) through a maximum-likelihood method (Gould 1995; Lombardi & Bertin 1998a). Depending on the shape of the intrinsic ellipticity distribution, this approach can yield estimates of the distortion which have a smaller *rms* error than the estimates discussed above. However, if the intrinsic ellipticity distribution is approximately Gaussian, the *rms* errors of both methods are identical. It should be noted that the intrinsic ellipticity distribution is likely to depend on the apparent magnitude of the galaxies, possibly on their redshifts, and on the wavelength at which they are observed, so that this distribution is not easily determined observationally. Knowledge of the intrinsic ellipticity distribution can also be used to determine δ from the orientation of the images (that is, the phase of χ) only (Kochanek 1990; Schneider & Seitz 1995; Deiser 1995, unpublished). This may provide a useful alternative to the method above since the orientation of images is much less affected by seeing than the modulus of χ. We return to the practical estimate of the image ellipticities and the corresponding distortion in Sect. 4.5..

In the case of weak lensing, defined by $\kappa \ll 1$ and $|\gamma| \ll 1$, implying $|g| \ll 1$, we find from (4.11–4.16) that

$$\gamma \approx g \approx \frac{\delta}{2} \approx \langle\epsilon\rangle \approx \frac{\langle\chi\rangle}{2}\;. \tag{4.18}$$

4.3.2. Sources Distributed in Redshift

So far, we assumed that all source galaxies are at the same redshift, or more precisely, that the ratio $D_{\rm ds}/D_{\rm s}$ between the lens-source and observer-source distances is the same for all sources. This ratio enters into the scaling (3.7, page 17) of the physical surface mass density Σ to the dimension-less convergence κ. The deflection angle, the deflection potential, and the shear are all linear in κ, so that the distance ratio $D_{\rm ds}/D_{\rm s}$ is sufficient to specify the lens strength as a function of source redshift. Provided $z_{\rm d} \lesssim 0.2$, this ratio is fairly constant for sources with redshift $z_{\rm s} \gtrsim 0.8$, so that the approximation used so far applies to relatively low-redshift deflectors. However, for higher-redshift lenses, the redshift distribution of the sources must explicitly be taken into account.

For a fixed lens redshift $z_{\rm d}$, the dimension-less surface mass density and the shear depend on the source redshift. We define

$$\begin{aligned} Z(z) &\equiv \frac{\lim_{z\to\infty}\Sigma_{\rm cr}(z_{\rm d},z)}{\Sigma_{\rm cr}(z_{\rm d},z)}\,\mathrm{H}(z-z_{\rm d}) \\ &= \frac{f_K[w(z_{\rm d},z)]}{f_K[w(0,z)]}\,\frac{f_K[w(0,\infty)]}{f_K[w(z_{\rm d},\infty)]}\,\mathrm{H}(z-z_{\rm d})\;, \end{aligned} \tag{4.19}$$

using the notation of Sect. 2.1. (page 3). The Heaviside step function accounts for the fact that sources closer than the deflector are not lensed. Then, $\kappa(\vec\theta,z) = Z(z)\kappa(\vec\theta)$, and $\gamma(\vec\theta,z) = Z(z)\gamma(\vec\theta)$ for a source at z, and κ and γ refer to a fictitious source at redshift infinity. The function $Z(z)$ is readily evaluated for any cosmological model using (2.41, page 7) and (2.4, page 4). We plot $Z(z)$ for various cosmologies and lens redshifts in Fig. 12.

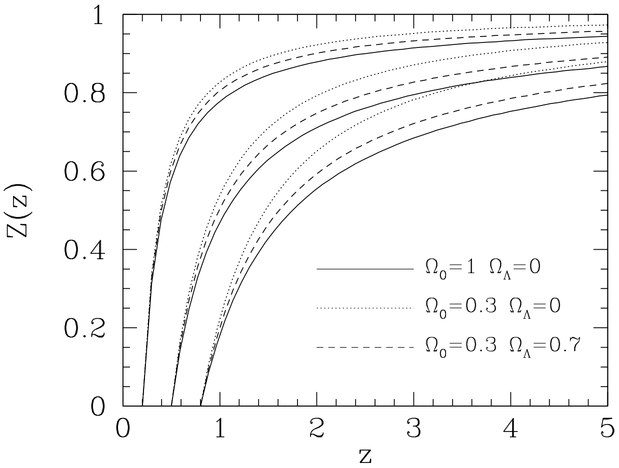

FIG. 12.—The function $Z(z)$ defined in eq. (4.19) describes the relative lens strength as a function of source redshift z. We show $Z(z)$ for three cosmological models as indicated in the figure, and for three values for the lens redshift, $z_{\rm d} = 0.2, 0.5, 0.8$. By definition, $Z(z) \to 0$ as $z \to z_{\rm d}$, and $Z(z) \to 1$ as $z \to \infty$. For sources close to the deflector, $Z(z)$ varies strongly in a way depending relatively weakly on cosmology.

The expectation value for the ellipticity of images with redshift z now becomes

$$\mathrm{E}[\epsilon(z)] = g(z) = \begin{cases} \dfrac{Z(z)\,\gamma}{1-Z(z)\,\kappa} & \text{for} \quad \mu(z) \ge 0 \\[2ex] \dfrac{1-Z(z)\,\kappa}{Z(z)\,\gamma^*} & \text{for} \quad \mu(z) < 0 \end{cases}\;, \tag{4.20}$$

where $\mu(z)$ is the magnification as a function of source redshift,

$$\mu(z) = \left\{[1-Z(z)\kappa]^2 - Z^2(z)|\gamma|^2\right\}^{-1}\;. \tag{4.21}$$

We refer to *sub-critical* lensing if $\mu(z) > 0$ for all redshifts, which is equivalent to $1-\kappa-|\gamma| > 0$.

Without redshift information, only the mean ellipticity averaged over all redshifts can be observed. We first consider this case, for which the source redshift distribution is assumed to be known. We define the probability $p_z(z)\mathrm{d}z$ that a galaxy image (in the selected magnitude range) has a redshift within $\mathrm{d}z$ of z. The image redshift distribution will in general be different from the source redshift distribution since magnified sources can be seen to higher redshifts than unlensed ones. Therefore, the redshift distribution will depend on the local lens parameters κ and

γ through the magnification (4.21). If, however, the magnification is small, or if the redshift distribution depends only weakly on the flux, the simplification of identifying the two redshift distributions is justified. We shall drop it later. Given $p_z(z)$, the expectation value of the image ellipticity becomes the weighted average

$$\mathrm{E}(\varepsilon) = \int \mathrm{d}z\, p_z(z)\, \mathrm{E}[\varepsilon(z)] = \gamma \left[X(\kappa,\gamma) + |\gamma|^{-2} Y(\kappa,\gamma) \right] , \quad (4.22)$$

with

$$\begin{aligned} X(\kappa,\gamma) &= \int_{\mu(z)\geq 0} \mathrm{d}z\, p_z(z)\, \frac{Z(z)}{1-Z(z)\kappa}, \\ Y(\kappa,\gamma) &= \int_{\mu(z)<0} \mathrm{d}z\, p_z(z)\, \frac{1-Z(z)\kappa}{Z(z)} , \end{aligned} \quad (4.23)$$

and the integration boundaries depend on the values of κ and $|\gamma|$ through the magnification.

If the lens is sub-critical, $\mu(z) > 0$ for all z. Then $Y = 0$, and only the first term in (4.22) remains. Also, X no longer depends on γ in this case, and $\mathrm{E}(\varepsilon) = \gamma X(\kappa)$. An accurate approximation for $X(\kappa)$, valid for $\kappa \lesssim 0.6$, has been derived in Seitz & Schneider (1997),

$$\gamma = \frac{\mathrm{E}(\varepsilon)}{\langle Z\rangle} \left(1 - \frac{\langle Z^2\rangle}{\langle Z\rangle} \kappa \right) , \quad (4.24)$$

where $\langle Z^n\rangle \equiv \int \mathrm{d}z\, p_z(z)\, Z^n$.

Specialising further to the weak-lensing regime, the expectation value of the image ellipticity is simply

$$\mathrm{E}(\varepsilon) \approx \langle Z\rangle \gamma . \quad (4.25)$$

Thus, in the weak-lensing case, a source redshift distribution can be collapsed on a single redshift z_s satisfying $Z(z_\mathrm{s}) = \langle Z\rangle$.

We now drop the simplification introduced above and define $n_0(S,z)\mathrm{d}S\mathrm{d}z$ as the number of galaxy images per unit solid angle with flux within $\mathrm{d}S$ of S and redshift within $\mathrm{d}z$ of z in the absence of lensing. At a point $\vec\theta$ with surface mass density κ and shear γ, the number density can be changed by magnification. Images of a fixed set of sources are distributed over a larger solid angle, reducing the number density by a factor $\mu^{-1}(z)$. On the other hand, the magnification allows the observation of fainter sources. In total, the expected number density becomes

$$n(S,z) = \frac{1}{\mu^2(z)}\, n_0 \left(\frac{S}{\mu(z)}, z \right) , \quad (4.26)$$

with $\mu(z)$ given in (4.21). This yields the redshift distribution

$$p(z;S,\kappa,\gamma) = \frac{n_0\left[\mu^{-1}(z)S,z\right]}{\mu^2(z) \int \mathrm{d}z'\, \mu^{-2}(z')\, n_0\left[\mu^{-1}(z)S,z'\right]} , \quad (4.27)$$

which depends on the flux S and the local lens parameters κ and γ through the magnification. This function can now be substituted for $p_z(z)$ in eq. (4.22).

4.3.3. Practical Estimates of the Shear

We saw before that $\langle\varepsilon\rangle = \sum_i u_i \varepsilon_i / \sum_i u_i$ is an unbiased estimate of the local reduced shear g if all sources are at the same redshift. We now generalise this result for sources distributed in redshift. Then, the expectation value of ε is no longer a simple function of κ and γ, and therefore estimates of γ for an assumed value for κ will be derived.

We first assume that redshifts for individual galaxies are unavailable, but that only the normalised redshift distribution $p_z(z)$ is known, or the distribution in eq. (4.27). Replacing the expectation value of the image ellipticity by the mean, eq. (4.22) implies that the solution $\gamma^{(1)}$ of

$$\gamma = \left[X(\kappa,\gamma) + |\gamma|^{-2} Y(\kappa,\gamma) \right]^{-1} \langle\varepsilon\rangle \quad (4.28)$$

provides an unbiased estimator for the shear γ. This is not a particularly explicit expression for the shear estimate, but it is still extremely useful, as we shall see in the next section. The shear estimate considerably simplifies if we assume a sub-critical lens. Then,

$$\gamma^{(1,\mathrm{sc})} = \langle\varepsilon\rangle X^{-1}(\kappa) \approx \frac{\langle\varepsilon\rangle}{\langle Z\rangle} \left(1 - \frac{\langle Z^2\rangle}{\langle Z\rangle}\kappa \right) , \quad (4.29)$$

where we used eq. (4.24) in the second step. Specialising further to weak lensing, the shear estimate simplifies to

$$\gamma^{(1,\mathrm{wl})} = \langle\varepsilon\rangle\, \langle Z\rangle^{-1} . \quad (4.30)$$

Next, we assume that the redshifts of all galaxy images are known. At first sight, this appears entirely unrealistic, because the galaxy images are so faint that a complete spectroscopic survey at the interesting magnitude limits seems to be out of reach. However, it has become clear in recent years that accurate redshift estimates, the so-called photometric redshifts, can be obtained from multi-colour photometry alone (see, e.g., Connolly et al. 1995). The accuracy of photometric redshifts depends on the number of wave bands for which photometry is available, the photometric accuracy, and the galaxy type; typical errors are $\Delta z \sim 0.1$ for faint, high-redshift galaxies. This uncertainty is small compared to the range over which the function $Z(z)$ varies appreciably, so that photometric redshifts are (almost) as good as precise spectroscopic redshifts for our purposes.

If the redshifts z_i of the galaxies are known, more precise shear estimates than before can be derived. Consider the weighted sum $F \equiv \sum_i u_i\, |\varepsilon_i - \mathrm{E}(\varepsilon_i)|^2$, where the expectation value is given by eq. (4.20), and $Z = Z_i \equiv Z(z_i)$. For an assumed value of κ, an unbiased estimate of γ is given by the $\gamma^{(2)}$ minimising F. Due to the case distinction in eq. (4.20), this estimator is complicated to write down analytically, but can easily be calculated numerically.

This case distinction is no longer necessary in the sub-critical case, for which the resulting estimator reads

$$\gamma^{(2,\mathrm{sc})} = \frac{\sum_i u_i\, Z_i\, \varepsilon_i\, (1-Z_i\kappa)^{-1}}{\sum_i u_i\, Z_i^2\, (1-Z_i\kappa)^{-2}} . \quad (4.31)$$

In the case of weak lensing, this becomes

$$\gamma^{(2,\mathrm{wl})} = \frac{\sum_i u_i\, Z_i\, \varepsilon_i}{\sum_i u_i\, Z_i^2} . \quad (4.32)$$

We now compare the accuracy of the shear estimates with and without redshift information of the individual galaxies. For simplicity, we assume sub-critical lensing and set all weight factors to unity, $u_i = 1$. The dispersion of the estimate $\gamma^{(1,\mathrm{sc})} = (NX)^{-1} \sum_i \varepsilon_i$ for N galaxy images is

$$\begin{aligned} \sigma^2\left(\gamma^{(1,\mathrm{sc})}\right) &= \mathrm{E}\left(|\gamma^{(1,\mathrm{sc})}|^2\right) - |\gamma|^2 \\ &= [NX(\kappa)]^{-2}\, \mathrm{E}\left(\sum_{ij} \varepsilon_i\, \varepsilon_j^* \right) - |\gamma|^2 . \end{aligned} \quad (4.33)$$

The expectation value in the final expression can be estimated noting that the image ellipticity is to first order given by $\varepsilon_i = \varepsilon_i^{(s)} + \gamma$, and that the intrinsic ellipticities are uncorrelated. If we further assume that the redshifts of any two galaxies are uncorrelated, we find

$$\begin{aligned} \mathrm{E}\left(\varepsilon_i \varepsilon_j^*\right) &\approx \left\langle \frac{Z_i Z_j}{(1-Z_i\kappa)(1-Z_j\kappa)} \right\rangle |\gamma|^2 + \delta_{ij}\sigma_\varepsilon^2 \\ &= X^2(\kappa)|\gamma|^2 + \delta_{ij}\left(\sigma_X^2|\gamma|^2 + \sigma_\varepsilon^2\right) \,, \end{aligned} \tag{4.34}$$

where we used the definition (4.23) of $X(\kappa)$, and defined $\sigma_X^2(\kappa) \equiv \langle Z^2(1-Z\kappa)^{-2}\rangle - X^2$. Angular brackets denote averages over the redshift distribution p_z. Inserting (4.34) into (4.33) yields

$$\sigma^2\left(\gamma^{(1,\mathrm{sc})}\right) = \frac{\sigma_X^2|\gamma|^2 + \sigma_\varepsilon^2}{NX^2} \,. \tag{4.35}$$

Likewise, the dispersion of the estimate $\gamma^{(2,\mathrm{sc})}$ is

$$\begin{aligned} &\sigma^2\left(\gamma^{(2,\mathrm{sc})}\right) \\ &= \frac{\sum_{ij} Z_i Z_j (1-Z_i\kappa)^{-1}(1-Z_j\kappa)^{-1}\mathrm{E}\left(\varepsilon_i\varepsilon_j^*\right)}{\left[\sum_i Z_i^2(1-Z_i\kappa)^{-2}\right]^2} - |\gamma|^2 \\ &= \frac{\sigma_\varepsilon^2}{\sum_i Z_i^2(1-Z_i\kappa)^{-2}} \approx \frac{\sigma_\varepsilon^2}{N[X^2(\kappa)+\sigma_X^2(\kappa)]} \,. \end{aligned} \tag{4.36}$$

We used eq. (4.34), but noted that Z is now no longer a statistical variable, so that we can put $\sigma_X^2 = 0$ in (4.34). In the final step, we have replaced the denominator by its expectation value under ensemble averaging. We then find the ratio of the dispersions,

$$\frac{\sigma^2\left(\gamma^{(1,\mathrm{sc})}\right)}{\sigma^2\left(\gamma^{(2,\mathrm{sc})}\right)} = \left(1 + |\gamma|^2\frac{\sigma_X^2}{\sigma_\varepsilon^2}\right)\left(1 + \frac{\sigma_X^2}{X^2}\right) \,. \tag{4.37}$$

We thus see that the relative accuracy of these two estimates depends on the fractional width of the distribution of $Z/(1-Z\kappa)$, and on the ratio between the dispersion of this quantity and the ellipticity dispersion. Through its explicit dependence on $|\gamma|^2$, and through the dependence of σ_X and X on κ, the relative accuracy also depends on the lens parameters. Quantitative estimates of (4.37) are given in Fig. 13.

The figure shows that the accuracy of the shear estimate is noticeably improved, in particular once the lens redshift becomes a fair fraction of the mean source redshift. The dependence of the lens strength on the deflector redshift implies that the lens signal will become smaller for increasing deflector redshift, so that the accuracy gained by redshift information becomes significant. In addition, the assumptions used to derive (4.35) were quite optimistic, since we have assumed in (4.34) that the sample of galaxies over which the average is taken is a fair representation of the galaxy redshift distribution $p_z(z)$. Given that these galaxies come from a small area (small enough to assume that κ and γ are constant across this area), and that the redshift distribution of observed galaxies in pencil beams shows strong correlations (see, e.g., Broadhurst et al. 1990, Steidel et al. 1998, Cohen et al. 1999), this assumption is not very realistic. Indeed, the strong clustering of galaxy redshifts means that the effective σ_X will be considerably larger than the analytical estimate used above. The noise in the local determination of the shear due to the correlated galaxy redshifts does not decrease with the number N of galaxies used, and, therefore, its relative contribution becomes more important for larger number densities of source galaxies (Schneider & Morales-Merino 2000). In any case, redshift information on the source galaxies will substantially improve the accuracy of weak lensing results.

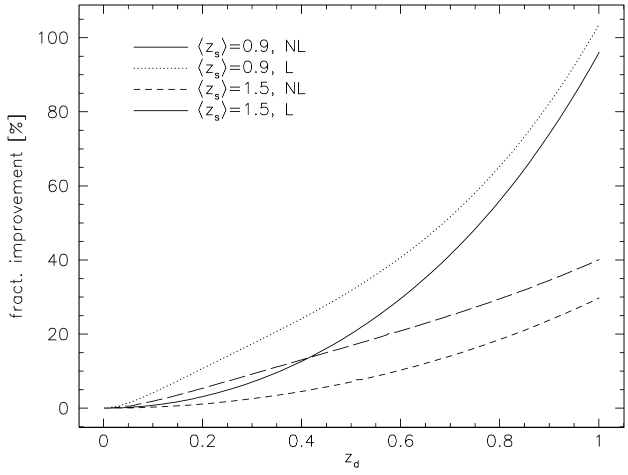

FIG. 13.—The fractional accuracy gain in the shear estimate due to the knowledge of the source redshifts is plotted, more precisely the deviation of the square root of (4.37) from unity in per cent. The four curves shown correspond to two different values of the mean source redshift, and to the cases without lensing ($\kappa = 0 = \gamma$), and with lensing ($\kappa = 0.3 = |\gamma|$), labelled NL and L, respectively. We assumed the redshift distribution (2.69) with $\beta = 3/2$, and an Einstein-de Sitter cosmology. As expected, the higher the lens redshift z_d, the more substantially is the shear estimate improved by redshift information, since for low values of z_d, the function $Z(z)$ is nearly constant. Furthermore, the lower the mean redshift of the source distribution, the more important the knowledge of individual redshifts becomes, for example to distinguish between foreground and background galaxies. Finally, redshift information is relatively more important for larger lens strength.

4.4. *Magnification Effects*

In addition to the distortion of image *shapes*, by which the (reduced) shear can be measured locally, gravitational light deflection also magnifies the images, leaving the surface brightness invariant. The magnification changes the size, and therefore the flux, of individual galaxy images. Moreover, for a fixed set of sources, the number density of images decreases by a factor μ as the sky is locally stretched. Combining the latter effect with the flux magnification, the lensed and unlensed source counts are changed according to (4.26). Two strategies to measure the magnification effect have been suggested in the literature, namely either through the change in the local source counts, perhaps combined with the associated change (4.27) in the redshift distribution (Broadhurst et al. 1995), or through the change of image sizes at fixed surface brightness (Bartelmann & Narayan 1995).

4.4.1. *Number Density Effect*

Let $n_0(>S,z)\mathrm{d}z$ be the unlensed number density of galaxies with redshift within $\mathrm{d}z$ of z and with flux larger than S. Then, at an angular position $\vec\theta$ where the magnification is $\mu(\vec\theta,z)$, the number counts are changed according to (4.26),

$$n(>S,z) = \frac{1}{\mu(\vec\theta,z)}\, n_0\left(>\frac{S}{\mu(\vec\theta,z)}, z\right) \,. \tag{4.38}$$

Accordingly, magnification can either increase or decrease the local number counts, depending on the shape of the unlensed number-count function. This change of number counts is called *magnification bias*, and is a very important effect for gravitational lensing of QSOs (see Schneider et al. 1992 for references).[7]

Magnification allows the observation of fainter sources. Since the flux from the sources is correlated with their redshift, the redshift distribution is changed accordingly,

$$p(z;>S,\kappa,\gamma) = \frac{n_0\left[>\mu^{-1}(z)S,z\right]}{\mu(z)\int \mathrm{d}z'\,\mu^{-1}(z')\,n_0\left[>\mu^{-1}(z')S,z'\right]}\;, \tag{4.39}$$

in analogy to the redshift distribution (4.27) at fixed flux S. Since the objects of interest here are very faint, spectroscopic redshift information is in general difficult to obtain, and so one can only observe the redshift-integrated counts

$$n(>S) = \int \mathrm{d}z\,\frac{1}{\mu(z)}\,n_0\left(>\mu^{-1}(z)S,z\right)\;. \tag{4.40}$$

The number counts of faint galaxies are observed to very closely follow a power law over a wide range of fluxes, and so we write the unlensed counts as

$$n_0(>S,z) = a\,S^{-\alpha}\,p_0(z;S)\;, \tag{4.41}$$

where the exponent α depends on the wave band of the observation (e.g. Smail et al. 1995a), and $p_0(z;S)$ is the redshift probability distribution of galaxies with flux $> S$. Whereas this redshift distribution is fairly well known for brighter galaxies which are accessible to current spectroscopy, little is known about the faint galaxies of interest here. The ratio of the lensed and unlensed source counts is then found by inserting (4.41) into (4.40),

$$\frac{n(>S)}{n_0(>S)} = \int \mathrm{d}z\,\mu^{\alpha-1}(z)\,p_0\left(z;\mu^{-1}(z)S\right)\;. \tag{4.42}$$

We should note that the lensed counts do not strictly follow a power law in S, for p_0 depends on z. Since the redshift distribution $p_0(z,S)$ is currently unknown, the change of the number counts due to the magnification cannot be predicted. For very faint flux thresholds, however, the redshift distribution is likely to be dominated by galaxies at relatively high redshift. For lenses at fairly small redshift (say $z_\mathrm{d} \lesssim 0.3$), we can approximate the redshift-dependent magnification $\mu(z)$ by the magnification μ of a fiducial source at infinity, in which case

$$\frac{n(>S)}{n_0(>S)} = \mu^{\alpha-1}\;. \tag{4.43}$$

Thus, a local estimate of the magnification can be obtained through (4.43) and from a measurement of the local change of the number density of images. If the slope of the source counts is unity, $\alpha = 1$, there will be no magnification bias, while it will cause a decrease of the local number density for flatter slopes. Broadhurst et al. (1995) pointed out that one can immediately obtain (for sub-critical lensing, i.e. $\det\mathcal{A} > 0$) an estimate for the local surface mass density from a measurement of the local magnification and the local reduced shear g, $\kappa = 1 - [\mu(1-|g|^2)]^{-1/2}$. In the absence of shape information, (4.43) can be used in the weak lensing limit [where $\kappa \ll 1$, $|\gamma| \ll 1$, so that $\mu \approx (1+2\kappa)$] to obtain an estimate of the surface mass density,

$$\kappa \approx \frac{n(>S)-n_0(>S)}{n_0(>S)}\,\frac{1}{2(\alpha-1)}\;. \tag{4.44}$$

4.4.2. Size Effect

Since lensing conserves surface brightness, the magnification can be obtained from the change in galaxy-image sizes at fixed surface brightness. Let I be some convenient measure of the surface brightness. For example, if ω is the solid angle of an image, defined by the determinant of the tensor of second brightness moments as in (4.3), one can set $I = S/\omega$.

Denoting by $n(\omega,I,z)\mathrm{d}\omega$ the number density of images with surface brightness I, redshift z, and solid angle within $\mathrm{d}\omega$ of ω, the relation between the lensed and the unlensed number density can be written

$$n(\omega,I,z) = \frac{1}{\mu^2}\,n_0\left(\frac{\omega}{\mu},I,z\right)\;. \tag{4.45}$$

For simplicity, we only consider the case of a moderately small lens redshift, so that the magnification can be assumed to be locally constant for all images, irrespective of galaxy redshift. We can then drop the variable z here. The mean image size $\langle\omega\rangle(I)$ at fixed surface brightness I is then related to the mean image size $\langle\omega\rangle_0(I)$ in the absence of lensing through

$$\langle\omega\rangle(I) = \mu\langle\omega\rangle_0(I)\;. \tag{4.46}$$

If the mean image size in the absence of lensing can be measured (e.g. by deep *HST* exposures of blank fields), the local value μ of the magnification can therefore be determined by comparing the observed image sizes to those in the blank fields. This method has been discussed in detail in Bartelmann & Narayan (1995). For instance, if we assume that the logarithm of the image size is distributed as a Gaussian with mean $\langle\ln\omega\rangle_0(I)$ and dispersion $\sigma(I)$, we obtain an estimate for the local magnification from a set of N galaxy images,

$$\ln\mu = \sum_{i=1}^{N}\frac{\ln\omega_i - \langle\ln\omega\rangle_0(I_i)}{\sigma^2(I_i)}\left(\sum_{i=1}^{N}\frac{1}{\sigma^2(I_i)}\right)^{-1}\;. \tag{4.47}$$

A typical value for the dispersion is $\sigma(I) \approx 0.5$ (Bartelmann & Narayan 1995).

4.4.3. Relative Merits of Shear and Magnification Effect

It is interesting to compare the prospects of measuring shear and magnification caused by a deflector. We consider a small patch of the sky containing an expected number N of galaxy images (in the absence of lensing), which is sufficiently small so that the lens parameters κ and γ can be assumed to be constant. We also restrict the discussion to weak lensing case.

The dispersion of a shear estimate from averaging over galaxy ellipticities is σ_ε^2/N, so that the signal-to-noise ratio is

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{shear}} = \frac{|\gamma|}{\sigma_\varepsilon}\sqrt{N}\;. \tag{4.48}$$

According to (4.44), the expected change in galaxy number counts is $|\Delta N| = 2\kappa|\alpha-1|N$. Assuming Poissonian noise, the signal-to-noise ratio in this case is

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{counts}} = 2\kappa|\alpha-1|\sqrt{N}\;. \tag{4.49}$$

[7] Bright QSOs have a very steep number-count function, and so the flux enhancement of the sources outweighs the number reduction due to the stretching of the sky by a large margin. Whereas the lensing probability even for a high-redshift QSO is probably too small to affect the overall sources counts significantly, the fraction of multiply-imaged QSOs in flux-limited samples is increased through the magnification bias by a substantial factor over the probability that any individual QSO is multiply imaged (see, e.g. Turner et al. 1984; Narayan & Wallington 1993 and references therein).

Finally, the signal-to-noise ratio for the magnification estimate (4.47) is

$$\left(\frac{\mathrm{S}}{\mathrm{N}}\right)_{\mathrm{size}} = \frac{2\kappa}{\sigma(I)}\sqrt{N}\,, \tag{4.50}$$

assuming all $\sigma(I)$ are equal.

Comparing the three methods, we find

$$\frac{(\mathrm{S/N})_{\mathrm{shear}}}{(\mathrm{S/N})_{\mathrm{counts}}} = \frac{|\gamma|}{\kappa}\frac{1}{2\sigma_\varepsilon|\alpha-1|}\,, \quad \frac{(\mathrm{S/N})_{\mathrm{counts}}}{(\mathrm{S/N})_{\mathrm{size}}} = 2\sigma(I)|\alpha-1|\,. \tag{4.51}$$

If the lens situation is such that $\kappa \approx |\gamma|$ as for isothermal spheres, the first of eqs. (4.51) implies that the signal-to-noise of the shear measurement is considerably larger than that of the magnification. Even for number-count slopes as flat as $\alpha \sim 0.5$, this ratio is larger than five, with $\sigma_\varepsilon \sim 0.2$. The second of eqs. (4.51) shows that the size effect yields a somewhat larger signal-to-noise ratio than the number-density effect. We therefore conclude from these considerations that shear measurements should yield more significant results than magnification measurements.

This, however, is not the end of the story. Several additional considerations come into play when these three methods of measuring lensing effects are compared. First, the shear measurement is the only one for which we know precisely what to expect in the absence of lensing, whereas the other two methods need to compare the measurements with calibration fields void of lensing. These comparisons require very accurate photometry. Second, eq. (4.49) overestimates the signal-to-noise ratio since we assumed Poissonian errors, while real galaxies are known to cluster even at very faint magnitudes (e.g., Villumsen et al. 1997), and so the error is substantially underestimated. A particularly bad example for this effect has been found by Athreya et al. (1999) where a cluster at $z \sim 0.9$ seems to be behind the cluster (at $z = 0.3$) they investigated with weak-lensing techniques, as identified with photometric redshifts. Third, as we shall discuss in Sect. 4.6., observational effects such as atmospheric seeing affect the observable ellipticities and sizes of galaxy images, whereas the observed flux of galaxies is much less affected. Hence, the shear and size measurements require better seeing conditions than the number-count method. Both the number counts and the size measurements (at fixed surface brightness) require accurate photometry, which is not very important for the shear measurements. As we shall see in the course of this article, most weak-lensing measurements have indeed been obtained from galaxy ellipticities.

A more detailed study on the relative merits of shear and magnification methods has been performed by Schneider et al. (2000). Both methods were used to determine the parameters of mass profiles of spherically symmetric clusters. The results of this study can be summarised as follows: The magnification in many cases yields tighter constraints on the slope of the mass profiles, whereas the shear provides a more accurate determination of its amplitude (or lens strength). However, for the magnification methods to yield accurate results, the value of the unlensed number density n_0 needs to be known fairly accurately. In particular, for measurements out to large distances from the cluster centre (e.g., more than ~ 10 Einstein radii), even an error of a few per cent on n_0 destroys its relative advantage in the estimate of the shape relative to that of the shear. But, as we shall see in the next section, the magnification effect is very important for breaking an invariance transformation in the lens reconstruction that is permitted by shear measurements alone.

4.5. *Minimum Lens Strength for its Weak Lensing Detection*

After our detailed discussion of shear estimates and signal-to-noise ratios for local lensing measurements, it is interesting to ask how strong a deflecting mass distribution needs to be for a weak lensing measurement to recognise it. Our simplified consideration here suffices to gain insight into the dependence on the lens mass of the signal-to-noise ratio for a lens detection, and on the redshifts of lens and sources.

We model the deflector as a singular isothermal sphere (see Sect. 3.1.5., page 18). Let there be N galaxy images with ellipticities ε_i in an annulus centred on the lens and bounded by angular radii $\theta_{\mathrm{in}} \le \theta_i \le \theta_{\mathrm{out}}$. For simplicity, we restrict ourselves to weak lensing, so that $\mathrm{E}(\varepsilon) \approx \gamma$. For an axially-symmetric mass distribution, the shear is always tangentially oriented relative to the direction towards the mass centre, which is expressed by eq. (3.18) on page 18. We therefore consider the ellipticity component projected onto the tangential direction. It is formally defined by $\varepsilon_{\mathrm{t}} \equiv -\Re(\varepsilon\,\mathrm{e}^{-2\mathrm{i}\varphi})$, where φ is the polar angle of the galaxy position relative to the lens centre [see (3.18), page 18]. We now define an estimator for the lens strength by

$$X \equiv \sum_{i=1}^{N} a_i\,\varepsilon_{\mathrm{t}i}\,. \tag{4.52}$$

The factors $a_i = a(\theta_i)$ are arbitrary at this point, and will be chosen later such as to maximise the signal-to-noise ratio of the estimator (4.52). Note that the expectation value of X is zero in the absence of lensing, so that a significant non-zero value of X signifies the presence of a lens. The expectation value for an isothermal sphere is $\mathrm{E}(X) = \theta_{\mathrm{E}}\sum_i a_i/(2\theta_i)$, where we used (3.18, page 18), and

$$\mathrm{E}(X^2) = \sum_{i,j=1}^{N} a_i a_j\,\mathrm{E}(\varepsilon_{\mathrm{t}i}\varepsilon_{\mathrm{t}j}) = [\mathrm{E}(X)]^2 + \frac{\sigma_\varepsilon^2}{2}\sum_{i=1}^{N} a_i^2\,. \tag{4.53}$$

We employed $\mathrm{E}(\varepsilon_{\mathrm{t}i}\varepsilon_{\mathrm{t}j}) = \gamma_{\mathrm{t}}(\theta_i)\gamma_{\mathrm{t}}(\theta_j) + \delta_{ij}\sigma_\varepsilon^2/2$ here, and the factor two is due to the fact that the ellipticity dispersion only refers to one component of the ellipticity, while σ_ε is defined as the dispersion of the two-component ellipticity. Therefore, the signal-to-noise ratio for a detection of the lens is

$$\frac{\mathrm{S}}{\mathrm{N}} = \frac{\theta_{\mathrm{E}}}{\sqrt{2}\sigma_\varepsilon}\frac{\sum_i a_i\,\theta_i^{-1}}{\sqrt{\sum_i a_i^2}}\,. \tag{4.54}$$

Differentiating (S/N) with respect to a_j, we find that (S/N) is maximised if the a_i are chosen $\propto \theta_i^{-1}$. Inserting this choice into (4.54) yields $\mathrm{S/N} = 2^{-1/2}\theta_{\mathrm{E}}\sigma_\varepsilon^{-1}\left(\sum_i\theta_i^{-2}\right)^{1/2}$. We now replace the sum by its ensemble average over the annulus, $\left\langle\sum_i\theta_i^{-2}\right\rangle = N\left\langle\theta^{-2}\right\rangle = 2n\pi\ln(\theta_{\mathrm{out}}/\theta_{\mathrm{in}})$, where we used $N = \pi n(\theta_{\mathrm{out}}^2 - \theta_{\mathrm{in}}^2)$, with the number density of galaxy images n. Substituting this result into (4.54), and using the definition of the Einstein radius (3.17, page 18), the signal-to-noise ratio becomes

$$\begin{aligned}\frac{\mathrm{S}}{\mathrm{N}} &= \frac{\theta_{\mathrm{E}}}{\sigma_\varepsilon}\sqrt{\pi n}\sqrt{\ln(\theta_{\mathrm{out}}/\theta_{\mathrm{in}})} \qquad (4.55)\\ &= 12.7\left(\frac{n}{30\,\mathrm{arc\,min}^{-2}}\right)^{1/2}\left(\frac{\sigma_\varepsilon}{0.2}\right)^{-1}\left(\frac{\sigma_v}{600\,\mathrm{km\,s}^{-1}}\right)^2\\ &\times \left(\frac{\ln(\theta_{\mathrm{out}}/\theta_{\mathrm{in}})}{\ln 10}\right)^{1/2}\left\langle\frac{D_{\mathrm{ds}}}{D_{\mathrm{s}}}\right\rangle\,.\end{aligned}$$

As expected, the signal-to-noise ratio is proportional the square root of the number density of galaxies and the inverse of the intrinsic ellipticity dispersion. Furthermore, it is proportional to the square of the velocity dispersion σ_v. Assuming the fiducial values given in eq. (4.55) and a typical value of $(D_{\rm ds}/D_{\rm s}) \sim 0.5$, lenses with velocity dispersion in excess of $\sim 600\,{\rm km\,s^{-1}}$ can be detected with a signal-to-noise $\gtrsim 6$. This shows that galaxy clusters will yield a significant weak lensing signal, and explains why clusters have been the main target for weak-lensing research up to now. Individual galaxies with $\sigma_v \sim 200\,{\rm km\,s^{-1}}$ cannot be detected with weak-lensing techniques. If one is interested in the statistical properties of the mass distribution of galaxies, the lensing effects of $N_{\rm gal}$ galaxies need to be statistically superposed, increasing (S/N) by a factor of $\sqrt{N_{\rm gal}}$. Thus, it is necessary to superpose several hundred galaxies to obtain a significant galaxy-galaxy lensing signal. We shall return to this topic in Sect. 7. on page 60.

We finally note that (4.55) also demonstrates that the detection of lenses will become increasingly difficult with increasing lens redshift, as the last factor is a sensitive function of $z_{\rm d}$. Therefore, most lenses so far investigated with weak-lensing techniques have redshifts below 0.5. High-redshift clusters have only recently become the target of detailed lensing studies.

4.6. Practical Consideration for Measuring Image Shapes

4.6.1. General Discussion

Real astronomical data used for weak lensing are supplied by CCD images. The steps from a CCD image to a set of galaxy images with measured ellipticities are highly non-trivial and cannot be explained in any detail in the frame of this review. Nevertheless, we want to mention some of the problems together with the solutions which were suggested and applied.

The steps from CCD frames to image ellipticities can broadly be grouped into four categories; data reduction, image detection, shape determination, and corrections for the point-spread function. The data-reduction process is more or less standard, involving de-biasing, flat-fielding, and removal of cosmic rays and bad pixels. For the latter purpose, it is essential to have several frames of the same field, slightly shifted in position. This also allows the the flat field to be determined from the images themselves (a nice description of these steps is given in Mould et al. 1994). To account for telescope and instrumental distortions, the individual frames have to be re-mapped before being combined into a final image. In order to do this, the geometric distortion has to be either known or stable. In the latter case, it can be determined by measuring the positions and shapes of stellar images (e.g., from a globular cluster). In Mould et al. (1994), the classical optical aberrations were determined and found to be in good agreement with the system's specifications obtained from ray-tracing analysis.

With the individual frames stacked together in the combined image, the next step is to detect galaxies and to measure their shapes. This may appear simple, but is in fact not quite as straightforward, for several reasons. Galaxy images are not necessarily isolated on the image, but they can overlap, e.g. with other galaxies. Since weak-lensing observations require a large number density of galaxy images, such merged images are not rare. The question then arises whether a detected object is a single galaxy, or a merged pair, and depending on the choice made, the measured ellipticities will be much different. Second, the image is noisy because of the finite number of photons per pixel and the noise intrinsic to the CCD electronics. Thus, a local enhancement of counts needs to be classified as a statistically significant source detection, and a conservative signal-to-noise threshold reduces the number of galaxy images. Third, galaxy images have to be distinguished from stars. This is not a severe problem, in particular if the field studied is far from the Galactic plane where the number density of stars is small.

Several data-analysis software packages exist, such as FOCAS (Jarvis & Tyson 1981) and SExtractor (Bertin & Arnouts 1996). They provide routines, based on algorithms developed from experience and simulated data, for objective selection of objects and measuring their centroids, their multipole moments, their magnitudes, and classify them as stars or extended objects. Kaiser et al. (1995) developed their own object detection algorithm. It is based on convolving the CCD image with two-dimensional Mexican hat-shaped filter functions of variable width $\theta_{\rm s}$. For each value of $\theta_{\rm s}$, the maxima of the smoothed intensity map are localised. Varying $\theta_{\rm s}$, these maxima form curves in the three-dimensional space spanned by $\vec\theta$ and $\theta_{\rm s}$. Along each such curve, the significance of a source detection is calculated, and the maximum of the significance is defined as the location $\vec\theta$ of an object with corresponding size $\theta_{\rm s}$.

Once an object is found, the quadrupole moments can in principle be obtained from (4.2). In practice, however, this is not necessarily the most practical definition of the moment tensor. The function $q_I(I)$ in (4.2) should be chosen such that it vanishes for surface brightnesses close to and smaller than the sky brightness; otherwise, one would sample too much noise. On the other hand, if q_I is cut off at too bright values of I, the area within which the quadrupole moments are measured becomes too small, and the effects of seeing (see below) become overwhelming. Also, with a too conservative cut-off, many galaxy images would be missed. Assume, for instance, that $q_I(I) = I\,{\rm H}(I - I_{\rm th})$. One would then choose $I_{\rm th}$ such that it is close to, but a few $\sigma_{\rm noise}$ above the sky background, and the quadrupole moments would then be measured inside the resulting limiting isophote. Since this isophote is close to the sky background, its shape is affected by sky noise. This implies that the measured quadrupole moments will depend highly non-linearly on the brightness on the CCD; in particular, the effect of noise will enter the measured ellipticities in a non-linear fashion. A more robust measurement of the quadrupole moments is obtained by replacing the weight function $q_I[I(\vec\theta)]$ in (4.2) by $IW(\vec\theta)$, where $W(\vec\theta)$ explicitly depends on $\vec\theta$. Kaiser et al. (1995) use a Gaussian of size $\theta_{\rm s}$ as their weight function W, i.e., the size of their W is the scale on which the object was detected at highest significance. It should be noted that the quadrupole moments obtained with a weight function $W(\vec\theta)$ do not obey the transformation law (4.5), and therefore, the expectation value of the ellipticity, ${\rm E}(\varepsilon)$, will be different from the reduced shear g. We return to this issue further below.

Another severe difficulty for the determination of the local shear is atmospheric seeing. Due to atmospheric turbulence, a point-like source will be seen from the ground as an extended image; the source is smeared-out. Mathematically, this can be described as a convolution. If $I(\vec\theta)$ is the surface brightness before passing the Earth's atmosphere, the observed brightness distribution $I^{\rm (obs)}(\vec\theta)$ is

$$I^{\rm (obs)}(\vec\theta) = \int {\rm d}^2\vartheta\, I(\vec\vartheta)\, P(\vec\theta - \vec\vartheta)\;, \tag{4.56}$$

where $P(\vec\theta)$ is the *point-spread function* (PSF) which describes the brightness distribution of a point source on the CCD. $P(\vec\theta)$ is normalised to unity and centred on $\vec 0$. The characteristic width of the PSF is called the size of the seeing disc. The smaller it is, the less smeared the images are. A seeing well below 1 arc second is required for weak-lensing observations, and there are only a handful of telescope sites where such seeing conditions are regularly met. The reason for this strong requirement on the data quality lies in the fact that weak-lensing studies require a high number density of galaxy images, i.e., the observations have to be extended to faint magnitudes. But the characteristic angular size of faint galaxies is below 1 arc second. If the seeing is larger than that, the shape information is diluted or erased.

The PSF includes not only the effects of the Earth's atmosphere, but also pointing errors of the telescope (e.g., caused by wind shake). Therefore, the PSF will in general be slightly anisotropic. Thus, seeing has two important effects on the observed image ellipticities: Small elliptical images become rounder, and the anisotropy of the PSF introduces a systematic, spurious image ellipticity. The PSF can be determined directly from the CCD once a number of isolated stellar images are identified. The shape of the stars (which serve as point sources) reflects the PSF. Note that the PSF is not necessarily constant across the CCD. If the number density of stellar images is sufficiently large, one can empirically describe the PSF variation across the field by a low-order polynomial. An additional potential difficulty is the chromaticity of the PSF, i.e. the dependence of the PSF on the spectral energy distribution of the radiation. The PSF as measured from stellar images is not necessarily the same as the PSF which applies to galaxies, due to their different spectra. The difference of the PSFs is larger for broader filters. However, it is assumed that the PSF measured from stellar images adequately represents the PSF for galaxies.

In the idealised case, in which the quadrupole moments are defined with the weight function $q_I(I) = I$, the effect of the PSF on the observed image ellipticities can easily be described. If P_{ij} denotes the quadrupole tensor of the PSF, defined in complete analogy to (4.2), then the observed quadrupole tensor $Q_{ij}^{(\rm obs)}$ is related to the true one by $Q_{ij}^{(\rm obs)} = P_{ij} + Q_{ij}$ (see Valdes et al. 1983). The ellipticity χ then transforms like

$$\chi^{(\rm obs)} = \frac{\chi + T\chi^{(\rm PSF)}}{1+T} \;, \tag{4.57}$$

where

$$T = \frac{P_{11}+P_{22}}{Q_{11}+Q_{22}} \;; \quad \chi^{(\rm PSF)} = \frac{P_{11}-P_{22}+2{\rm i}P_{12}}{P_{11}+P_{22}} \;. \tag{4.58}$$

Thus, T expresses the ratio of the PSF size to the image size before convolution, and $\chi^{(\rm PSF)}$ is the PSF ellipticity. It is evident from (4.57) that the smaller T, the less $\chi^{(\rm obs)}$ deviates from χ. In the limit of very large T, $\chi^{(\rm obs)}$ approaches $\chi^{(\rm PSF)}$. In principle, the relation (4.57) could be inverted to obtain χ from $\chi^{(\rm obs)}$. However, this inversion is unstable unless T is sufficiently small, in the sense that noise affecting the measurement of $\chi^{(\rm obs)}$ is amplified by the inversion process. Unfortunately, these simple transformation laws only apply for the specific choice of the weight function. For weighting schemes that can be applied to real data, the resulting transformation becomes much more complicated.

If a galaxy image features a bright compact core which emits a significant fraction of the galaxy's light, this core will be smeared out by the PSF. In that case, $\chi^{(\rm obs)}$ may be dominated by the core and thus contain little information about the galaxy ellipticity. This fact motivated Bonnet & Mellier (1995) to define the quadrupole moments with a weight function $W(\vec\theta)$ which not only cuts off at large angular separations, but which is also small near $\vec\theta = \vec 0$. Hence, their weight function q is significantly nonzero in an annulus with radius and width both being of the order of the size of the PSF.

The difficulties mentioned above prohibit the determination of the local reduced shear by straight averaging over the directly measured image ellipticities. This average is affected by the use of a angle-dependent weight function W in the practical definition of the quadrupole moments, by the finite size of the PSF and its anisotropy, and by noise. Bonnet & Mellier (1995) have performed detailed simulations of CCD frames which resemble real observations as close as possible, including an anisotropic PSF. With these simulations, the efficiency of object detection, the accuracy of their centre positions, and the relation between true and measured image ellipticities can be investigated in detail, and so the relation between mean ellipticity and (reduced) shear can approximately be calibrated. Wilson et al. (1996) followed a very similar approach, except that the analysis of their simulated CCD frames was performed with FOCAS. Assuming an isotropic PSF, the mean image ellipticity is proportional to the reduced shear, $g \approx f\langle\epsilon\rangle$, with a correction factor f depending on the limiting galaxy magnitude, the photometric depth of the image, and the size of the seeing disk. For a seeing of $0\farcs8$, Bonnet & Mellier obtained a correction factor $f \sim 6$, whereas the correction factor in Wilson et al. for the same seeing is $f \sim 1.5$. This large difference is not a discrepancy, but due to the different definitions of the quadrupole tensor. Although the correction factor is much larger for the Bonnet & Mellier method, they show that their measured (and calibrated) shear estimate is more accurate than that obtained with FOCAS. Kaiser et al. (1995) used CCD frames taken with WFPC2 on board *HST* which are unaffected by atmospheric seeing, sheared them, and degraded the resulting images by a PSF typical for ground-based images and by adding noise. In this way, they calibrated their shear measurement and tested their removal of an anisotropic contribution of the PSF.

However, calibrations relying on simulated images are not fully satisfactory since the results will depend on the assumptions underlying the simulations. Kaiser et al. (1995) and Luppino & Kaiser (1997) presented a perturbative approach for correcting the observed image ellipticities for PSF effects, with additional modifications made by Hoekstra et al. (1998) and Hudson et al. (1998). Since the measurement of ellipticities lies at the heart of weak lensing studies, we shall present this approach in the next subsection, despite its being highly technical.

4.6.2. The KSB Method

Closely following the work by Kaiser et al. (1995), this subsection provides a relation between the observed image ellipticity and a source ellipticity known to be isotropically distributed. The relation corrects for PSF smearing and its anisotropy, and it also takes into account that the transformation (4.5) no longer applies if the weight factor explicitly depends on $\vec\theta$.

We consider the quadrupole tensor

$$Q_{ij} = \int {\rm d}^2\theta\,(\theta_i-\bar\theta_i)(\theta_j-\bar\theta_j)\,I(\vec\theta)\,W\left(|\vec\theta-\bar{\vec\theta}|^2/\sigma^2\right) \;, \tag{4.59}$$

where W contains a typical scale σ, and $\bar{\vec\theta}$ is defined as in (4.1), but with the new weight function. Note that, in contrast to the definition (4.2), this tensor is no longer normalised by the flux, but this does not affect the definition (4.4) of the complex ellipticity.

The relation between the observed surface brightness $I^{\rm obs}(\vec\theta)$ and the true surface brightness I is given by (4.56). We assume in the following that P is nearly isotropic, so that the anisotropic part of P is small. Then, we define the isotropic part $P^{\rm iso}$ of P as the azimuthal average over P, and decompose P into an isotropic and an anisotropic part as

$$P(\vec\vartheta) = \int {\rm d}^2\varphi\, q(\vec\varphi)\, P^{\rm iso}(\vec\vartheta-\vec\varphi) \;, \tag{4.60}$$

which defines q uniquely. In general, $q(\vec\varphi)$ will be an almost singular function, but we shall show later that it has well-behaved moments. Both $P^{\rm iso}$ and q are normalised to unity and have vanishing first moments. With $P^{\rm iso}$, we define the brightness profiles

$$
\begin{aligned}
I^{\rm iso}(\vec\theta) &= \int {\rm d}^2\varphi\, I(\vec\varphi)\, P^{\rm iso}(\vec\theta-\vec\varphi) \\
I^{0}(\vec\theta) &= \int {\rm d}^2\varphi\, I^{\rm s}(\vec\varphi)\, P^{\rm iso}(\vec\theta-\vec\varphi)\;. \qquad (4.61)
\end{aligned}
$$

The first of these would be observed if the true image was smeared only with an isotropic PSF, and the second is the unlensed source smeared with $P^{\rm iso}$. Both of these brightness profiles are unobservable, but convenient for the following discussion. For each of them, we can define a quadrupole tensor as in (4.59). From each quadrupole tensor, we define the complex ellipticity $\chi=\chi_1+{\rm i}\chi_2$, in analogy to (4.4).

If we define the centres of images including a spatial weight function, the property that the centre of the image is mapped onto the centre of the source through the lens equation is no longer strictly true. However, the deviations are expected to be very small in general and will be neglected in the following. Hence, we choose coordinates such that $\bar{\vec\theta}=\vec 0$, and approximate the other centres to be at the origin as well.

According to our fundamental assumption that the intrinsic ellipticities are randomly oriented, this property is shared by the ellipticities χ^0 defined in terms of I^0 [see (4.61)], because it is unaffected by an isotropic PSF. Therefore, we can replace (4.13) by ${\rm E}(\chi^0)=0$ in the determination of g. The task is then to relate the observed image ellipticity $\chi^{\rm obs}$ to χ^0. We break it into several steps.

From $\chi^{\rm iso}$ to $\chi^{\rm obs}$. We first look into the effect of an anisotropic PSF on the observed ellipticity. According to (4.60) and (4.61),

$$
I^{\rm obs}(\vec\theta)=\int {\rm d}^2\varphi\, q(\vec\theta-\vec\varphi)\, I^{\rm iso}(\vec\varphi)\;. \qquad (4.62)
$$

Let $f(\vec\theta)$ be an arbitrary function, and consider

$$
\begin{aligned}
&\int {\rm d}^2\theta\, f(\vec\theta)\, I^{\rm obs}(\vec\theta)=\int {\rm d}^2\varphi\, I^{\rm iso}(\vec\varphi)\int {\rm d}^2\vartheta\, f(\vec\varphi+\vec\vartheta)\, q(\vec\vartheta) \\
&=\int {\rm d}^2\varphi\, I^{\rm iso}(\vec\varphi)\, f(\vec\varphi) \\
&+\frac{1}{2}q_{kl}\int {\rm d}^2\varphi\, I^{\rm iso}(\vec\varphi)\frac{\partial^2 f}{\partial\varphi_k\partial\varphi_l}+O(q^2)\;. \qquad (4.63)
\end{aligned}
$$

We used the fact that q is normalised and has zero mean, and defined

$$
q_{ij}=\int {\rm d}^2\varphi\, q(\vec\varphi)\,\varphi_i\varphi_j\;,\quad q_1\equiv q_{11}-q_{22}\;,\quad q_2\equiv 2q_{12}\;. \quad (4.64)
$$

The tensor q_{ij} is trace-less, $q_{11}=-q_{22}$, following from (4.60). We consider in the following only terms up to linear order in q. To that order, we can replace $I^{\rm iso}$ by $I^{\rm obs}$ in the final term in (4.63), since the difference would yield a term $\propto O(q^2)$. Hence,

$$
\begin{aligned}
&\int {\rm d}^2\varphi\, I^{\rm iso}(\vec\varphi)\, f(\vec\varphi)\approx\int {\rm d}^2\theta\, f(\vec\theta)\, I^{\rm obs}(\vec\theta) \\
&-\frac{1}{2}q_{kl}\int {\rm d}^2\varphi\, I^{\rm obs}(\vec\varphi)\frac{\partial^2 f}{\partial\varphi_k\partial\varphi_l}\;. \qquad (4.65)
\end{aligned}
$$

Setting $\sigma_{\rm iso}=\sigma_{\rm obs}\equiv\sigma$ in the definition of the quadrupole tensors $Q^{\rm iso}$ and $Q^{\rm obs}$, and choosing $f(\vec\theta)=\theta_i\theta_j W(|\vec\theta|^2/\sigma^2)$, yields

$$
Q^{\rm iso}_{ij}=Q^{\rm obs}_{ij}-\frac{1}{2}Z_{ijkl}q_{kl}\;, \qquad (4.66)
$$

where the Einstein summation convention was adopted, and where

$$
Z_{ijkl}=\int {\rm d}^2\varphi\, I^{\rm obs}(\vec\varphi)\,\frac{\partial^2}{\partial\varphi_k\partial\varphi_l}\left[\varphi_i\varphi_j\, W\left(\frac{|\vec\varphi|^2}{\sigma^2_{\rm iso}}\right)\right]\;. \qquad (4.67)
$$

This then yields

$$
\begin{aligned}
{\rm tr}(Q^{\rm iso}) &= {\rm tr}(Q^{\rm obs})-x_\alpha q_\alpha\;, \\
(Q^{\rm iso}_{11}-Q^{\rm iso}_{22}) &= (Q^{\rm obs}_{11}-Q^{\rm obs}_{22})-X_{1\alpha}q_\alpha\;, \quad \text{and} \\
2Q^{\rm iso}_{12} &= 2Q^{\rm obs}_{12}-X_{2\alpha}q_\alpha\;, \qquad (4.68)
\end{aligned}
$$

where the sums run over $\alpha=1,2$.[8] Up to linear order in q_α,

$$
\chi^{\rm iso}_\alpha=\chi^{\rm obs}_\alpha-P^{\rm sm}_{\alpha\beta}q_\beta\;, \qquad (4.69)
$$

with the definitions

$$
\begin{aligned}
P^{\rm sm}_{\alpha\beta} &= \frac{1}{{\rm tr}Q^{\rm obs}}\left(X_{\alpha\beta}-\chi^{\rm obs}_\alpha x_\beta\right) \\
X_{\alpha\beta} &= \int {\rm d}^2\varphi\, I^{\rm obs}(\vec\varphi)\left[\left(W+2|\vec\varphi|^2\frac{W'}{\sigma^2_{\rm iso}}\right)\delta_{\alpha\beta}\right. \\
&+ \left.\eta_\alpha(\vec\varphi)\eta_\beta(\vec\varphi)\frac{W''}{\sigma^4_{\rm iso}}\right]\;, \\
x_\alpha &= \int {\rm d}^2\varphi\, I^{\rm obs}(\vec\varphi)\,\eta_\alpha(\vec\varphi)\left(\frac{2W'}{\sigma^2_{\rm iso}}+|\vec\varphi|^2\frac{W''}{\sigma^4_{\rm iso}}\right)\;, \qquad (4.70)
\end{aligned}
$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol, and

$$
\eta_1(\vec\theta)=\theta_1^2-\theta_2^2\;;\quad \eta_2(\vec\theta)=2\theta_1\theta_2\;. \qquad (4.71)
$$

$P^{\rm sm}_{\alpha\beta}$ was dubbed *smear polarisability* in Kaiser et al. (1995). It describes the (linear) response of the ellipticity to a PSF anisotropy. Note that $P^{\rm sm}_{\alpha\beta}$ depends on the observed brightness profile. In particular, its size decreases for larger images, as expected: The ellipticities of larger images are less affected by a PSF anisotropy than those of smaller images.

The determination of q_α. Equation (4.69) provides a relation between the ellipticities of an observed image and a hypothetical image smeared by an isotropic PSF. In order to apply this relation, the anisotropy term q_α needs to be known. It can be determined from the shape of stellar images.

Since stars are point-like and unaffected by lensing, their isotropically smeared images have zero ellipticity, $\chi^{*,\rm iso}=0$. Hence, from (4.69),

$$
q_\alpha=(P^{*,\rm sm})^{-1}_{\alpha\beta}\,\chi^{*,\rm obs}_\beta\;. \qquad (4.72)
$$

In general, the PSF varies with the position of an image. If this variation is sufficiently smooth, q can be measured for a set of stars, and approximated by a low-order polynomial across the data field. As pointed out by Hoekstra et al. (1998), the scale size σ in the measurement of q is best chosen to be the same as that of the galaxy image under consideration. Hence, for each value of σ, such a polynomial fit is constructed. This approach works well and provides an estimate of q at the position of all galaxies, which can then be used in the transformation (4.69).

[8] We use Greek instead of Latin indices $\alpha,\beta=1,2$ to denote that they are not tensor indices. In particular, the components of χ do not transform like a vector, but like the trace-less part of a symmetric tensor.

From χ^0 to $\chi^{\rm iso}$. We now relate $\chi^{\rm iso}$ to the ellipticity χ^0 of a hypothetical image obtained from isotropic smearing of the source. To do so, we use (4.61) and (3.10) in the form $I(\vec\theta) = I^s(\mathcal{A}\vec\theta)$, and consider

$$
\begin{aligned}
I^{\rm iso}(\vec\theta) &= \int {\rm d}^2\varphi\, I^s(\mathcal{A}\vec\varphi)\, P^{\rm iso}(\vec\theta-\vec\varphi) \qquad (4.73)\\
&= \frac{1}{\det\mathcal{A}}\int {\rm d}^2\zeta\, I^s(\vec\zeta)\, P^{\rm iso}(\vec\theta-\mathcal{A}^{-1}\vec\zeta) \equiv \hat I(\mathcal{A}\vec\theta)\;.
\end{aligned}
$$

The second step is merely a transformation of the integration variable, and in the final step we defined the brightness moment

$$
\hat I(\vec\theta) = \int {\rm d}^2\varphi\, I^s(\vec\varphi)\, \hat P(\vec\theta-\vec\varphi) \quad \text{with} \quad \hat P(\vec\theta) \equiv \frac{1}{\det\mathcal{A}} P^{\rm iso}(\mathcal{A}^{-1}\vec\theta)\;. \tag{4.74}
$$

The function $\hat P$ is normalised and has zero mean. It can be interpreted as a PSF relating $\hat I$ to I^s. The presence of shear renders $\hat P$ anisotropic.

We next seek to find a relation between the ellipticities of $I^{\rm iso}$ and $\hat I$:

$$
\begin{aligned}
\hat Q_{ij} &= \int {\rm d}^2\beta\, \beta_i\beta_j\, \hat I(\vec\beta)\, W\left(\frac{|\vec\beta|^2}{\hat\sigma^2}\right) \qquad (4.75)\\
&= \det\mathcal{A}\,\mathcal{A}_{ik}\mathcal{A}_{jl}\int {\rm d}^2\theta\,\theta_k\theta_l\, I^{\rm iso}(\vec\theta)\, W\left(\frac{|\vec\theta|^2-\delta_\alpha\eta_\alpha(\vec\theta)}{\sigma^2}\right)\;.
\end{aligned}
$$

The relation between the two filter scales is given by $\hat\sigma^2 = (1-\kappa)^2(1+|g|^2)\sigma^2$, and δ is the distortion (4.15). For small δ, we can employ a first-order Taylor expansion of the weight function W in the previous equation. This results in the following relation between $\hat\chi$ and $\chi^{\rm iso}$:

$$
\chi^{\rm iso}_\alpha - \hat\chi_\alpha = C_{\alpha\beta}\, g_\beta\;, \tag{4.76}
$$

where

$$
\begin{aligned}
C_{\alpha\beta} &= 2\delta_{\alpha\beta} - 2\chi^{\rm iso}_\alpha\chi^{\rm iso}_\beta + \frac{2}{{\rm tr}(Q^{\rm iso})}\chi^{\rm iso}_\alpha L_\beta - \frac{2}{{\rm tr}(Q^{\rm iso})} B_{\alpha\beta}\;,\\
B_{\alpha\beta} &= -\int {\rm d}^2\theta\, I^{\rm iso}(\vec\theta)\, W'\left(\frac{|\vec\theta|^2}{\sigma^2}\right)\frac{1}{\sigma^2}\,\eta_\alpha(\vec\theta)\,\eta_\beta(\vec\theta)\\
L_\alpha &= -\int {\rm d}^2\theta\, |\vec\theta|^2 I^{\rm iso}(\vec\theta)\, W'\left(\frac{|\vec\theta|^2}{\sigma^2}\right)\frac{1}{\sigma^2}\,\eta_\alpha(\vec\theta)\;. \qquad (4.77)
\end{aligned}
$$

C is the *shear polarisability* of Kaiser et al. (1995). Whereas C is defined in terms of $I^{\rm iso}$, owing to the assumed smallness of q, the difference of C calculated with $I^{\rm iso}$ and $I^{\rm obs}$ would cause a second-order change in (4.76) and is neglected, so that we can calculate C directly from the observed brightness profile.

In analogy to (4.60), we can decompose $\hat P$ into an isotropic and an anisotropic part, the latter one being small due to the assumed smallness of the shear,

$$
\hat P(\vec\theta) = \int {\rm d}^2\varphi\, \hat P^{\rm iso}(\vec\varphi)\, \hat q(\vec\theta-\vec\varphi)\;. \tag{4.78}
$$

Defining the brightness profile which would be obtained from smearing the source with the isotropic PSF $\hat P^{\rm iso}$, $\hat I^0(\vec\theta) = \int {\rm d}^2\varphi\, I^s(\vec\varphi)\, \hat P^{\rm iso}(\vec\theta-\vec\varphi)$, one finds

$$
\hat I(\vec\theta) = \int {\rm d}^2\varphi\, \hat I^0(\vec\varphi)\, \hat q(\vec\theta-\vec\varphi)\;. \tag{4.79}
$$

Thus, the relation between $\hat I$ and $\hat I^0$ is the same as that between $I^{\rm obs}$ and $I^{\rm iso}$, and we can write

$$
\hat\chi^0_\alpha = \hat\chi_\alpha - P^{\rm sm}_{\alpha\beta}\,\hat q_\beta\;. \tag{4.80}
$$

Note that $P^{\rm sm}$ should in principle be calculated by using $\hat I$ instead of $I^{\rm obs}$ in (4.69). However, due to the assumed smallness of g and q, the differences between $I^{\rm obs}$, $I^{\rm iso}$, and $\hat I$ are small, namely of first order in g and q. Since $\hat q$ is of order g [as is obvious from its definition, and will be shown explicitly in (4.82)], this difference in the calculation of $P^{\rm sm}$ would be of second order in (4.80) and is neglected here.

Eliminating $\hat\chi$ from (4.76) and (4.80), we obtain

$$
\chi^{\rm iso}_\alpha = \hat\chi^0_\alpha + C_{\alpha\beta}\, g_\beta + P^{\rm sm}_{\alpha\beta}\,\hat q_\beta\;. \tag{4.81}
$$

Now, for stellar objects, both $\hat\chi^0$ and $\chi^{\rm iso}$ vanish, which implies a relation between $\hat q$ and g,

$$
\hat q_\alpha = -(P^{\rm sm*})^{-1}_{\alpha\beta}\, C^*_{\beta\gamma}\, g_\gamma\;, \tag{4.82}
$$

where the asterisk indicates that $P^{\rm sm}$ and C are to be calculated from stellar images. Whereas the result should in principle not depend on the choice of the scale length in the weight function, it does so in practice. As argued in Hoekstra et al. (1998), one should use the same scale length in $P^{\rm sm*}$ and C^* as for the galaxy object for which the ellipticities are measured. Defining now

$$
P^g_{\alpha\beta} = C_{\alpha\beta} - P^{\rm sm}_{\alpha\gamma}(P^{\rm sm*})^{-1}_{\gamma\delta}\, C^*_{\delta\beta}\;, \tag{4.83}
$$

and combining (4.69) and (4.81), we finally obtain

$$
\hat\chi^0 = \chi^{\rm obs} - P^{\rm sm}_{\alpha\beta}\, q_\beta - P^g_{\alpha\beta}\, g_\beta\;. \tag{4.84}
$$

This equation relates the observed ellipticity to that of the source smeared by an isotropic PSF, using the PSF anisotropy and the reduced shear g. Since the expectation value of $\hat\chi^0$ is zero, (4.84) yields an estimate of g. The two tensors $P^{\rm sm}$ and P^g can be calculated from the brightness profile of the images. Whereas the treatment has been confined to first order in the PSF anisotropy and the shear, the simulations in Kaiser et al. (1995) and Hoekstra et al. (1998) show that the resulting equations can be applied even for moderately large shear. A numerical implementation of these relations, the `imcat` software, is provided by N. Kaiser (see `http://www.ifa.hawaii.edu/~kaiser`). We also note that modifications of this scheme were recently suggested (Rhodes et al. 1999, Kaiser 1999), as well as a completely different approach to shear measurements (Kuijken 1999). Kaiser et al. (1999b) provide a detailed description of the image analysis of weak-lensing data from a large CCD-array camera.

5. WEAK LENSING BY GALAXY CLUSTERS

5.1. *Introduction*

So far, weak gravitational lensing has chiefly been applied to determine the mass distribution of medium-redshift galaxy clusters. The main reason for this can be seen from eq. (4.55): Clusters are massive enough to be individually detected by weak lensing. More traditional methods to infer the matter distribution in clusters are (a) dynamical methods, in which the observed line-of-sight velocity distribution of cluster galaxies is used in conjunction with the virial theorem, and (b) the investigation of the diffuse X–ray emission from the hot ($\sim 10^7$ K) intra-cluster gas residing in the cluster potential well (see, e.g., Sarazin 1986).

Both of these methods are based on rather strong assumptions. For the dynamical method to be reliable, the cluster must be in or near virial equilibrium, which is not guaranteed because the typical dynamical time scale of a cluster is not much shorter than the Hubble time H_0^{-1}, and the substructure abundantly observed in clusters indicates that an appreciable fraction of them is still in the process of formation. Projection effects and the anisotropy of galaxy orbits in clusters further affect the mass determination by dynamical methods. On the other hand, X–ray analyses rely on the assumption that the intra-cluster gas is in hydrostatic equilibrium. Owing to the finite spatial and energy resolution of existing X–ray instruments, one often has to conjecture the temperature profile of the gas. Here, too, the influence of projection effects is difficult to assess.

Whereas these traditional methods have provided invaluable information on the physics of galaxy clusters, and will continue to do so, gravitational lensing offers a welcome alternative approach, for it determines the projected mass distribution of a cluster independent of the physical state and nature of the matter. In particular, it can be used to calibrate the other two methods, especially for clusters showing evidence of recent merger events, for which the equilibrium assumptions are likely to fail. Finally, as we shall show below, the determination of cluster mass profiles by lensing is theoretically simple, and recent results show that the observational challenges can also be met with modern telescopes and instruments.

Both shear and magnification effects have been observed in a number of galaxy clusters. In this chapter, we discuss the methods by which the projected mass distribution in clusters can be determined from the observed lensing effects, and show some results of mass reconstructions, together with a brief discussion of their astrophysical relevance. In principle, voids could also be measured using the same methods, but as shown in Amendola et al. (1999), their (negative) density contrast is too small for a detection under realistic assumptions. Sect. 5.2. presents the principles of cluster mass reconstruction from estimates of the (reduced) shear obtained from image ellipticities (also recently reviewed by Umetsu et al. 1999). In contrast to the two-dimensional mass maps generated by these reconstructions, the aperture mass methods discussed in Sect. 5.3. determine a single number to characterise the bulk properties of the cluster mass. Observational results are presented in Sect. 5.4.. We outline further developments in the final section, including the combined analysis of shear and magnification effects, maximum-likelihood methods for the mass reconstruction, and a method for measuring local lens parameters from the extragalactic background noise.

5.2. Cluster Mass Reconstruction from Image Distortions

We discussed in detail in Sect. 4. how the distortion of image shapes can be used to determine the local tidal gravitational field of a cluster. In this section, we describe how this information can be used to construct two-dimensional mass maps of clusters.

Shortly after the discovery of giant luminous arcs (Soucail et al. 1987a; Lynds & Petrosian 1989), Fort et al. (1988) detected a number of distorted galaxy images in the cluster A 370. They also interpreted these *arclets* as distorted background galaxy images, but on a weaker level than the giant luminous arc in the same cluster. The redshift determination of one arclet by Mellier et al. (1991) provided early support for this interpretation. Tyson et al. (1990) discovered a coherent distortion of faint galaxy images in the clusters A 1689 and Cl 1409+52, and constrained their (dark) mass profiles from the observed 'shear'. Kochanek (1990) and Miralda-Escude (1991) studied in detail how parameterised mass models for clusters can be constrained from such distortion measurements.

The field began to flourish after Kaiser & Squires (1993) found that the distortions can be used for parameter-free reconstructions of cluster surface mass densities. Their method, and several variants of it, will be described in this section. It has so far been applied to about 15 clusters, and this number is currently limited by the number of available dark nights with good observing conditions at the large telescopes which are required for observations of weak lensing.

5.2.1. Linear Inversion of Shear Maps

Equation (3.15, page 18) shows that the shear γ is a convolution of the surface mass density κ with the kernel $\mathcal{D}$. This relation is easily inverted in Fourier space to return the surface mass density in terms of a linear functional of the shear. Hence, if the shear can be observed from image distortions, the surface mass density can directly be obtained. Let the Fourier transform of $\kappa(\vec{\theta})$ be

$$\hat{\kappa}(\vec{l}) = \int_{\mathbb{R}^2} \mathrm{d}^2\theta\, \kappa(\vec{\theta})\, \exp(\mathrm{i}\vec{\theta}\cdot\vec{l}) \,. \tag{5.1}$$

The Fourier transform of the complex kernel $\mathcal{D}$ defined in (3.15, page 18) is

$$\hat{\mathcal{D}}(\vec{l}) = \pi \frac{\left(l_1^2 - l_2^2 + 2\mathrm{i} l_1 l_2\right)}{|\vec{l}|^2} \,. \tag{5.2}$$

Using the convolution theorem, eq. (3.15, page 18) can be written $\hat{\gamma}(\vec{l}) = \pi^{-1}\hat{\mathcal{D}}(\vec{l})\,\hat{\kappa}(\vec{l})$ for $\vec{l} \neq \vec{0}$. Multiplying both sides of this equation with $\tilde{\mathcal{D}}^*$ and using $\tilde{\mathcal{D}}\tilde{\mathcal{D}}^* = \pi^2$ gives

$$\hat{\kappa}(\vec{l}) = \pi^{-1}\hat{\gamma}(\vec{l})\,\hat{\mathcal{D}}^*(\vec{l}) \quad \text{for} \quad \vec{l} \neq \vec{0} \,, \tag{5.3}$$

and the convolution theorem leads to the final result

$$\begin{aligned} \kappa(\vec{\theta}) - \kappa_0 &= \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta'\, \mathcal{D}^*(\vec{\theta}-\vec{\theta}')\gamma(\vec{\theta}') \\ &= \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta'\, \Re\left[\mathcal{D}^*(\vec{\theta}-\vec{\theta}')\gamma(\vec{\theta}')\right] \end{aligned} \tag{5.4}$$

(Kaiser & Squires 1993). The constant κ_0 in (5.4) appears because a constant surface mass density does not cause any shear and is thus unconstrained by γ. The two expressions in (5.4) are equivalent because $\Im(\hat{\mathcal{D}}^*\hat{\gamma}) \equiv 0$, as can be shown from the Fourier transforms of equations (3.12, page 17). In applications, the second form of (5.4) should be used to ensure that κ is real. Relation (5.4) can either be applied to a case where all the sources are at the same redshift, in which case κ and γ are defined as in eqs. (3.7) and (3.12), or where the sources are distributed in redshift, because κ and γ are interpreted as convergence and shear for a hypothetical source at infinite redshift, as discussed in Sect. 4.3.2..

In the case of a weak lens ($\kappa \ll 1$, $|\gamma| \ll 1$), the shear map is directly obtained from observations, cf. (5.17). When inserted into (5.4), this map provides a parameter-free reconstruction of the surface mass density, apart from an overall additive constant. The importance of this result is obvious, as it provides us with a novel and simple method to infer the mass distribution in galaxy clusters.

There are two basic ways to apply (5.4) to observational data. Either, one can derive a shear map from averaging over galaxy images by calculating the local shear on a grid in $\vec{\theta}$-space, as described in Sect. 4.3.; or, one can replace the integral in (5.4) by a sum over galaxy images at positions $\vec{\theta}_i$,

$$\kappa(\vec{\theta}) = \frac{1}{n\pi}\sum_i \Re\left[\mathcal{D}^*(\vec{\theta}-\vec{\theta}_i)\,\varepsilon_i\right] \,. \tag{5.5}$$

Unfortunately, this estimate of κ has infinite noise (Kaiser & Squires 1993) because of the noisy sampling of the shear at the discrete background galaxy positions. Smoothing is therefore necessary to obtain estimators of κ with finite noise. The form of eq. (5.5) is preserved by smoothing, but the kernel $\mathcal{D}$ is modified to another kernel $\tilde{\mathcal{D}}$. In particular, Gaussian smoothing with smoothing length θ_s leads to

$$\tilde{\mathcal{D}}(\vec{\theta}) = \left[1 - \left(1 + \frac{|\vec{\theta}|^2}{\theta_s^2}\right) \exp\left(-\frac{|\vec{\theta}|^2}{\theta_s^2}\right)\right] \mathcal{D}(\vec{\theta}) \tag{5.6}$$

(Seitz & Schneider 1995a). The *rms* error of the resulting κ map is of order $\sigma_\varepsilon N^{-1/2}$, where N is the number of galaxy images per smoothing window, $N \sim n\pi\theta_s^2$. However, the errors will be strongly spatially correlated.

van Waerbeke (2000) showed that the covariance of a mass map obtained with the kernel (5.6) is

$$\mathrm{Cov}(\kappa(\vec{\theta}), \kappa(\vec{\theta}')) = \frac{\sigma_\varepsilon^2}{4\pi\theta_s^2 n} \exp\left(-\frac{|\vec{\theta} - \vec{\theta}'|^2}{2\theta_s^2}\right) . \tag{5.7}$$

Thus, the correlation extends to scales of order the smoothing scale θ_s (see also Lombardi & Bertin 1998b). Indeed, this result is surprising, as by reducing the smoothing scale, the correlation length of the noise can accordingly be reduced to small scales – although the surface mass density at each point depends on the galaxy ellipticities at all distances. It should be noted that the covariance in (5.7) is derived under the assumption of no lensing, $\gamma \equiv 0$. In the presence of a shear – the interesting situation of course – an additional effect contributes to the noise, namely that the galaxy images are not uniformly, but randomly distributed. This effect contributes shot-noise to the covariance, quadratically in γ (Schneider & Morales-Merino 2000). Therefore, whereas the estimate (5.5) with $\mathcal{D}$ replaced by $\tilde{\mathcal{D}}$ uses the observational data more directly than by first constructing a smoothed shear map and applying (5.4) to it, it turns out that the latter method yields a mass map which is less noisy than the estimate obtained from (5.5), because (5.5) contains the 'shot noise' from the random angular position of the galaxy images (Seitz & Schneider 1995a).

A lower bound to the smoothing length θ_s follows from the spatial number density of background galaxies, i.e. their mean separation. More realistically, a smoothing window needs to encompass several galaxies. In regions of strong shear signals, $N \sim 10$ may suffice, whereas mass maps in the outskirts of clusters where the shear is small may be dominated by noise unless $N \sim 100$. These remarks illustrate that a single smoothing scale across a whole cluster may be a poor choice. We shall return to this issue in Sect. 5.5.1., where improvements will be discussed.

Before applying the mass reconstruction formula (5.4) to real data, one should be aware of the following difficulties:

1. The integral in (5.4) extends over $\mathbb{R}^2$, while real data fields are relatively small (most of the applications shown in Sect. 5.4. are based on CCDs with side lengths of about 7 arc min). Since there is no information on the shear outside the data field, the integration has to be restricted to the field, which is equivalent to setting $\gamma = 0$ outside. This is done explicitly in (5.5). This cut-off in the integration leads to boundary artefacts in the mass reconstruction. Depending on the strength of the lens, its angular size relative to that of the data field, and its location within the data field, these boundary artefacts can be more or less severe. They are less important if the cluster is weak, small compared to the data field, and centred on it.
2. The shear is an approximate observable only in the limit of weak lensing. The surface mass density obtained by (5.4) is biased low in the central region of the cluster where the weak lensing assumption may not hold (and does not hold in those clusters which show giant arcs). Thus, if the inversion method is to be applied also to the inner parts of a cluster, the relation between γ and the observable δ has to be taken into account.
3. The surface mass density is determined by (5.4) only up to an additive constant. We demonstrate in the next subsection that there exists a slightly different general invariance transformation which is present in all mass reconstructions based solely on image shapes. However, this invariance transformation can be broken by including the magnification effect.

In the next three subsections, we shall consider points (1) and (2). In particular, we show that the first two problems can easily be cured. The magnification effects will be treated in Sect. 5.4..

5.2.2. Non–Linear Generalisation of the Inversion, and an Invariance Transformation

In this section, we generalise the inversion equation (5.4) to also account for strong lensing, i.e. we shall drop the assumption $\kappa \ll 1$ and $|\gamma| \ll 1$. In this case, the shear γ is no longer a direct observable, but at best the reduced shear g, or in general the distortion δ. In this case, the relation between κ and the observable becomes non-linear. Furthermore, we shall assume here that all sources are at the same redshift, so that the reduced shear is well-defined.

Consider first the case that the cluster is sub-critical everywhere, i.e. $\det\mathcal{A} > 0$ for all $\vec{\theta}$, which implies $|g(\vec{\theta})| < 1$. Then, the mean image ellipticity ε is an unbiased estimate of the local reduced shear, so that

$$\gamma(\vec{\theta}) = \left[1 - \kappa(\vec{\theta})\right] \langle\varepsilon\rangle(\vec{\theta}) , \tag{5.8}$$

where the field $\langle\varepsilon\rangle(\vec{\theta})$ is determined by the local averaging procedure described in Sect. 4.3.1.. Inserting this into (5.4) leads to an integral equation for $\kappa(\vec{\theta})$,

$$\kappa(\vec{\theta}) - \kappa_0 = \frac{1}{\pi} \int_{\mathbb{R}^2} \mathrm{d}^2\theta' \left[1 - \kappa(\vec{\theta}')\right] \Re\left[\mathcal{D}^*(\vec{\theta} - \vec{\theta}') \langle\varepsilon\rangle(\vec{\theta}')\right] , \tag{5.9}$$

(Seitz & Schneider 1995a), which is readily solved by iteration. Starting from $\kappa \equiv 0$, a first estimate of $\kappa(\vec{\theta})$ is obtained from (5.9), which after insertion into the right-hand side of (5.8) yields an update of $\gamma(\vec{\theta})$, etc. This iteration process converges quickly to the unique solution.

The situation becomes only slightly more complicated if critical clusters are included. We only need to keep track of $\det\mathcal{A}$ while iterating, because γ must be derived from $1/\langle\varepsilon\rangle^*$ rather than from $\langle\varepsilon\rangle$ where $\det\mathcal{A} < 0$. Hence, the local invariance between g and $1/g^*$ is broken due to non-local effects: A local jump from g to $1/g^*$ cannot be generated by any smooth surface mass density.

After a minor modification[9], this iteration process converges quickly. See Seitz & Schneider (1995a) for more details on this

[9] At points where $\kappa = 1$, $1/g^* = 0$ and $E(\varepsilon) = 0$, while γ remains finite. During the iteration, there will be points $\vec{\theta}$ where the field κ is very close to unity, but where $\langle\varepsilon\rangle$ is not necessarily small. This leads to large values of γ, which render the iteration unstable. However, this instability can easily be removed if a damping factor like $\left(1 + |\gamma^2(\vec{\theta}')|\right) \exp\left(-|\gamma^2(\vec{\theta}')|\right)$ is included in (5.4). This modification leads to fast convergence and affects the result of the iteration only very slightly.

method and for numerical tests done with a cluster mass distribution produced by a cosmological N-body simulation. It should have become clear that the non-linear inversion process poses hardly any additional problem to the mass reconstruction compared to the linear inversion (5.4).

This non-linear inversion still contains the constant κ_0, and so the result will depend on this unconstrained constant. However, in contrast to the linear (weak lensing) case, this constant does not correspond to adding a sheet of constant surface mass density. In fact, as can be seen from (5.9), the transformation

$$\begin{aligned} \kappa(\vec{\theta}) \to \kappa'(\vec{\theta}) &= \lambda\kappa(\vec{\theta}) + (1-\lambda) \quad \text{or} \\ \left[1-\kappa'(\vec{\theta})\right] &= \lambda\left[1-\kappa(\vec{\theta})\right] \end{aligned} \tag{5.10}$$

leads to another solution of the inverse problem for any value of $\lambda \neq 0$. Another and more general way to see this is that the transformation $\kappa \to \kappa'$ changes γ to $\gamma'(\vec{\theta}) = \lambda\gamma(\vec{\theta})$, cf. (3.15, page 18). Hence, the reduced shear $g = \gamma(1-\kappa)^{-1}$ is invariant under the transformation (5.10), so that the relation between intrinsic and observed ellipticity is unchanged under the *invariance transformation* (5.10). This is the mass-sheet degeneracy pointed out by Falco et al. (1985) in a different context. We thus conclude that the degeneracy due to the invariance transformation (5.10) cannot be lifted if only image shapes are used. However, the magnification transforms like

$$\mu'(\vec{\theta}) = \lambda^{-2}\mu(\vec{\theta}) \;, \tag{5.11}$$

so that the degeneracy can be lifted if magnification effects are taken into account (see Sect. 4.4.).

The invariance transformation leaves the critical curves of the lens mapping invariant. Therefore, even the location of giant luminous arcs which roughly trace the critical curves does not determine the scaling constant λ. In addition, the curve $\kappa = 1$ is invariant under (5.10). However, there are at least two ways to constrain λ. First, it is reasonable to expect that on the whole the surface mass density in clusters decreases with increasing separation from the cluster 'centre', so that $\lambda > 0$. Second, since the surface mass density κ is non-negative, upper limits on λ are obtained by enforcing this condition.

5.2.3. *Finite–Field Inversion Techniques*

We shall now turn to the problem that the inversion (5.4) in principle requires data on the whole sky, whereas the available data field is finite. A simple solution of this problem has been attempted by Seitz & Schneider (1995a). They extrapolated the measured shear field on the finite region $\mathcal{U}$ outside the data field, using a parameterised form for the radial decrease of the shear. From a sample of numerically generated cluster mass profiles, Bartelmann (1995a) showed that this extrapolation yields fairly accurate mass distributions. However, in these studies the cluster was always assumed to be isolated and placed close to the centre of the data field. If these two conditions are not met, the extrapolation can produce results which are significantly off. In order to remove the boundary artefacts inherent in applying (5.4) to a finite field, one should therefore aim at constructing an unbiased finite-field inversion method.

The basis of most finite-field inversions is a result first derived by Kaiser (1995). Equation (3.12, page 17) shows that shear and surface mass density are both given as second partial derivatives of the deflection potential ψ. After partially differentiating (3.12, page 17) and combining suitable terms we find

$$\nabla\kappa = \begin{pmatrix} \gamma_{1,1}+\gamma_{2,2} \\ \gamma_{2,1}-\gamma_{1,2} \end{pmatrix} \equiv \vec{u}_\gamma(\vec{\theta}) \;. \tag{5.12}$$

The gradient of the surface mass density can thus be expressed by the first derivatives of the shear, hence $\kappa(\vec{\theta})$ can be determined, up to an additive constant, by integrating (5.12) along appropriately selected curves. This can be done in the weak lensing case where the observed smoothed ellipticity field $\langle\epsilon\rangle(\vec{\theta})$ can be identified with γ, and $\vec{u}_\gamma(\vec{\theta})$ can be constructed by finite differencing. If we insert $\gamma = (1-\kappa)\,g$ into (5.12), we find after some manipulations

$$\begin{aligned} \nabla K(\vec{\theta}) &= \frac{-1}{1-g_1^2-g_2^2}\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}\begin{pmatrix} g_{1,1}+g_{2,2} \\ g_{2,1}-g_{1,2} \end{pmatrix} \\ &\equiv \vec{u}_g(\vec{\theta}) \;, \end{aligned} \tag{5.13}$$

where

$$K(\vec{\theta}) \equiv \ln[1-\kappa(\vec{\theta})] \;. \tag{5.14}$$

Hence, using the smoothed ellipticity field $\langle\epsilon\rangle(\vec{\theta})$ as an unbiased estimator for $g(\vec{\theta})$, and assuming a sub-critical cluster, one can obtain the vector field $\vec{u}_g(\vec{\theta})$ by finite differencing, and thus determine $K(\vec{\theta})$ up to an additive constant from line integration, or, equivalently, $1-\kappa(\vec{\theta})$ up to an overall multiplicative constant. This is again the invariance transformation (5.10).

In principle, it is now straightforward to obtain $\kappa(\vec{\theta})$ from the vector field $\vec{u}_\gamma(\vec{\theta})$, or $K(\vec{\theta})$ from $\vec{u}_g(\vec{\theta})$, simply by a line integration of the type

$$\kappa(\vec{\theta},\vec{\theta}_0) = \kappa(\vec{\theta}_0) + \int_{\vec{\theta}_0}^{\vec{\theta}} \mathrm{d}\vec{l}\cdot\vec{u}_\gamma(\vec{l}) \;, \tag{5.15}$$

where $\vec{l}$ is a smooth curve connecting $\vec{\theta}$ with $\vec{\theta}_0$. If $\vec{u}_\gamma$ is a gradient field, as it ideally is, the resulting surface mass density is independent of the choice of the curves $\vec{l}$. However, since $\vec{u}_\gamma$ is obtained from noisy data (at least the noise resulting from the intrinsic ellipticity distribution), it will in general not be a gradient field, so that (5.12) has no solution. Therefore, the various line integration schemes proposed (Schneider 1995, Kaiser et al. 1994, Bartelmann 1995a) yield different results.

Realising that eq. (5.12) has no exact solution for an observed field $\vec{u}_\gamma$, we wish to find a mass distribution $\kappa(\vec{\theta})$ which satisfies (5.12) 'best'. In general, $\vec{u}_\gamma$ can be split into a gradient field and a curl component, but this decomposition is not unique. However, as pointed out in Seitz & Schneider (1996), since the curl component is due to noise, its mean over the data field is expected to vanish. Imposing this condition, which determines the decomposition uniquely, they showed that

$$\kappa(\vec{\theta}) - \bar{\kappa} = \int_{\mathcal{U}} \mathrm{d}^2\theta'\,\vec{H}(\vec{\theta}',\vec{\theta})\cdot\vec{u}_\gamma(\vec{\theta}') \;, \tag{5.16}$$

where $\bar{\kappa}$ is the average of $\kappa(\vec{\theta})$ over the data field $\mathcal{U}$, and the kernel $\vec{H}$ is the gradient of a scalar function which is determined through a von Neumann boundary value problem, with singular source term. This problem can be solved analytically for circular and rectangular data fields, as detailed in the Appendix of Seitz & Schneider (1996). If the data field has a more complicated geometry, an analytic solution is no longer possible, and the boundary value problem with a singular source term cannot be solved numerically.

An alternative method starts with taking the divergence of (5.12) and leads to the new boundary value problem,

$$\nabla^2\kappa = \nabla\cdot\vec{u}_\gamma \quad \text{with} \quad \vec{n}\cdot\nabla\kappa = \vec{n}\cdot\vec{u}_\gamma \quad \text{on} \quad \partial\mathcal{U} \;, \tag{5.17}$$

where $\vec{n}$ is the outward-directed normal on the boundary of $\mathcal{U}$. As shown in Seitz & Schneider (1998), eqs. (5.16) and (5.17) are equivalent. An alternative and very elegant way to derive (5.17) has been found by Lombardi & Bertin (1998b). They noticed that the 'best' approximation to a solution of (5.12) minimises the 'action'

$$\int_{\mathcal{U}} \mathrm{d}^2\theta \, |\nabla\kappa(\vec{\theta}) - \vec{u}_\gamma(\vec{\theta})|^2 \; . \tag{5.18}$$

Euler's equations of the variational principle immediately reproduce (5.17). This von Neumann boundary problem is readily solved numerically, using standard numerical techniques (see Sect. 19.5 of Press et al. 1986). Lombardi & Bertin (1999a) proposed a direct method for solving the variational principle (5.18) which, for rectangular fields, is equivalent to Fourier methods for the solution of the Neumann problem (5.17).

A comparison between these different finite-field inversion equations was performed in Seitz & Schneider (1996) and in Squires & Kaiser (1996) by numerical simulations. Of all the inversions tested, the inversion (5.17) performs best on all scales (Seitz & Schneider 1996; Fig. 6 of Squires & Kaiser 1996). Indeed, Lombardi & Bertin (1998b) showed analytically that the solution of eq. (5.17) provides the best unbiased estimate of the surface mass density. The relations (5.15) through (5.18) can be generalised to the non-weak case by replacing κ with K and $\vec{u}_\gamma$ with $\vec{u}_g$.

5.2.4. *Accounting for a redshift distribution of the sources*

We now describe how the preceding mass reconstructions must be modified if the sources have a broad redshift distribution. In fact, only minor modifications are needed. The relation $\langle\varepsilon\rangle = g$ for a single source redshift is replaced by eq. (4.28), which gives an estimate for the shear in terms of the mean image ellipticities and the surface mass density. This relation can be applied iteratively:

Begin with $\kappa^{(0)} = 0$; then, eq. (4.28) yields a first guess for the shear $\gamma^{(1)}(\vec{\theta})$ by setting $\gamma = 0$ on the right-hand side. From (5.16), or equivalently by solving (5.17), the corresponding surface mass density $\kappa^{(1)}(\vec{\theta})$ is obtained. Inserting $\kappa^{(1)}$ and $\gamma^{(1)}$ on the right-hand side of eq. (4.28), a new estimate $\gamma^{(2)}(\vec{\theta})$ for the shear is obtained, and so forth.

This iteration process quickly converges. Indeed, the difficulty mentioned in footnote 9 (page 33) no longer occurs since the critical curves and the curve(s) $\kappa = 1$ are effectively smeared out by the redshift distribution, and so the iteration converges even faster than in the case of a single source redshift.

Since $\kappa^{(n)}$ is determined only up to an additive constant for any $\gamma^{(n)}$, the solution of the iteration depends on the choice of this constant. Hence, one can obtain a one-parameter family of mass reconstructions, like in (5.10). However, the resulting mass-sheet degeneracy can no longer be expressed analytically due to the complex dependence of (4.28) on κ and γ. In the case of weak lensing, it corresponds to adding a constant, as before. An approximate invariance transformation can also be obtained explicitly for mildly non-linear clusters with $\kappa \lesssim 0.7$ and $\det\mathcal{A} > 0$ everywhere. In that case, eq. (4.29) holds approximately, and can be used to show (Seitz & Schneider 1997) that the invariance transformation takes the form

$$\kappa(\vec{\theta}) \to \kappa'(\vec{\theta}) = \lambda\,\kappa(\vec{\theta}) + \frac{(1-\lambda)\langle Z\rangle}{\langle Z^2\rangle} \; . \tag{5.19}$$

In case of a single redshift z_s, such that $Z(z_\mathrm{s}) = \langle Z\rangle$, this transformation reduces to (5.10) for $\langle Z\rangle\kappa$.

We point out that the invariance transformation (5.19) in the case of a redshift distribution of sources is of different nature than that for a single source redshift. In the latter case, the reduced shear $g(\vec{\theta})$ is invariant under the transformation (5.10). Therefore, the probability distribution of the observed galaxy ellipticities is invariant, since it involves only the intrinsic ellipticity distribution and g. For a redshift distribution, the invariance transformation keeps the mean image ellipticities invariant, but the probability distributions are changed. Several strategies were explored in Seitz & Schneider (1997) to utilise this fact for breaking the invariance transformation. (See also Lombardi & Bertin 1999b and Gautret et al. 2000.) While possible in principle, the corresponding effect on the observed ellipticity distribution is too small for this approach to be feasible with existing data.

5.2.5. *Breaking the Mass-Sheet Degeneracy*

Equation (5.11) shows that the invariance transformation (5.10) affects the magnification. Hence, the degeneracy can be lifted with magnification information. As discussed in Sect. 4.4., two methods to obtain magnification information have been proposed. Detections of the number-density effect have so far been reported for two clusters (Cl 0024+16, Fort et al. 1997; Abell 1689, Taylor et al. 1998). Whereas the information provided by the number density effect is less efficient than shear measurements (see Sect. 4.4.3.), these two clusters appear to be massive enough to allow a significant detection. In fact, Taylor et al. (1998) obtained a two-dimensional mass reconstruction of the cluster A 1689 from magnification data.

In the case of weak lensing, and thus small magnifications, the magnification can locally be translated into a surface mass density – see (4.44). In general, the relation between μ and κ is non-local, since μ also depends on the shear. Various attempts to account for this non-locality have been published (van Kampen 1998, Dye & Taylor 1998). However, it must be noted that the surface mass density cannot be obtained from magnification alone since the magnification also depends on the shear caused by matter outside the data field. In practice, if the data field is sufficiently large and no mass concentration lies close to but outside the data field, the mass reconstruction obtained from magnification can be quite accurate.

In order to break the mass-sheet degeneracy, it suffices in principle to measure one value of the magnification: Either the magnification at one location in the cluster, or the average magnification over a region. We shall see later in Sect. 5.5.1. how local magnification information can be combined with shear measurements. Doing it the naive way, expressing κ in terms of μ and γ, is a big waste of information: Since there is only one independent scalar field (namely the deflection potential ψ) describing the lens, one can make much better use of the measurements of γ and μ than just combining them locally; the relation between them should be used to reduce the error on κ.

5.2.6. *Accuracy of cluster mass determinations*

The mass-sheet degeneracy fundamentally limits the accuracy with which cluster masses can be determined from shear measurements if no additional assumptions are introduced. Furthermore, cosmologists are traditionally interested in the masses of clusters inside spherical volumes (e.g., inside the virial radius), whereas lensing measures the mass in cylinders, i.e., the projected mass. On the other hand, cosmological simulations show that cluster mass profiles are quite similar in shape (e.g., Navarro et al. 1996b). Assuming such a universal density profile, both of these effects can approximately be accounted for.

The relation between projected mass within the virial radius and that inside a sphere with virial radius has been investigated by Reblinsky & Bartelmann (1999a) and Metzler et al. (1999), using numerical cluster simulations. The ratio of these two masses is by definition ≥ 1, but as these authors show, this ratio can be larger than unity by several tens of a per cent, due to projection of additional mass in front of or behind the cluster proper. As clusters are preferentially located inside filaments, the largest deviations occur when the filament is oriented along the line-of-sight to the cluster. The amplitude of this effect decreases with higher cluster masses. The projection bias is of interest only when comparing lensing masses with cosmological predictions of spherical masses. However, at least when cosmological predictions are derived from numerical simulations, one can equally well extract the projected masses; the projection bias therefore does not affect the use of cluster mass estimates from lensing for cosmology.

Using cluster mass models obtained from N-body simulations, Brainerd et al. (1999) showed that, when the observed shear signal is related to the mass using the isothermal relation (cf. Sect. 3.1.5.) $M_3(<r) = 4r^2\gamma(r)\sigma_{\rm cr}$ for the mass inside spheres of radius r, or $M_2(<r) = 2\pi r^2\gamma(r)\Sigma_{\rm cr}$ for the projected mass inside $r = D_{\rm d}\theta$, fairly accurate masses of clusters can be derived from weak lensing. In particular, the virial masses of clusters can be determined with high accuracy, provided the shear measurements extend to such large distances. Whereas most of the previous weak-lensing cluster studies, as described in Sect. 5.4., do not cover such large an area, the upcoming wide-field imaging cameras will allow one to do this in the near future. Nevertheless, the projection bias needs to be kept in mind when masses of clusters are quoted from weak-lensing analyses using relatively small angular fields.

5.3. *Aperture Mass and Multipole Measures*

Having reconstructed the mass distribution, we can estimate the local dispersion of κ (e.g., Lombardi & Bertin 1998b, van Waerbeke 2000); cf. eq. (5.7). However, the errors at different points are strongly correlated, and so it makes little sense to attach an error bar to each point of the mass map. Although mass maps contain valuable information, it is sometimes preferable to reduce them to a small set of numbers such as the mass-to-light ratio, or the correlation coefficient between the mass map and the light distribution. One of the quantities of interest is the total mass inside a given region. As became clear in the last section, this quantity by itself cannot be determined from observed image ellipticities due to the invariance transformation. But a quantity related to it,

$$\zeta(\vec{\theta};\vartheta_1,\vartheta_2) \equiv \bar{\kappa}(\vec{\theta};\vartheta_1) - \bar{\kappa}(\vec{\theta};\vartheta_1,\vartheta_2)\;, \tag{5.20}$$

the difference between the mean surface mass densities in a circle of radius ϑ_1 around $\vec{\theta}$ and in an annulus of inner and outer radii ϑ_1 and ϑ_2, respectively, can be determined in the weak-lensing case, since then the invariance transformation corresponds to an additive constant in κ which drops out of (5.20). We show in this section that quantities like (5.20) can directly be obtained from the image ellipticities without the need for a two-dimensional mass map. In Sect. 5.3.1., we derive a generalised version of (5.20), whereas we consider the determination of mass multipoles in Sect. 5.3.2.. The prime advantage of all these aperture measures is that the error analysis is relatively straightforward.

5.3.1. *Aperture Mass Measures*

Generally, aperture mass measures are weighted integrals of the local surface mass density,

$$M_{\rm ap}(\vec{\theta}_0) = \int {\rm d}^2\theta\,\kappa(\vec{\theta})\,U(\vec{\theta}-\vec{\theta}_0)\;, \tag{5.21}$$

with weight function $U(\vec{\theta})$. Assume now that the weight function is constant on self-similar concentric curves. For example, the ζ-statistics (5.20), introduced by Kaiser (1995), is of the form (5.21), with a weight function that is constant on circles, $U(\vartheta) = (\pi\vartheta_1^2)^{-1}$ for $0 \leq \vartheta \leq \vartheta_1$, $U(\vartheta) = [\pi(\vartheta_2^2 - \vartheta_1^2]^{-1}$ for $\vartheta_1 < \vartheta \leq \vartheta_2$, and zero otherwise.

Let the shape of the aperture be described by a closed curve $\vec{c}(\lambda)$, $\lambda \in I$, where I is a finite interval, such that $\vec{c}\times\dot{\vec{c}} \equiv c_1\dot{c}_2 - c_2\dot{c}_1 > 0$ for all $\lambda \in I$. We can then uniquely define a new coordinate system (b,λ) by choosing a centre $\vec{\theta}_0$ and defining $\vec{\theta} = \vec{\theta}_0 + b\vec{c}(\lambda)$. The weight function should be constant on the curves $\vec{c}(\lambda)$ so that it depends only on b. In the new coordinate system, (5.21) reads

$$M_{\rm ap}(\vec{\theta}_0) = \int_0^\infty {\rm d}b\,b\,U(b)\oint_I {\rm d}\lambda\,\vec{c}\times\dot{\vec{c}}\,\kappa[\vec{\theta}_0 + b\vec{c}(\lambda)]\;, \tag{5.22}$$

where the factor $b\,\vec{c}\times\dot{\vec{c}}$ is the Jacobian determinant of the coordinate transformation. Equation (5.22) can now be transformed in three steps; first, by a partial integration with respect to b; second, by replacing partial derivatives of κ with partial derivatives of γ using eq. (5.12); and third by removing partial derivatives of γ in another partial integration. In carrying out these steps, we assume that the weight function is compensated,

$$\int {\rm d}b\,b\,U(b) = 0\;. \tag{5.23}$$

Introducing

$$Q(b) \equiv \frac{2}{b^2}\int_0^b {\rm d}b'\,b'\,U(b') - U(b) \tag{5.24}$$

and writing the curve $\vec{c}$ in complex notation, $C(\lambda) = c_1(\lambda) + {\rm i}c_2(\lambda)$, leads to the final result (Schneider & Bartelmann 1997)

$$M_{\rm ap}(\vec{\theta}_0) = \int {\rm d}^2\theta\,Q[b(\vec{\theta})]\,\frac{\Im[\gamma(\vec{\theta})C^*\dot{C}^*]}{\Im[C^*\dot{C}]}\;, \tag{5.25}$$

where the argument λ of C is to be evaluated at position $\vec{\theta} = \vec{\theta}_0 + b\vec{c}(\lambda)$.[10] The numerator in the final term of (5.25) projects out a particular component of the shear, whereas the denominator is part of the Jacobian of the coordinate transformation. The constraint (5.23) assures that an additive constant in κ does not affect $M_{\rm ap}$. The expression (5.25) has several nice properties which render it useful:

1. If the function $U(b)$ is chosen such that it vanishes for $b > b_2$, then from (5.23) and (5.24), $Q(b) = 0$ for $b > b_2$. Thus, the aperture mass can be derived from the shear in a finite region.

2. If $U(b) =$ const for $0 \leq b \leq b_1$, then $Q(b) = 0$ in that interval. This means that the aperture mass can be determined solely from the shear in an annulus $b_1 < b < b_2$. This has two advantages which are relevant in practice. First, if the aperture is centred on a cluster, the bright central cluster galaxies may prevent the detection of a large number of faint background galaxies there, so that the shear in the central part of the cluster may be difficult to measure. In that case it is still possible to determine the total mass inside the cluster core using (5.25) with an appropriately chosen weight function U. Second, although in general the shear

[10]There are of course other ways to derive (5.25), e.g. by inserting (5.4) into (5.21). See Squires & Kaiser (1996) for a different approach using Gauss's law.

cannot be determined directly from the image ellipticities [but only the reduced shear $\gamma(1-\kappa)^{-1}$], we can choose the size b_1 of the inner boundary of the annulus sufficiently large that $\kappa \ll 1$ in the annulus, and then $\gamma \approx g$ is an accurate approximation. Hence, in that case the mean image ellipticity directly yields an estimate of the shear. Then, the integral (5.25) can be transformed into a sum over galaxy images lying in the annulus, yielding $M_{\rm ap}$ directly in terms of the observables. This in turn has the great advantage that an error analysis of $M_{\rm ap}$ is fairly simple.

We consider circular apertures as an example, for which $(b,\lambda) = (\theta,\varphi)$ and $C(\varphi) = \exp(i\varphi)$. Then, $\Im(C^*\dot{C}) = 1$, and

$$\begin{aligned}\Im(\gamma C^* \dot{C}^*) &= \gamma_{\rm t}(\vec{\theta};\vec{\theta}_0) := -\left[\gamma_1 \cos(2\varphi) + \gamma_2 \sin(2\varphi)\right] \\ &= -\Re[\gamma(\vec{\theta}+\vec{\theta}_0){\rm e}^{-2{\rm i}\varphi}] \,, \end{aligned} \tag{5.26}$$

where we have defined the *tangential component* $\gamma_{\rm t}$ of the shear relative to the point $\vec{\theta}_0$. Hence, for circular apertures (5.25) becomes

$$M_{\rm ap}(\vec{\theta}_0) = \int {\rm d}^2\theta\, Q(|\vec{\theta}|)\, \gamma_{\rm t}(\vec{\theta};\vec{\theta}_0) \tag{5.27}$$

(Kaiser et al. 1994; Schneider 1996b). The ζ-statistics (5.20) is obtained from (5.27) by setting $Q(\theta) = \vartheta_2^2\theta^{-2}\left[\pi(\vartheta_2^2-\vartheta_1^2)\right]^{-1}$ for $\vartheta_1 \le \theta \le \vartheta_2$ and $Q(\theta) = 0$ otherwise, so that

$$\zeta(\vec{\theta}_0;\vartheta_1,\vartheta_2) = \frac{\vartheta_2^2}{\left[\pi(\vartheta_2^2-\vartheta_1^2)\right]} \int {\rm d}^2\theta\, \frac{\gamma_{\rm t}(\vec{\theta};\vec{\theta}_0)}{|\vec{\theta}|^2} \,, \tag{5.28}$$

where the integral is taken over the annulus $\vartheta_1 \le \theta \le \vartheta_2$.

For practical purposes, the integral in (5.27) is transformed into a sum over galaxy images. Recalling that ε is an estimator for γ in the weak-lensing case, and that the weight function can be chosen to avoid the strong-lensing regime, we can write

$$M_{\rm ap}(\vec{\theta}_0) = \frac{1}{n}\sum_i Q(|\vec{\theta}_i - \vec{\theta}_0|)\, \varepsilon_{{\rm t}i}(\vec{\theta}_0) \,, \tag{5.29}$$

where we have defined, in analogy to $\gamma_{\rm t}$, the tangential component $\varepsilon_{{\rm t}i}$ of the ellipticity of an image at $\vec{\theta}_i$ relative to the point $\vec{\theta}_0$ by

$$\varepsilon_{{\rm t}i} = -\Re(\varepsilon\, {\rm e}^{-2{\rm i}\varphi}) \,, \tag{5.30}$$

φ is the polar angle of $\vec{\theta}-\vec{\theta}_0$, and n is the number density of galaxy images. The *rms* dispersion $\sigma(M_{\rm ap})$ of $M_{\rm ap}$ in the case of no lensing is found from the (two-dimensional) dispersion σ_ε of the intrinsic ellipticity of galaxies,

$$\sigma(M_{\rm ap}) = \frac{\sigma_\varepsilon}{2^{1/2}n}\left[\sum_i Q^2(|\vec{\theta}_i - \vec{\theta}_0|)\right]^{1/2} \,. \tag{5.31}$$

The *rms* dispersion in the presence of lensing will deviate only weakly from $\sigma(M_{\rm ap})$ as long as the assumption of weak lensing in the annulus is satisfied. Hence, $\sigma(M_{\rm ap})$ can be used as an error estimate for the aperture mass and as an estimate for the signal-to-noise ratio of a mass measurement.

This opens the interesting possibility to search for (dark) mass concentrations using the aperture mass (Schneider 1996b). Consider a weight function U with the shape of a Mexican hat, and a data field $\mathcal{U}$ on which apertures of angular size θ can be placed. For each aperture position, one can calculate $M_{\rm ap}$ and the dispersion. The dispersion can be obtained either from the analytical formula (5.31), or it can be obtained directly from the data, by randomising the position angles of all galaxy images within the aperture. The dispersion can be obtained from many realisations of this randomisation process. Large values of $M_{\rm ap}$ will be obtained for mass concentrations whose characteristic size and shape is close to that of the chosen filter function U. Thus, by varying the size θ of the filter, different mass concentrations will preferentially be selected. The aperture mass is insensitive to mass concentrations of much smaller and much larger angular scales than the filter size.

We have considered in Sect. 4.5. the signal-to-noise ratio for the detection of a singular isothermal sphere from its weak lensing effect. The estimate (4.54) was obtained by an optimal weighting scheme for this particular mass distribution. Since real mass concentrations will deviate from this profile, and also from the assumed symmetry, the filter function U should have a more generic shape. In that case, the S/N will have the same functional behaviour as in (4.54), but the prefactor depends on the exact shape of U. For the filter function used in Schneider (1996b), S/N is about 25% smaller than in (4.54). Nevertheless, one expects that the aperture-mass method will be sensitive to search for intermediate-redshift halos with characteristic velocity dispersions above $\sim 600\,{\rm km\,s^{-1}}$.

This expectation has been verified by numerical simulations, which also contained larger and smaller scale mass perturbations. In addition, a detailed strong-lensing investigation of the cluster MS 1512+62 has shown that its velocity dispersion is very close to $\sim 600\,{\rm km\,s^{-1}}$, and it can be seen from the weak-lensing image distortion alone with very high significance (Seitz et al. 1998b), supporting the foregoing quantitative prediction. Thus, this method appears to be a very promising way to obtain a *mass-selected* sample of halos which would be of great cosmological interest (cf. Reblinsky & Bartelmann 1999b). We shall return to this issue in Sect. 6.7.2..

5.3.2. *Aperture Multipole Moments*

Since it is possible to express the weighted mass within an aperture as an integral over the shear, with the advantage that in the weak lensing regime this integral can be replaced by a sum over galaxy ellipticities, it is natural to ask whether a similar result holds for multipole moments of the mass. As shown in Schneider & Bartelmann (1997), this is indeed possible, and we shall briefly outline the method and the result.

Consider a circular aperture[11] centred on a point $\vec{\theta}_0$. Let $U(|\vec{\theta}|)$ be a radial weight function, and define the n-th multipole moment by

$$Q^{(n)} \equiv \int_0^\infty {\rm d}\theta\, \theta^{n+1}\, U(\theta) \int_0^{2\pi} {\rm d}\varphi\, {\rm e}^{n{\rm i}\varphi}\, \kappa(\vec{\theta}_0+\vec{\theta}) \,. \tag{5.32}$$

This can be replaced by an integral over γ in two ways: (5.32) can be integrated by parts with respect to φ (for $n \neq 0$), or with respect to θ, again utilising (5.12). The resulting expressions are assumed to contain no boundary terms, which restricts the choice for the weight function $U(\theta)$. The remaining integrals then contain partial derivatives of κ with respect to φ and θ, respectively. Writing (5.12) in polar coordinates, these partial derivatives can be replaced by partial derivatives of the shear components with respect to φ and θ. Integrating those by parts with respect to the appropriate coordinate, and enforcing vanishing boundary terms, we find two different expressions for the $Q^{(n)}$:

$$Q^{(n)}_{\varphi,\theta} = \int {\rm d}^2\theta\, q^{(n)}_{\varphi,\theta}(\vec{\theta})\, \gamma(\vec{\theta}_0+\vec{\theta}) \,. \tag{5.33}$$

[11] The method is not restricted to circular apertures, but this case will be most relevant for measuring multipole moments.

The two expressions for $q^{(n)}$ are formally very different, although it can be shown that the resulting two expressions for $Q^{(n)}$ are equivalent. The two very different equations for the same result are due to the fact that the two components of the shear γ are not mutually independent, which was not used in the derivation of (5.33).

We now have substantial freedom to choose the weight function and to select one of the two expressions for $Q^{(n)}$, or even to take a linear combination of them. We note the following interesting examples:

1. The weight function $U(\theta)$ can be chosen to vanish outside an annulus, to be piece-wise differentiable, and to be zero on the inner and outer boundary of the annulus. The $Q^{(n)}$ for $n \neq 0$ can then be expressed as integrals of the shear over the annulus, with no further restrictions on U. In particular, $U(\theta)$ does not need to be a compensated weight function.

2. $U(\theta)$ can be a piece-wise differentiable weight function which is constant for $\theta \leq \theta_1$, and decreases smoothly to zero at $\theta = \theta_2 > \theta_1$. Again, $Q^{(n)}$ for $n \neq 0$ can be expressed as an integral of the shear in the annulus $\theta_1 \leq \theta \leq \theta_2$. Hence, as for the aperture mass, multipole moments in the inner circle can be probed with the shear in the surrounding annulus.

3. One can choose, for $n > 2$, a piece-wise differentiable weight function $U(\theta)$ which behaves like θ^{-2n} for $\theta > \theta_2$ and decreases to zero at $\theta = \theta_1 < \theta_2$. In that case, the multipole moments of the matter outside an annulus can be probed with data inside the annulus.

For practical applications, the integral in (5.33) is replaced by a sum over galaxy ellipticities. The dispersion of this sum is easily obtained in the absence of lensing, with an expression analogous to (5.31). Therefore, the signal-to-noise ratio for the multipole moments is easily defined, and thus also the significance of a multipole-moment detection.

5.4. Application to Observed Clusters

Soon after the parameter-free two-dimensional mass reconstruction was suggested by Kaiser & Squires (1993), their method was applied to the cluster MS 1224 (Fahlman et al. 1994). Since then, several groups have used it to infer the mass profiles of clusters. In parallel to this, several methods have been developed to measure the shear from CCD data, accounting for PSF smearing and anisotropy, image distortion by the telescope, noise, blending etc. – see the discussion in Sect. 4.6. We will now summarise and discuss several of these observational results.

Tyson et al. (1990) made the first attempt to constrain the mass distribution of a cluster from a weak-lensing analysis. They discovered a statistically significant tangential alignment of faint galaxy images relative to the centre of the clusters A 1689 and Cl 1409+52. Their "lens distortion map" obtained from the image ellipticities yields an estimate of the mass distribution in these clusters. A detailed analysis of their method is given in Kaiser & Squires (1993). From a comparison with numerical simulations, Tyson et al. showed that the best isothermal sphere model for the clusters has a typical velocity dispersion of $\sigma_v \sim 1300 \pm 200\,\mathrm{km\,s^{-1}}$ for both clusters. In particular, their analysis showed that diffuse dark matter in the cluster centres is needed to account for the observed image distortions.

The inversion method developed by Kaiser & Squires (1993) provided a systematic approach to reconstruct the mass distribution in clusters. It was first applied to the cluster MS 1224+20 (Fahlman et al. 1994) at redshift $z_\mathrm{d} = 0.33$, which had been selected for its high X-ray luminosity. Their square data field with side-length $\sim 14'$ was composed of several exposures, most of them with excellent seeing. They estimated the shear from image ellipticities, corrected for the PSF anisotropy, and applied a correction factor f as defined in Sect. 4.6.1.. They found $f \sim 1.5$ in simulations, in very good agreement with Wilson et al. (1996). The resulting shear pattern, obtained from 2147 galaxy images, clearly shows a circular pattern around the cluster centre as defined by the centroid of the optical and X-ray light. Using the Kaiser & Squires reconstruction method (5.8), Fahlman et al. produced maps of the dimension-less surface mass density $\kappa(\vec\theta)$, both by taking all galaxy images into account, and after splitting the galaxy sample into a 'brighter' and 'fainter' sample of roughly equal size. Although differing in detail, the resulting mass show an overall similarity. In particular, the position of the mass centre is very similar in all maps.

Fahlman et al. applied the aperture mass method to determine the cluster mass – see (5.21) and (5.29) – in an annulus centred on the cluster centre with inner radius $\vartheta_1 = 2\overset{'}{.}76$ and an outer radius such that the annulus nearly fits into their data field. The lower limit to the mean surface mass density in the annulus is $\bar\kappa(2\overset{'}{.}76) \geq \zeta = 0.06 \pm 0.013$. To convert this into an estimate of the physical surface mass density and the total mass inside the aperture, the mean distance ratio $D_\mathrm{ds}/D_\mathrm{s}$ for the galaxy population has to be estimated, or equivalently the mean value of Z as defined after (5.19).

While the redshift distribution is known statistically for the brighter sub-sample from redshift surveys, the use of the fainter galaxies requires an extrapolation of the galaxy redshifts. From that, Fahlman et al. estimated the mass within a cylinder of radius $\vartheta_1 = 2\overset{'}{.}76$, corresponding to $0.48h^{-1}$ Mpc for an Einstein-de Sitter cosmology, to be $\sim 3.5 \times 10^{14} h^{-1} M_\odot$. This corresponds to a mass-to-light ratio (in solar units) of $M/L \sim 800\,h$. Carlberg et al. (1994) obtained 75 redshifts of galaxies in the cluster field, of which 30 are cluster members. From their line-of-sight velocity dispersion, the cluster mass can be estimated by a virial analysis. The resulting mass is lower by a factor ~ 3 than the weak-lensing estimate. The mass-to-light ratio from the virial analysis is much closer to typical values in lower-redshift clusters like Coma, which has $M/L \approx 270\,h^{-1}$. The high mass estimate of this cluster was recently confirmed in a completely independent study by Fischer (1999).

The origin of this large apparent discrepancy is not well understood yet, and several possibilities are discussed in Kaiser et al. (1994). It should be pointed out that lensing measures the total mass inside a cone, weighted by the redshift-dependent factor $D_\mathrm{d}D_\mathrm{ds}/D_\mathrm{s}$, and hence the lensing mass estimate possibly includes substantial foreground and background material. While this may cause an overestimate of the mass, it is quite unlikely to cause an overestimate of the mass-to-light ratio of the total material inside the cone. Foreground material will contribute much more strongly to the light than to the measured mass, and additional matter behind the cluster will not be very efficient as a lens. The uncertainty in the redshift distribution of the faint galaxies translates into an uncertainty in the mass. However, all background galaxies would have to be put at a redshift ~ 4 to explain the mass discrepancy, while redshift surveys show that the brighter sub-sample of Fahlman et al. has a mean redshift below unity. The mass estimate is only weakly dependent on the assumed cosmological model. On the other hand, the light distribution of the cluster MS 1224 is not circular, and it cannot be excluded that this cluster is not in virial equilibrium.

Two images of the cluster Cl 0024+17 were analysed by Bonnet et al. (1994). One was centred on the cluster itself and yielded the shear in the inner part of the cluster. The second image was off-centred by several arc minutes and allowed, for the first time, a shear measurement out to large radial distances.

They detected a clear shear signal out to distances $\gtrsim 1h^{-1}$ Mpc. In addition, they found an apparent distortion of the nearly circular shear pattern from the cluster which is most directly interpreted as a mass concentration. However, it does not show an obvious concentration of galaxies. In fact, an X-ray observation of this cluster reveals a weak X-ray source close to the position where the mass concentration was seen in the shear map (Soucail et al. 2000), although with marginal significance. This cluster (at $z = 0.39$) hosts a giant arc system and has an Einstein radius of $\sim 30''$; together with the redshift of $z = 1.675$ of the arc (Broadhurst et al. 1999), this indicates that the cluster is indeed very massive. Despite that, the cluster is a comparatively faint and cool X-ray source, indicating a clear and interesting discrepancy between mass estimates from the X-rays and both strong and weak lensing.

Squires et al. (1996a) compared the mass profiles derived from weak lensing data and the X-ray emission of the cluster A 2218. Under the assumption that the hot X-ray-emitting intracluster gas is in hydrostatic equilibrium between gravity and thermal pressure support, the mass profile of the cluster can be constrained. The reconstructed mass map qualitatively agrees with the optical and X-ray light distributions. Using the aperture mass estimate, a mass-to-light ratio of $M/L = (440 \pm 80)\,h$ in solar units is found. The radial mass profile appears to be flatter than isothermal. Within the error bars, it agrees with the mass profile obtained from the X-ray analysis, with a slight indication that at large radii the lensing mass is larger than the mass inferred from X-rays.

Abell 2218 also contains a large number of arcs and multiply-imaged galaxies which have been used by Kneib et al. (1996) to construct a detailed mass model of the cluster's central region. In addition to the main mass concentration, there is a secondary clump of cluster galaxies whose effects on the arcs is clearly visible. The separation of these two mass centres is $67''$. Whereas the resolution of the weak lensing mass map as obtained by Squires et al. is not sufficient to reveal a distinct secondary peak, the elongation of the central density contours extend towards the secondary galaxy clump.

General agreement between the reconstructed mass map and the distribution of cluster galaxies and X-ray emission has also been found for the two clusters Cl 1455+22 ($z = 0.26$) and Cl 0016+16 ($z = 0.55$) by Smail et al. (1995a). Both are highly X-ray luminous clusters in the *Einstein* Extended Medium Sensitivity Survey (EMSS; Stocke et al. 1991). The orientation and ellipticity of the central mass peak is in striking agreement with those of the galaxy distribution and the X-ray map. However, the authors find some indication that the mass is more centrally condensed than the other two distributions. In addition, given the finite angular resolution of the mass map, the core size derived from weak lensing is most likely only an upper bound to the true value, and in both clusters the derived core size is significantly larger than found in clusters with giant luminous arcs (see, e.g., Fort & Mellier 1994).

The mass-to-light ratios for the two clusters are $\sim 1000\,h$ and $\sim 740\,h$, respectively. However, at least for Cl 0016+16, the mass scale is fairly uncertain, owing to the high cluster redshift and the unknown redshift distribution of the faint galaxies. The mean value of $D_{\rm ds}/D_{\rm s}$ must be estimated from an assumed distribution $p(z)$.

The unprecedented imaging quality of the refurbished *Hubble Space Telescope* (HST) can be used profitably for weak lensing analyses. Images taken with the *Wide Field Planetary Camera 2* (WFPC2) have an angular resolution of order $0.''1$, limited by the pixel size. Because of this superb resolution and the lower sky background, the number density of galaxy images for which a shape can reliably be measured is considerably larger than from the ground, so that higher-resolution mass maps can be determined. The drawback is the small field covered by the WFPC2, which consists of 3 CCD chips with $80''$ side-length each. Using the first publicly available deep image of a cluster obtained with the WFPC2, Seitz et al. (1996) have constructed a mass map of the cluster Cl 0939+47 ($z = 0.41$). Figure 14 clearly shows a mass peak near the left boundary of the frame shown. This maximum coincides with the cluster centre as determined from the cluster galaxies (Dressler & Gunn 1992). Furthermore, a secondary maximum is clearly visible in the mass map, as well as a pronounced minimum. When compared to the optical image, a clear correlation with the bright (cluster) galaxies is obvious. In particular, the secondary maximum and the minimum correspond to the same features in the bright galaxy distribution. A formal correlation test confirms this similarity. Applying the maximum-likelihood mass reconstruction technique (Seitz et al. 1998c; see Sect. 5.4.) to the same HST image, Geiger & Schneider (1999) constructed a higher-resolution map of this cluster. The angular resolution achieved is much higher in the cluster centre, predicting a region in which strong lensing effects may occur. Indeed, Trager et al. (1997) reported on a highly elongated arc and a triple image, with both source galaxies having a redshift $z \approx 3.97$.

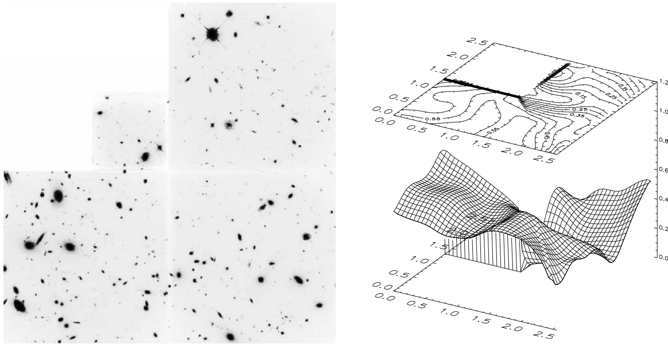

FIG. 14.—*Left panel:* WFPC2 image of the cluster Cl0939+4713 (A 851); North is at the bottom, East to the right. The cluster centre is located at about the upper left corner of the left CCD, a secondary maximum of the bright (cluster) galaxies is seen close to the interface of the two lower CCDs, and a minimum in the cluster light is at the interface between the two right CCDs. In the lensing analysis, the data from the small CCD (the Planetary Camera) were not used. *Right panel:* The reconstructed mass distribution of A 851, assuming a mean redshift of the $N = 295$ galaxies with $24 \le R \le 25.5$ of $\langle z \rangle = 1$. (from Seitz et al. 1996)

The X-ray map of this cluster (Schindler & Wambsganss 1997) shows that the two mass peaks are also close to two X-ray components. The determination of the total mass inside the WFPC2 frame is difficult, for two reasons: First, the high redshift of the cluster implies that the mean value of $D_{\rm ds}/D_{\rm s}$ depends quite sensitively on the assumed redshift distribution of the background galaxies. Second, the small field of the WFPC2 precludes the measurement of the surface mass density at large distance where κ tends to zero, and thus the mass-sheet degeneracy implies a considerable uncertainty in the mass scale. Attempting to lift the mass sheet degeneracy with the number-density effect – see Sect. 4.4.1. –, a mass-to-light ratio of $\sim 250\,h$ was derived within the WFPC2 aperture. This value is also affected by the unknown fraction of cluster members in the catalog of faint galaxies. Seitz et al. (1996) assumed that the spatial distribution of faint cluster galaxies follows that of brighter cluster galaxies. The striking difference between the M/L ratios for this and the other clusters described above may be related to the fact that Cl 0939+47 is the highest-redshift cluster in the Abell catalog (A 851). Hence, it was selected by its high optical luminosity, whereas the previously mentioned clusters are all X-ray selected. The X-ray luminosity of Cl 0939+47 is fairly small for such a rich cluster (Schindler & Wambsganss 1996). Since X-ray luminosity and cluster mass are generally well correlated, the small M/L-ratio found from the weak lensing analysis is in agreement with the expectations based on the high optical flux and the low X-ray flux. Note that the large spread of mass-to-light ratios as found by the existing cluster mass reconstructions is unexpected in the frame of hierarchical models of structure formation and thus poses an interesting astrophysical problem.

Hoekstra et al. (1998) reconstructed the mass distribution in the cluster MS 1358+62 from a mosaic of HST images, so that their data field in substantially larger than for a single HST pointing (about $8' \times 8'$). This work uses the correction method presented in Sect. 4.6.2., thus accounting for the relatively strong PSF anisotropy at the edges of each WFPC2 chip. A weak-lensing signal out to 1.5 Mpc is found. The X-ray mass is found to be slightly lower than the dynamical mass estimate, but seems to agree well with the lensing mass determination.

Luppino & Kaiser (1997) found a surprisingly strong weak-lensing signal in the field of the high-redshift cluster MS 1054−03 ($z = 0.83$). This implies that the sheared galaxies must have an appreciably higher redshift than the cluster, thus strongly constraining their redshift distribution. In fact, unless the characteristic redshift of these faint background galaxies is $\gtrsim 1.5$, this cluster would have an unrealistically large mass. It was also found that the lensing signal from the bluer galaxies is stronger than from the redder ones, indicating that the characteristic redshift of the bluer sample is higher. In fact, the mass estimated assuming $\langle z_{\rm s} \rangle = 1.5$ agrees well with results from analyses of the X–ray emission (Donahue et al. 1998) and galaxy kinematics (Tran et al. 1999).

Using an HST mosaic image in two filters, Hoekstra et al. (2000) also studied MS 1054. They found a tangential distortion which is smaller than that obtained by Luppino & Kaiser (1997) by about a factor of 1.5, but fairly well in agreement with that obtained by Clowe et al. (2000) from Keck imaging. They estimated the redshift distribution of the background galaxies from the photometric redshifts obtained in the Hubble Deep Fields, both as a function of magnitude and of colour. This enabled them to study the relative lensing strength $\langle Z \rangle$ as a function of these two observables, finding, as expected, the lensing strength increasing towards fainter magnitudes and, in agreement with Luppino & Kaiser (1997) and Clowe et al. (2000), with bluer colour. The estimated mass is in very good agreement with that obtained from the X-ray temperature of this cluster. The mass map shows three distinct peaks which are in good correspondence with the observed distribution of cluster galaxies. Clowe et al. (1998) derived weak lensing maps for two additional clusters at $z \sim 0.8$, namely MS 1137+66 at $z = 0.783$ and RXJ 1716+67 at $z = 0.813$.

The large-format CCD cameras allow weak-lensing studies of low-redshift clusters which subtend a larger solid angle on the sky. As a first example, Joffre et al. (1999) obtained the mass map for the cluster Abell 3667 ($z = 0.055$). Investigations of low-redshift clusters are particularly useful since for them more detailed X-ray and optical information is available than for higher-redshift ones.

The mass distribution in the supercluster MS 0302+17 at $z = 0.42$ was reconstructed by Kaiser et al. (1998) in a wide-field image of size $\sim 30'$. The supercluster consists of three clusters which are very close together on the sky and in redshift. The image contains about 30,000 galaxies from which a shear can be measured. This shear was found to correlate strongly with the distribution of the early-type (foreground) galaxies in the field, provided that the overall mass-to-light ratio is about $250\,h$. Each of the three clusters, which are also seen in X-rays, is recovered in the mass map. The ratios between mass and light or X-ray emission differ slightly across the three clusters, but the differences are not highly significant.

A magnification effect was detected from the depletion of the number counts (see Sect. 4.4.1.) in two clusters. Fort et al. (1997) discovered that the number density of very faint galaxies drops dramatically near the critical curve of the cluster Cl 0024+16, and remains considerably lower than the mean number density out to about twice the Einstein radius. This is seen in photometric data with two filters. Fort et al. (1997) interpret this broad depletion curve in terms of a broad redshift distribution of the background galaxies, so that the location of the critical curve of the cluster varies over a large angular scale. A spatially-dependent number depletion was detected in the cluster A 1689 by Taylor et al. (1998).

These examples should suffice to illustrate the current status of weak lensing cluster mass reconstructions. For additional results, see Squires et al. (1996b), Squires et al. (1997), Fischer et al. (1997), Fischer & Tyson (1997), and Athreya et al. (1999). Many of the difficulties have been overcome; e.g., the method presented in Sect. 4.6.2. appears to provide an accurate correction method for PSF effects. The quantitative results, for example for the M/L-ratios, are somewhat uncertain due to the lack of sufficient knowledge on the source redshift distribution, which applies in particular to the high-redshift clusters.

Further large-format HST mosaic images either are already or will soon become available, e.g. for the clusters A 2218, A 1689, and MS 1054−03. Their analysis will substantially increase the accuracy of cluster mass determinations from weak lensing compared to ground-based imaging.

5.5. *Outlook*

We have seen in the preceding subsection that first results on the mass distribution in clusters were derived with the methods described earlier. Because weak lensing is now widely regarded as the most reliable method to determine the mass distribution of clusters, since it does not rely on assumptions on the physical state and symmetries of the matter distribution, further attempts at improving the method are in progress, and some of them will briefly be outlined below.

In particular, we describe a method which simultaneously accounts for shear and magnification information, and which can incorporate constraints from strong-lensing features (such as arcs and multiple images of background sources). A method for the determination of the local shear is described next which does not rely on the detection and the quadrupole measurement of individual galaxies, and instead makes use of the light from very faint galaxies which need not be individually detected. We will finally consider the potential of weak lensing for determining the redshift distribution of galaxies which are too faint to be investigated spectroscopically, and report on first results.

5.5.1. *Maximum-Likelihood Cluster Reconstructions*

The mass reconstruction method described above is a direct method: The locally averaged observed image ellipticities $\langle \varepsilon \rangle$ are inserted into an inversion equation such as (5.10) to find the mass map $\kappa(\vec\theta)$. The beauty of this method is its simplicity and computational speed. Mass reconstructions from the observed image ellipticities are performed in a few CPU seconds.

The drawback of this method is its lack of flexibility. No additional information can be incorporated into the inversion process. For example, if strong-lensing features like giant arcs or multiple galaxy images are observed, they should be included in the mass reconstruction. Since such strong-lensing features typically occur in the innermost parts of the clusters (at $\lesssim 30''$ from cluster centres), they strongly constrain the mass distribution in cluster cores which can hardly be probed by weak lensing alone due to its finite angular resolution. A further example is the incorporation of magnification information, as described in Sect. 4.4., which can in principle not only be used to lift the mass-sheet degeneracy, but also provides local information on the shape of the mass distribution.

An additional problem of direct inversion techniques is the choice of the smoothing scale which enters the weight factors u_i in (5.16). We have not given a guideline on how this scale should be chosen. Ideally, it should be adapted to the data. In regions of strong shear, the signal-to-noise ratio of a shear measurement for a fixed number of galaxy images is larger than in regions of weak shear, and so the smoothing scale can be smaller there.

Recently, these problems have been attacked with inverse methods. Suppose the mass distribution of a cluster is parameterised by a set of model parameters p_k. These model parameters could then be varied until the best-fitting model for the observables is found. Considering for example the observed image ellipticities ε_i and assuming a non-critical cluster, the expectation value of ε_i is the reduced shear g at the image position, and the dispersion is determined (mainly) by the intrinsic dispersion of galaxy ellipticities σ_ε. Hence, one can define a χ^2-function

$$\chi^2 = \sum_{i=1}^{N_{\rm g}} \frac{|\varepsilon_i - g(\vec\theta_i)|^2}{\sigma_\varepsilon^2} \tag{5.34}$$

and minimise it with respect to the p_k. A satisfactory model is obtained if χ^2 is of order $N_{\rm g}$ at its minimum, as long as the number of parameters is much smaller than $N_{\rm g}$. If the chosen parameterisation does not achieve this minimum value, another one must be tried. However, the resulting mass model will depend on the parameterisation which is a serious drawback relative to the parameter-free inversion methods discussed before.

This problem can be avoided with ‘generic’ mass models. For instance, the deflection potential $\psi(\vec\theta)$ can be composed of a finite sum of Fourier modes (Squires & Kaiser 1996), whose amplitudes are the parameters p_k.[12] The number of Fourier modes can be chosen such that the resulting χ^2 per degree of freedom is approximately unity. Additional modes would then start to fit the noise in the data.

Alternatively, the values of the deflection potential ψ on a (regular) grid can be used as the p_k. Bartelmann et al. (1996) employed the locally averaged image ellipticities and the size ratios $\langle\omega\rangle/\langle\omega\rangle_0$ – see (4.47) – on a grid. The corresponding expectation values of these quantities, the reduced shear g and the magnification μ, were calculated by finite differencing of the discretised deflection potential ψ. Since both γ and κ, and thus μ, are unchanged under the transformation $\psi(\vec\theta) \to \psi(\vec\theta) + \psi_0 + \vec a \cdot \vec\theta$, the deflection potential has to be kept fixed at three grid points. If no magnification information is used, the mass-sheet degeneracy allows a further transformation of ψ which leaves the expected image ellipticities invariant, and the potential has to be kept fixed at four grid points.

A χ^2-function was defined using the local dispersion of the image ellipticities and image sizes relative to unlensed sizes of galaxies with the same surface brightness, and it was minimised with respect to the values of ψ on the grid points. The grid spacing was chosen such that the resulting minimum χ^2 has approximately the correct value. Tests with synthetic data sets, using a numerically generated cluster mass distribution, showed that this method reconstructs very satisfactory mass maps, and the total mass of the cluster was accurately reproduced.

If a finer grid is used, the model for the deflection potential will reproduce noise features in the data. On the other hand, the choice of a relatively coarse grid which yields a satisfactory χ^2 implies that the resolution of the mass map is constant over the data field. Given that the signal increases towards the centre of the cluster, one would like to use a finer grid there. To avoid overfitting of noise, the maximum-likelihood method can be complemented by a regularisation term (see Press et al. 1986, Chap. 18). As shown by Seitz et al. (1998c), a maximum-entropy regularisation (Narayan & Nityananda 1986) is well suited for the problem at hand. As in maximum-entropy image restoration (e.g., Lucy 1994), a prior is used in the entropy term which is a smoothed version of the current density field, and thus is being adapted during the minimisation. The relative weight of the entropy term is adjusted such that the resulting minimum χ^2 is of order unity per degree of freedom.

In this scheme, the expectation values and dispersions of the individual image ellipticities and sizes are found by bi-linear interpolation of κ and γ on the grid which themselves are obtained by finite differencing of the potential. When tested on synthetic data sets, this refined maximum-likelihood method produces mass maps with considerably higher resolution near the cluster centre without over-fitting the noise at larger clustercentric distances. The practical implementation of this method is somewhat complicated. In particular, if critical clusters are studied, some modifications have to be included to allow the minimisation algorithm to move critical curves across galaxy images in the lens plane. However, the quality of the reconstruction justifies the additional effort, especially if high-quality data from HST images are available. A first application of this method is presented by Geiger & Schneider (1999).

Inverse methods such as the ones described here are likely to become the standard tool for cluster mass profile reconstruction, owing to their flexibility. As mentioned before, additional constraints from strong lensing signatures such as arcs and multiply-imaged sources, can straightforwardly be incorporated into these methods. The additional numerical effort is negligible compared to the efforts needed to gain the observational data. Direct inversion methods will certainly retain an important role in this field, to obtain quick mass maps during the galaxy image-selection process (e.g., cuts in colour and brightness can be applied). Also, a mass map obtained by a direct method as a starting model in the inverse methods reduces the computational effort.

5.5.2. The Auto-Correlation Function of the Extragalactic Background Light

So far, we described how shear can be determined from ellipticities of individual galaxy images on a CCD. In that context, a galaxy image is a statistically significant flux enhancement on the CCD covering several contiguous pixels and being more extended than the PSF as determined from stars. Reducing the threshold for the signal-to-noise per object, the number density of detected galaxies increases, but so does the fraction of misidentifications. Furthermore, the measured ellipticity of faint galaxies has larger errors than that of brighter and larger images. The detection threshold therefore is a compromise between high number density of images and significance per individual object.

Even the faintest galaxy images whose ellipticity cannot be measured reliably still contain information on the lens distortion.

[12] It is important to note that the deflection potential ψ rather than the surface mass density κ (as in Squires & Kaiser 1996) should be parameterised, because shear and surface mass density depend on the local behaviour of ψ, while the shear *cannot* be obtained from the local κ, and not even from κ on a finite field. In addition, the local dependence of κ and γ on ψ is computationally much more efficient than calculating γ by integrating over κ as in Bridle et al. (1998).

It is therefore plausible to use this information, by 'adding up' the faintest galaxies statistically. For instance, one could co-add their brightness profiles and measure the shear of the combined profiled. This procedure, however, is affected by the uncertainties in defining the centres of the faint galaxies. Any error in the position of the centre, as defined in (5.1), will affect the resulting ellipticity.

To avoid this difficulty, and also the problem of faint object definition at all, van Waerbeke et al. (1997) have suggested considering the auto-correlation function (ACF) of the 'background' light. Most of the sky brightness is due to atmospheric scattering, but this contribution is uniform. Fluctuations of the brightness on the scale of arc seconds is supposedly mainly due to very faint galaxies. Therefore, these fluctuations should intrinsically be isotropic. If the light from the faint galaxies propagates through a tidal gravitational field, the isotropy will be perturbed, and this provides a possibility to measure this tidal field.

Specifically, if $I(\vec{\theta})$ denotes the brightness distribution as measured on a CCD, and $\bar{I}$ is the brightness averaged over the CCD (or a part of it, see below), the auto-correlation function $\xi(\vec{\theta})$ of the brightness is defined as

$$\xi(\vec{\theta}) = \left\langle \left(I(\vec{\vartheta}) - \bar{I}\right) \left(I(\vec{\vartheta}+\vec{\theta}) - \bar{I}\right) \right\rangle_{\vec{\vartheta}} , \qquad (5.35)$$

where the average is performed over all pairs of pixels with separation $\vec{\theta}$. From the invariance of surface brightness (3.10, page 17) and the locally linearised lens mapping, $I(\vec{\theta}) = I^{(\mathrm{s})}(\mathcal{A}\vec{\theta})$, one finds that the observed ACF is related to the intrinsic ACF $\xi^{(\mathrm{s})}$, defined in complete analogy to (5.35), by

$$\xi(\vec{\theta}) = \xi^{(\mathrm{s})}(\mathcal{A}\vec{\theta}) . \qquad (5.36)$$

Thus the transformation from intrinsic to observed ACF has the same functional form as the transformation of surface brightness. In analogy to the definition of the quadrupole tensor Q for galaxy images – see (5.2) – the tensor of second moments of the ACF is defined as

$$\mathcal{M}_{ij} = \frac{\int \mathrm{d}^2\theta\, \xi(\vec{\theta})\, \theta_i \theta_j}{\int \mathrm{d}^2\theta\, \xi(\vec{\theta})} . \qquad (5.37)$$

The transformation between the observed quadrupole tensor $\mathcal{M}$ and the intrinsic one, $\mathcal{M}^{(\mathrm{s})}$, is the same as for the moment tensor of image ellipticities, (5.5), $\mathcal{M}^{(\mathrm{s})} = \mathcal{A}\mathcal{M}\mathcal{A}$. As shown by van Waerbeke et al. (1997), the tensor $\mathcal{M}$ directly determines the distortion δ,

$$\delta = \frac{\mathcal{M}_{11} - \mathcal{M}_{22} + 2\mathrm{i}\mathcal{M}_{12}}{\mathcal{M}_{11} + \mathcal{M}_{22}} . \qquad (5.38)$$

Hence, δ is related to $\mathcal{M}$ in the same way as the complex ellipticity χ is related to Q. In some sense, the ACF plays the role of a single 'equivalent' image from which the distortion can be determined, instead of an ensemble average over individual galaxy ellipticities.

Working with the ACF has several advantages. First, centres of galaxy images do not need to be determined, which avoids a potential source of error. Second, the ACF can be used with substantial flexibility. For instance, one can use all galaxy images which are detected with high significance, determine their ellipticity, and obtain an estimate of δ from them. Sufficiently large circles containing these galaxies can be cut out of the data frame, so that the remaining frame is reminiscent of a Swiss cheese. The ACF on this frame provides another estimate of δ, which is independent information and can statistically be combined with the estimate from galaxy ellipticities. Or one can use the ACF only on galaxy images detected within a certain magnitude range, still avoiding the need to determine centres.

Third, on sufficiently deep images with the brighter objects cut out as just described, one might assume that the intrinsic ACF is due to a very large number of faint galaxies, so that the intrinsic ACF becomes a universal function. This function can in principle be determined from deep HST images. In that case, one also knows the width of the intrinsic ACF, as measured by the trace or determinant of $\mathcal{M}$, and can determine the magnification from the width of the observed ACF, very similar to the method discussed in Sect. 4.4.2., but with the advantage of dealing with a single 'universal source'.

If this universal intrinsic ACF does exist, corrections of the measured $\mathcal{M}$ for a PSF considerably simplify compared to the case of individual image ellipticities, as shown by van Waerbeke et al. (1997). They performed several tests on synthetic data to demonstrate the potential of the ACF method for the recovery of the shear applied to the simulated images. van Waerbeke et al. determined shear fields of two clusters, with several magnitude thresholds for the images which were punched out. A comparison of these shear fields with those obtained from the standard method using galaxy ellipticities clearly shows that the ACF method is at least competitive, but since it provides additional information from those parts of the CCD which are unused by the standard method, it should in be employed any case. The optimal combination of standard method and ACF still needs to be investigated, but detailed numerical experiments indicate that the ACF may be the best method for measuring very weak shear amplitudes (L. van Waerbeke & Y. Mellier, private communication).

5.5.3. The Redshift Distribution of Very Faint Galaxies

Galaxy redshifts are usually determined spectroscopically. A successful redshift measurement depends on the magnitude of the galaxy, the exposure time, and the spectral type of the galaxy. If it shows strong emission or absorption lines, as star-forming galaxies do, a redshift can much easier be determined than in absence of strong spectral features. The recently completed Canadian-French Redshift Survey (CFRS) selected 730 galaxies in the magnitude interval $17.5 \le I \le 22.5$ (see Lilly et al. 1995 and references therein). For 591 of them (81%), redshifts were secured with multi-slit spectroscopy on a 3.6m telescope (CFHT) with a typical exposure time of ~ 8 hours. Whereas the upcoming 10m-class telescopes will be able to perform redshift surveys to somewhat fainter magnitude limits, it will be difficult to secure fairly complete redshift information of a flux-limited galaxy sample fainter than $I \sim 24$. In addition, it can be expected that many galaxies in a flux-limited sample with fainter threshold will have redshifts between ~ 1.2 and ~ 2.2, where the cleanest spectral features, the OII emission line at $\lambda = 372.7$ nm and the $\lambda = 400$ nm break are shifted beyond the region where spectroscopy can easily be done from the ground.

As we have seen, the calibration of cluster mass distributions depends on the assumed redshift distribution of the background galaxies. Most of the galaxies used for the reconstruction are considerably fainter than those magnitude limits for which complete redshift samples are available, so that this mass calibration requires an extrapolation of the redshift distribution from brighter galaxy samples. The fact that lensing is sensitive to the redshift distribution is not only a source of uncertainty, but also offers the opportunity to investigate the redshift distribution of galaxies too faint to be investigated spectroscopically. Several approaches towards a redshift estimate of faint galaxies by lensing have been suggested, and some of them have already shown spectacular success, as will be discussed next.

First of all, a strongly lensed galaxy (e.g. a giant luminous arc) is highly magnified, and so the gravitational lens effect allows to obtain spectra of objects which would be too faint for a spectroscopic investigation without lensing. It was possible in this way to measure the redshifts of several arcs, e.g., the giant arc in A 370 at $z = 0.724$ (Soucail et al. 1988), the arclet A 5 in A 370 at $z = 1.305$ (Mellier et al. 1991), the giant arc in Cl 2244−02 at $z = 2.237$ (Mellier et al. 1991), and the 'straight arc' in A 2390 at $z = 0.913$ (Pello et al. 1991). For a more complete list of arc redshifts, see Fort & Mellier (1994). A fair fraction of galaxies with redshift $z \gtrsim 4$ have been found behind clusters, for example two arclet sources at $z \approx 4.05$ behind A 2390 (Frye & Broadhurst 1998, Pelló et al. 1999b), two sources at $z \approx 3.97$ behind Cl 0939+4713 (Trager et al. 1997), and two sources behind MS 1358+62, which for a few months held the redshift record of $z = 4.92$ (Franx et al. 1997).

If the cluster contains several strong-lensing features, the mass model can be sufficiently well constrained to determine the arc magnifications (if they are resolved in width, which has become possible only from imaging with the refurbished HST), and thus to determine the unlensed magnitude of the source galaxies, some of which are fainter that $B \sim 25$.

Some clusters, such as A 370 and A 2218, were observed in great detail both from the ground and with HST, and show a large number of strongly lensed images. They can be used to construct very detailed mass models of the cluster centre (e.g., Kneib et al. 1993; Kneib et al. 1996). An example is A 2218, in which at least five multiply imaged systems were detected (Kneib et al. 1996), and several giant arcs were clearly seen. Refining the mass model for A 2218 constructed from ground-based data (Kneib et al. 1995) with the newly discovered or confirmed strong lensing features on the WFPC2 image, a strongly constrained mass model for the cluster can be computed and calibrated by two arc redshifts (a five-image system at $z = 0.702$, and at $z = 1.034$).

Up to now, the deepest HST image taken in the direction of a cluster was on A 1689, perhaps the strongest lensing cluster yet detected (HST proposal number 6004, PI J. A. Tyson). This impressive image provides a wealth of strong-lensing features which should allow the construction of a very detailed mass model for its central region. In addition, a large-scale, though fairly shallow, image mosaic has been obtained with HST (HST proposal number 5993, PI N. Kaiser). These two data sets will yield the most detailed mass profile currently obtainable.

Visual inspection of the WFPC2 image immediately shows a large number of arclets in A 2218, which surround the cluster centre in a nearly perfect circular pattern. These arclets have very small axis ratios, and most of them are therefore highly distorted. The strength of the distortion depends on the redshift of the corresponding galaxy. Assuming that the sources have a considerably smaller ellipticity than the observed images, one can then estimate a redshift range of the galaxy.

To be more specific, let $p^{(s)}(\epsilon^{(s)})$ be the probability density of the intrinsic source ellipticity, assumed for simplicity to be independent of redshift. The corresponding probability distribution for the image ellipticity is then

$$p(\epsilon) = p^{(s)}\left(\epsilon^{(s)}(\epsilon)\right) \det\left(\frac{\partial \epsilon^{(s)}}{\partial \epsilon}\right) , \tag{5.39}$$

where the transformation $\epsilon^{(s)}(\epsilon)$ is given by eq. (4.12, page 22), and the final term is the Jacobian of this transformation. For each arclet near the cluster centre where the mass profile is well constrained, the value of the reduced shear g is determined up to the unknown redshift of the source – see eq. (4.20, page 23).

One can now try to maximise $p(\epsilon)$ with respect to the source redshift, and in that way find the most likely redshift for the arc.[13] Depending on the ellipticity of the arclet and the local values of shear and surface mass density, three cases have to be distinguished: (1) the arclet has the 'wrong' orientation relative to the local shear, i.e., if the source lies behind the cluster, it must be even more elliptical than the observed arclet. For the arclets in A 2218, this case is very rare. (2) The most probable redshift is 'at infinity', i.e., even if the source is placed at very high redshift, the maximum of $p(\epsilon)$ is not reached. (3) $p(\epsilon)$ attains a maximum at a finite redshift. This is by far the most common case in A 2218.

This method, first applied to A 370 (Kneib et al. 1994), was used to estimate the redshifts of ~ 80 arclets in A 2218 brighter than $R \sim 25$. Their typical redshifts are estimated to be of order unity, with the fainter sub-sample $24 \le R \le 25$ extending to somewhat higher redshifts. For one of them, a redshift range $2.6 \lesssim z \lesssim 3.3$ was estimated, and a spectroscopic redshift of $z = 2.515$ was later measured (Ebbels et al. 1996), providing spectacular support for this method. Additional spectroscopic observations of arclets in A 2218 were conducted and further confirmed the reliability of the method for the redshift estimates of individual arclets (Ebbels et al. 1998).

Another success of this arclet redshift estimate was recently achieved in the cluster A 2390, which can also be modelled in great detail from HST data. There, two arclets with very strong elongation did not fit into the cluster mass model unless they are at very high redshift. Spectroscopic redshifts of $z \sim 4.05$ were recently measured for these two arclets (Frye & Broadhurst 1998, Pelló et al. 1999a).

However, several issues should be kept in mind. First, the arclets for which a reliable estimate of the redshift can be obtained are clearly magnified, and thus the sample is magnification biased. Since it is well known that the galaxy number counts are considerable steeper in the blue than in the red (see, e.g., Smail et al. 1995a), blue galaxies are preferentially selected as arclets – see eq. (4.42). This might also provide the explanation why most of the giant arcs are blue. Therefore, the arclets represent probably a biased sample of faint galaxies. Second, the redshift dependence of $p(\epsilon)$ enters through the ratio $D_{\rm ds}/D_{\rm s}$. For a cluster at relatively low redshift, such as A 2218 ($z_{\rm d} = 0.175$), this ratio does not vary strongly with redshift once the source redshift is larger than ~ 1. Hence, to gain more accurate redshift estimates for high-redshift galaxies, a moderately-high redshift cluster should be used.

The method just described is not a real 'weak lensing' application, but lies on the borderline between strong and weak lensing. With weak lensing, the redshifts of *individual* galaxy images cannot be determined, but some statistical redshift estimates can be obtained. Suppose the mass profile of a cluster has been reconstructed using the methods described in Sect. 5.2. or Sect. 5.5.1., for which galaxy images in a certain magnitude range were used. If the cluster contains strong-lensing features with spectroscopic information (such as a giant luminous arc with measured redshift), then the overall mass calibration can be determined, i.e., the factor $\langle Z \rangle$ – see Sect. 4.3.2. – can be estimated, which provides a first integral constraint on the redshift distribution.

Repeating this analysis with several such clusters at different redshifts, further estimates of $\langle Z \rangle$ with different $D_{\rm d}$ are obtained, and thus additional constraints on the redshift distribution. In addition, one can group the faint galaxy images into sub-samples, e.g., according to their apparent magnitude. Ignoring for simplicity the magnification bias (which can safely be done in the

[13] This simplified treatment neglects the magnification bias, i.e. the fact that at locations of high magnification the redshift probability distribution is changed – see Sect. 4.3.2..

outer parts of clusters), one can determine $\langle Z\rangle$ for each magnitude bin. Restricting our treatment to the regions of weak lensing only, such that $|\gamma| \ll 1$, $\kappa \ll 1$, the expectation value of the ellipticity ε_i of a galaxy at position $\vec{\theta}_i$ is $\langle Z\rangle\gamma(\vec{\theta}_i)$, and so an estimate of $\langle Z\rangle$ for the galaxy sub-sample under consideration is

$$\langle Z\rangle = \frac{\sum \Re\left(\gamma(\vec{\theta}_i)\varepsilon_i^*\right)}{\sum |\gamma(\vec{\theta}_i)|^2} \,. \tag{5.40}$$

In complete analogy, Bartelmann & Narayan (1995) suggested the 'lens parallax method', an algorithm for determining mean redshifts for galaxy sub-samples at fixed surface brightness, using the magnification effect as described in Sect. 4.4.2.. Since the surface brightness I is most likely much more strongly correlated with galaxy redshift than the apparent magnitude (due to the $(1+z)^{-4}$ decrease of bolometric surface brightness with redshift), a narrow bin in I will probably correspond to a fairly narrow distribution in redshift, allowing to relate $\langle Z\rangle$ of a surface brightness bin fairly directly to a mean redshift in that bin, while $\langle Z\rangle$ in magnitude bins can only be translated into redshift information with a parameterised model of the redshift distribution. On the other hand, apparent magnitudes are easier to measure than surface brightness and are much less affected by seeing.

Even if a cluster without strong lensing features is considered, the two methods just described can be applied. The mass reconstruction then gives the mass distribution up to an overall multiplicative constant. We assume here that the mass-sheet degeneracy can be lifted, either using the magnification effect as described in Sect. 5.4., or by extending the observations so sufficiently large distances so that $\kappa \approx 0$ near the boundary of the data field. The mass scale can then be fixed by considering the brightest sub-sample of galaxy images for which a shear signal is detected if they are sufficiently bright for their redshift probability distribution to be known from spectroscopic redshift surveys (Bartelmann & Narayan 1995).

Whereas these methods have not yet rigorously been applied, there is one observational result which indicates that the faint galaxy population has a relatively high median redshift. In a sequence of clusters with increasing redshift, more and more of the faint galaxies will lie in the foreground or very close behind the cluster and therefore be unlensed. The dependence of the observed lensing strength of clusters on their redshift can thus be used as a rough indication of the median redshift of the faint galaxies. This idea was put forward by Smail et al. (1994), who observed three clusters with redshifts $z = 0.26$, $z = 0.55$ and $z = 0.89$. In the two lower-redshift clusters, a significant weak lensing signal was detected, but no significant signal in the high-redshift cluster. From the detection, models for the redshift distribution of faint $I \le 25$ can be ruled out which predict a large fraction to be dwarf galaxies at low redshift. The non-detection in the high-redshift cluster cannot easily be interpreted since little information (e.g., from X-ray maps) is available for this cluster, and thus the absence of a lensing signal may be due to the cluster being not massive enough.

However, the detection of a strong shear signal in the cluster MS 1054−03 at $z = 0.83$ (Luppino & Kaiser 1997) implies that a large fraction of galaxies with $I \le 25.5$ must lie at redshifts larger than $z \sim 1.5$. They split their galaxy sample into red and blue sub-samples, as well as into brighter and fainter sub-samples, and found that the shear signal is mainly due to the fainter and the blue galaxies. If all the faint blue galaxies have a redshift $z_s = 1.5$, the mass-to-light ratio of this cluster is estimated to be $M/L \sim 580\,h$, and if they all lie at redshift $z_s = 1$, M/L exceeds $\sim 1000\,h$. This observational result, which is complemented by several additional shear detections in high-redshift clusters, one of them at $z = 0.82$ (G. Luppino, private communication), provides the strongest evidence for the high-redshift population of faint galaxies. In addition, it strongly constrains cosmological models; an $\Omega_0 = 1$ cosmological model predicts the formation of massive clusters only at relatively low redshifts (e.g., Richstone et al. 1992; Bartelmann et al. 1993) and has difficulties to explain the presence of strong lensing clusters at redshift $z \sim 0.8$.

Recently, Lombardi & Bertin (1999c) and Gautret et al. (2000) suggested that weak lensing by galaxy clusters can be used to constrain the cosmological parameters Ω_0 and Ω_Λ. Both of these two different methods assume that the redshift of background galaxies can be estimated, e.g. with sufficiently precise photometric-redshift techniques. Owing to the dependence of the lensing strength on the angular-diameter distance ratio D_{ds}/D_s, sufficiently detailed knowledge of the mass distribution in the lens and of the source redshifts can be employed to constrain these cosmological parameters. Such a determination through purely geometrical methods would be very valuable, although the observational requirements for applying these methods appear fairly demanding at present.

6. WEAK COSMOLOGICAL LENSING

In this section, we review how weak density perturbations in otherwise homogeneous and isotropic Friedmann-Lemaître model universes affect the propagation of light. We first describe how light propagates in the homogeneous and isotropic background models, and then discuss how local density inhomogeneities can be taken into account. The result is a propagation equation for the transverse separation between the light rays of a thin light bundle.

The solution of this equation leads to the deflection angle $\vec{\alpha}$ of weakly deflected light rays. In close analogy to the thin-lens situation, half the divergence of the deflection angle can be identified with an effective surface-mass density $\kappa_{\rm eff}$. The power spectrum of $\kappa_{\rm eff}$ is closely related to the power spectrum of the matter fluctuations, and it forms the central physical object of the further discussion. Any two-point statistics of cosmic magnification and cosmic shear can then be expressed in a fairly simple manner in terms of the effective-convergence power spectrum.

We discuss several applications, among which are the uncertainty in brightness determinations of cosmologically distant objects due to cosmic magnification, and several measures for cosmic shear, one of which is particularly suited for determining the effective-convergence power spectrum. At the end of this chapter, we turn to higher-order statistical measures of cosmic lensing effects, which reflect the non-Gaussian nature of the non-linearly evolved density perturbations.

When we give numerical examples, we generally employ four different model universes. All have the CDM power spectrum for density fluctuations, but different values for the cosmological parameters. They are summarised in Tab. 1. We choose two Einstein-de Sitter models, SCDM and σCDM, normalised either to the local abundance of rich clusters or to $\sigma_8 = 1$, respectively, and two low-density models, OCDM and ΛCDM, which are cluster normalised and either open or spatially flat, respectively.

Light propagation in inhomogeneous model universes has been the subject of numerous studies. Among them are Zel'dovich & Ya.B. (1964), Dashevskii et al. (1965), Kristian & Sachs (1966), Gunn (1967), Jaroszynski et al. (1990), Babul & Lee (1991), Bartelmann & Schneider (1991), Blandford et al. (1991), Miralda-Escudé (1991), and Kaiser (1992). Non-linear effects were included analytically by Jain & Seljak (1997), who also considered statistical effects of higher than second order, as did Bernardeau et al. (1997). A particularly suitable measure for

TABLE 1.—Cosmological models and their parameters used for numerical examples

Model	Ω_0	Ω_Λ	h	Normalisation	σ_8
SCDM	1.0	0.0	0.5	cluster	0.5
σCDM	1.0	0.0	0.5	σ_8	1.0
OCDM	0.3	0.0	0.7	cluster	0.85
ΛCDM	0.3	0.7	0.7	cluster	0.9

cosmic shear was introduced by Schneider et al. (1998a).

6.1. Light Propagation; Choice of Coordinates

As outlined in Sect. 3.2.1. (page 18), the governing equation for the propagation of thin light bundles through arbitrary space times is the equation of geodesic deviation (e.g. Misner et al. 1973, § 11; Schneider et al. 1992, § 3.5), or Jacobi equation (3.23, page 19). This equation implies that the transverse physical separation $\vec{\xi}$ between neighbouring rays in a thin light bundle is described by the second-order differential equation

$$\frac{\mathrm{d}^2\vec{\xi}}{\mathrm{d}\lambda^2} = \mathcal{T}\vec{\xi}\,, \tag{6.1}$$

where $\mathcal{T}$ is the *optical tidal matrix* (3.25, page 19) which describes the influence of space-time curvature on the propagation of light. The affine parameter λ has to be chosen such that it locally reproduces the proper distance and increases with decreasing time, hence $\mathrm{d}\lambda = -ca\,\mathrm{d}t$. The elements of the matrix $\mathcal{T}$ then have the dimension $[\text{length}]^{-2}$.

We already discussed in Sect. 3.2.1. that the optical tidal matrix is proportional to the unit matrix in a Friedmann-Lemaître universe,

$$\mathcal{T} = \mathcal{R}\,I\,, \tag{6.2}$$

where the factor $\mathcal{R}$ is determined by the Ricci tensor as in eq. (3.26, page 19). For a model universe filled with a perfect pressure-less fluid, $\mathcal{R}$ can be written in the form (3.28, page 19).

It will prove convenient for the following discussion to replace the affine parameter λ in eq. (6.1) by the comoving distance w, which was defined in eq. (2.3, page 4) before. This can be achieved using eqs. (3.31) and (3.32) together with the definition of Hubble's parameter, $H(a) = \dot{a}a^{-1}$. Additionally, we introduce the *comoving* separation vector $\vec{x} = a^{-1}\vec{\xi}$. These substitutions leave the propagation equation (6.1) in the exceptionally simple form

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}w^2} + K\vec{x} = \vec{0}\,, \tag{6.3}$$

where K is the spatial curvature given in eq. (2.30, page 6). Equation (6.3) has the form of an oscillator equation, hence its solutions are trigonometric or hyperbolic functions, depending on whether K is positive or negative. In the special case of spatial flatness, $K = 0$, the comoving separation between light rays is a linear function of distance.

6.2. Light Deflection

We now proceed by introducing density perturbations into the propagation equation (6.3). We assume throughout that the Newtonian potential Φ of these inhomogeneities is small, $|\Phi| \ll c^2$, that they move with velocities much smaller than the speed of light, and that they are localised, i.e. that the typical scales over which Φ changes appreciably are much smaller than the curvature scale of the background Friedmann-Lemaître model. Then, there exists a local neighbourhood around each density perturbation which is large enough to contain the perturbation completely and still small enough to be considered flat. Under these circumstances, the metric is well approximated by the first post-Newtonian order of the Minkowski metric (3.36, page 20). It then follows from eq. (3.36) that the effective local index of refraction in the neighbourhood of the perturbation is

$$\frac{\mathrm{d}l}{\mathrm{d}t} = n = 1 - \frac{2\Phi}{c^2}\,. \tag{6.4}$$

Fermat's principle (e.g. Blandford & Narayan 1986; Schneider 1985) demands that the light travel time along actual light paths is stationary, hence the variation of $\int n\mathrm{d}l$ must vanish. This condition implies that light rays are deflected locally according to

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}w^2} = -\frac{2}{c^2}\nabla_\perp\Phi\,. \tag{6.5}$$

In weakly perturbed Minkowski space, this equation describes how an *actual* light ray is curved away from a straight line in unperturbed Minkowski space. It is therefore appropriate for describing light propagation through e.g. the Solar system and other well-localised mass inhomogeneities.

This interpretation needs to be generalised for large-scale mass inhomogeneities embedded in an expanding cosmological background, since the meaning of a "straight" fiducial ray is then no longer obvious. In general, any physical fiducial ray will also be deflected by potential gradients along its way. We can, however, interpret $\vec{x}$ as the comoving separation vector between an arbitrarily chosen fiducial light ray and a closely neighbouring light ray. The right-hand side of eq. (6.5) must then contain the *difference* $\Delta(\nabla_\perp\Phi)$ of the perpendicular potential gradients between the two rays to account for the *relative* deflection of the two rays.

Let us therefore imagine a fiducial ray starting at the observer ($w = 0$) into direction $\vec{\theta} = \vec{0}$, and a neighbouring ray starting at the same point but into direction $\vec{\theta} \neq \vec{0}$. Let further $\vec{x}(\vec{\theta}, w)$ describe the comoving separation between these two light rays at comoving distance w. Combining the cosmological contribution given in eq. (6.3) with the modified local contribution (6.5) leads to the propagation equation

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}w^2} + K\vec{x} = -\frac{2}{c^2}\Delta\left\{\nabla_\perp\Phi[\vec{x}(\vec{\theta}, w), w]\right\}\,. \tag{6.6}$$

The notation on the right-hand side indicates that the difference of the perpendicular potential gradients has to be evaluated between the two light rays which have comoving separation $\vec{x}(\vec{\theta}, w)$ at comoving distance w from the observer.

Linearising the right-hand side of eq. (6.6) in $\vec{x}$ immediately returns the geodesic deviation equation (6.1) with the full optical tidal matrix, which combines the homogeneous cosmological contribution (3.28, page 19) with the contributions of local perturbations (3.37, page 20).

Strictly speaking, the comoving distance w, or the affine parameter λ, are changed in the presence of density perturbations. Here, we assume that the global properties of the weakly perturbed Friedmann-Lemaître models remain the same as in the homogeneous and isotropic case, and under this assumption the comoving distance w remains the same as in the unperturbed model.

To solve eq. (6.6), we first construct a Green's function $G(w, w')$, which has to be a suitable linear combination of either trigonometric or hyperbolic functions since the homogeneous equation (6.6) is an oscillator equation. We further have to specify two boundary conditions. According to the situation we have in mind, these boundary conditions read

$$\vec{x} = \vec{0} \,, \quad \frac{\mathrm{d}\vec{x}}{\mathrm{d}w} = \vec{\theta} \tag{6.7}$$

at $w = 0$. The first condition states that the two light rays start from the same point, so that their initial separation is zero, and the second condition indicates that they set out into directions which differ by $\vec{\theta}$.

The Green's function is then uniquely determined by

$$G(w,w') = \begin{cases} f_K(w-w') & \text{for} \quad w > w' \\ 0 & \text{otherwise} \end{cases} \,, \tag{6.8}$$

with $f_K(w)$ given in eq. (2.4, page 4). As a function of distance w, the comoving separation between the two light rays is thus

$$\begin{aligned} \vec{x}(\vec{\theta},w) &= f_K(w)\vec{\theta} \\ &- \frac{2}{c^2}\int_0^w dw'\, f_K(w-w')\,\Delta\left\{\nabla_\perp\Phi[\vec{x}(\vec{\theta},w'),w']\right\} \,. \end{aligned} \tag{6.9}$$

The perpendicular gradients of the Newtonian potential are to be evaluated along the true paths of the two light rays. In its exact form, eq. (6.10) is therefore quite involved.

Assuming that the change of the comoving separation vector $\vec{x}$ between the two *actual* rays due to light deflection is small compared to the comoving separation of *unperturbed* rays,

$$\frac{|\vec{x}(\vec{\theta},w') - f_K(w')\vec{\theta}|}{|f_K(w')\vec{\theta}|} \ll 1 \,, \tag{6.10}$$

we can replace $\vec{x}(\vec{\theta},w')$ by $f_K(w')\vec{\theta}$ in the integrand to arrive at a much simpler expression which corresponds to the Born approximation of small-angle scattering. The Born approximation allows us to replace the difference of the perpendicular potential gradients with the perpendicular gradient of the potential difference. Taking the potential difference then amounts to adding a term to the potential which depends on the comoving distance w' from the observer only. For notational simplicity, we can therefore rename the potential difference $\Delta\Phi$ between the two rays to Φ.

It is an important consequence of the Born approximation that the Jacobian matrix of the lens mapping (3.11, page 17; 6.28 below) remains symmetric even in the case of cosmological weak lensing. In a general multiple lens-plane situation, this is not the case (Schneider et al. 1992, chapter 9).

If the two light rays propagated through unperturbed spacetime, their comoving separation at distance w would simply be $\vec{x}'(\vec{\theta},w) = f_K(w)\vec{\theta}$, which is the first term on the right-hand side of eq. (6.10). The net deflection angle at distance w between the two rays is the difference between $\vec{x}'$ and $\vec{x}$, divided by the angular diameter distance to w, hence

$$\begin{aligned} \vec{\alpha}(\vec{\theta},w) &= \frac{f_K(w)\vec{\theta} - \vec{x}(\vec{\theta},w)}{f_K(w)} \\ &= \frac{2}{c^2}\int_0^w \mathrm{d}w'\, \frac{f_K(w-w')}{f_K(w)} \nabla_\perp\Phi[f_K(w')\vec{\theta},w'] \,. \end{aligned} \tag{6.11}$$

Again, this is the deflection angle of a light ray that starts out at the observer into direction $\vec{\theta}$ relative to a nearby fiducial ray. Absolute deflection angles cannot be measured. All measurable effects of light deflection therefore only depend on *derivatives* of the deflection angle (6.12), so that the choice of the fiducial ray is irrelevant for practical purposes. For simplicity, we call $\vec{\alpha}(\vec{\theta},w)$ the deflection angle at distance w of a light ray starting into direction $\vec{\theta}$ on the observer's sky, bearing in mind that it is the deflection angle relative to an arbitrarily chosen fiducial ray, so that $\vec{\alpha}(\vec{\theta},w)$ is far from unique.

In an Einstein-de Sitter universe, $f_K(w) = w$. Defining $y = w'/w$, eq. (6.12) simplifies to

$$\vec{\alpha}(\vec{\theta},w) = \frac{2w}{c^2}\int_0^1 \mathrm{d}y\,(1-y)\,\nabla_\perp\Phi(wy\vec{\theta},wy) \,. \tag{6.12}$$

Clearly, the deflection angle $\vec{\alpha}$ depends on the direction $\vec{\theta}$ on the sky into which the light rays start to propagate, and on the comoving distance w to the sources.

Recall the various approximations adopted in the derivation of eq. (6.12): (i) The density perturbations are well localised in an otherwise homogeneous and isotropic background, i.e. each perturbation can be surrounded by a spatially flat neighbourhood which can be chosen small compared to the curvature radius of the background model, and yet large enough to encompass the entire perturbation. In other words, the largest scale on which the density fluctuation spectrum $P_\delta(k)$ has appreciable power must be much smaller than the Hubble radius c/H_0. (ii) The Newtonian potential of the perturbations is small, $\Phi \ll c^2$, and typical velocities are much smaller than the speed of light. (iii) Relative deflection angles between neighbouring light rays are small enough so that the difference of the transverse potential gradient can be evaluated at the unperturbed path separation $f_K(w)\vec{\theta}$ rather than the actual one. Reassuringly, these approximations are very comfortably satisfied even under fairly extreme conditions. The curvature radius of the Universe is of order $cH_0^{-1} = 3000\,h^{-1}\,\mathrm{Mpc}$ and therefore much larger than perturbations of even several tens of Mpc's in size. Typical velocities in galaxy clusters are of order $10^3\,\mathrm{km\,s^{-1}}$, much smaller than the speed of light, and typical Newtonian potentials are of order $\Phi \lesssim 10^{-5}\,c^2$.

6.3. *Effective Convergence*

6.3.1. *Definition and Derivation*

In the thin-lens approximation, convergence κ and deflection angle $\vec{\alpha}$ are related by

$$\kappa(\vec{\theta}) = \frac{1}{2}\nabla_\theta\cdot\vec{\alpha}(\vec{\theta}) = \frac{1}{2}\frac{\partial\alpha_i(\vec{\theta})}{\partial\theta_i} \,, \tag{6.13}$$

where summation over i is implied. In exact analogy, an effective convergence $\kappa_{\mathrm{eff}}(w)$ can be defined for cosmological weak lensing,

$$\begin{aligned} \kappa_{\mathrm{eff}}(\vec{\theta},w) &= \frac{1}{2}\nabla_\theta\cdot\vec{\alpha}(\vec{\theta},w) \\ &= \frac{1}{c^2}\int_0^w \mathrm{d}w'\,\frac{f_K(w-w')f_K(w')}{f_K(w)} \\ &\times \frac{\partial^2}{\partial x_i\partial x_i}\Phi[f_K(w')\vec{\theta},w'] \,. \end{aligned} \tag{6.14}$$

Had we not replaced $\vec{x}(\vec{\theta},w')$ by $f_K(w')\vec{\theta}$ following eq. (6.10), eq. (6.15) would have contained second and higher-order terms in the potential derivatives. Since eq. (6.10) is a Volterra integral equation of the second kind, its solution (and derivatives thereof) can be expanded in a series, of which the foregoing expression for κ_{eff} is the first term. Equation (6.16) below shows that this term is of the order of the line-of-sight average of the density contrast δ. The next higher-order term, explicitly written down in the Appendix of Schneider et al. (1998a), is determined by the product $\delta(w')\,\delta(w'')$, averaged along the line-of-sight over $w' < w''$. Analogous estimates apply to higher-order terms. Whereas the density contrast may be large for individual density perturbations passed by a light ray, the average of δ is small compared to unity for most rays, hence $\kappa_{\rm eff} \ll 1$, and higher-order terms are accordingly negligible.

The effective convergence $\kappa_{\rm eff}$ in eq. (6.15) involves the two-dimensional Laplacian of the potential. We can augment it by $(\partial^2\Phi/\partial x_3^2)$ which involves only derivatives along the light path, because these average to zero in the limit to which we are working; the validity of this approximation has been verified with numerical simulations by White & Hu (1999). The three-dimensional Laplacian of the potential can then be replaced by the density contrast via Poisson's equation (2.65, page 11),

$$\Delta\Phi = \frac{3H_0^2\Omega_0}{2a}\,\delta\,. \tag{6.15}$$

Hence, we find for the effective convergence,

$$\kappa_{\rm eff}(\vec\theta,w) = \frac{3H_0^2\Omega_0}{2c^2}\int_0^w {\rm d}w'\,\frac{f_K(w')f_K(w-w')}{f_K(w)}\,\frac{\delta[f_K(w')\vec\theta,w']}{a(w')}\,. \tag{6.16}$$

The effective convergence along a light ray is therefore an integral over the density contrast along the (unperturbed) light path, weighted by a combination of comoving angular-diameter distance factors, and the scale factor a. The amplitude of $\kappa_{\rm eff}$ is proportional to the cosmic density parameter Ω_0.

Expression (6.16) gives the effective convergence for a fixed source redshift corresponding to the comoving source distance w. When the sources are distributed in comoving distance, $\kappa_{\rm eff}(\vec\theta,w)$ needs to be averaged over the (normalised) source-distance distribution $G(w)$,

$$\bar\kappa_{\rm eff}(\vec\theta) = \int_0^{w_{\rm H}} {\rm d}w\,G(w)\,\kappa_{\rm eff}(\vec\theta,w)\,, \tag{6.17}$$

where $G(w){\rm d}w = p_z(z){\rm d}z$. Suitably re-arranging the integration limits, we can then write the source-distance weighted effective convergence as

$$\bar\kappa_{\rm eff}(\vec\theta) = \frac{3H_0^2\Omega_0}{2c^2}\int_0^{w_{\rm H}} {\rm d}w\,\bar W(w)\,f_K(w)\,\frac{\delta[f_K(w)\vec\theta,w]}{a(w)}\,, \tag{6.18}$$

where the weighting function $\bar W(w)$ is now

$$\bar W(w) \equiv \int_w^{w_{\rm H}} {\rm d}w'\,G(w')\,\frac{f_K(w'-w)}{f_K(w')}\,. \tag{6.19}$$

The upper integration boundary $w_{\rm H}$ is the horizon distance, defined as the comoving distance obtained for infinite redshift. In fact, it is easily shown that the effective convergence can be written as

$$\kappa_{\rm eff} = \int {\rm d}z\,\frac{4\pi G}{c^2}\,\frac{D_{\rm d}D_{\rm ds}}{D_{\rm s}}\,\frac{{\rm d}D_{\rm prop}}{{\rm d}z}\,(\rho-\bar\rho)\,, \tag{6.20}$$

and the weighting function $\bar W$ is the distance ratio $\langle D_{\rm ds}/D_{\rm s}\rangle$, averaged over the source distances at fixed lens distance. Naively generalising the definition of the dimension-less surface-mass density (3.7, page 17) to a three-dimensional matter distribution would therefore directly have led to the cosmologically correct expression for the effective convergence.

6.4. *Effective-Convergence Power Spectrum*

6.4.1. *The Power Spectrum from Limber's Equation*

Here, we are interested in the statistical properties of the effective convergence $\kappa_{\rm eff}$, especially its power spectrum $P_\kappa(l)$. We refer the reader to Sect. 2.4. (page 14) for the definition of the power spectrum. We also note that the expression for $\bar\kappa_{\rm eff}(\vec\theta)$ is of the form (2.77, page 15), and so the power spectrum $P_\kappa(l)$ is given in terms of $P_\delta(k)$ by eq. (2.84, page 16), if one sets

$$q_1(w) = q_2(w) = \frac{3}{2}\,\frac{H_0^2}{c^2}\,\Omega_0\,\bar W(w)\,\frac{f_K(w)}{a(w)}\,. \tag{6.21}$$

We therefore obtain

$$P_\kappa(l) = \frac{9H_0^4\Omega_0^2}{4c^4}\int_0^{w_{\rm H}} {\rm d}w\,\frac{\bar W^2(w)}{a^2(w)}\,P_\delta\left(\frac{l}{f_K(w)},w\right)\,, \tag{6.22}$$

with the weighting function $\bar W$ given in eq. (6.19). This power spectrum is the central quantity for the discussion in the remainder of this chapter.

Figure 15 shows $P_\kappa(l)$ for five different realisations of the CDM cosmogony. These are the four models whose parameters are detailed in Tab. 1, all with non-linearly evolving density power spectrum P_δ, using the prescription of Peacock & Dodds (1996), plus the SCDM model with linearly evolving P_δ. Sources are assumed to be at redshift $z_{\rm s} = 1$. Curves 1 and 2 (solid and dotted; SCDM with linear and non-linear evolution, respectively) illustrate the impact of non-linear density evolution in an Einstein-de Sitter universe with cluster-normalised density fluctuations. Non-linear effects set in on angular scales below a few times $10'$, and increase the amplitude of $P_\kappa(l)$ by more than an order of magnitude on scales of $\approx 1'$. Curve 3 (short-dashed; σCDM), obtained for CDM normalised to $\sigma_8 = 1$ rather than the cluster abundance, demonstrates the potential influence of different choices for the power-spectrum normalisation. Curves 4 and 5 (dashed-dotted and long-dashed; OCDM and ΛCDM, respectively) show $P_\kappa(l)$ for cluster-normalised CDM in an open universe ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0$) and in a spatially flat, low-density universe ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$). It is a consequence of the normalisation to the local cluster abundance that the various $P_\kappa(l)$ are very similar for the different cosmologies on angular scales of a few arc minutes. For the low-density universes, the difference between the cluster- and the σ_8 normalisation is substantially smaller than for the Einstein-de Sitter model.

Figure 16 gives another representation of the curves in Fig. 15. There, we plot $l^2\,P_\kappa(l)$, i.e. the total power in the effective convergence per logarithmic l interval. This representation demonstrates that density fluctuations on angular scales smaller than $\approx 10'$ contribute most strongly to weak gravitational lensing by large-scale structures. On angular scales smaller than $\approx 1'$, the curves level off and then decrease very gradually. The solid curve in Fig 16 shows that, when linear density evolution is assumed, most power is contributed by structures on scales above $10'$, emphasising that it is crucial to take non-linear evolution into account to avoid misleading conclusions.

6.4.2. *Special Cases*

In the approximation of linear density evolution, applicable on large angular scales $\gtrsim 30'$, the density contrast grows in proportion with $a\,g(a)$, as described following eq. (2.52) on page 9. The power spectrum of the density contrast then evolves $\propto a^2\,g^2(a)$. Inserting this into eq. (6.22), the squared scale factor $a^2(w)$ can-

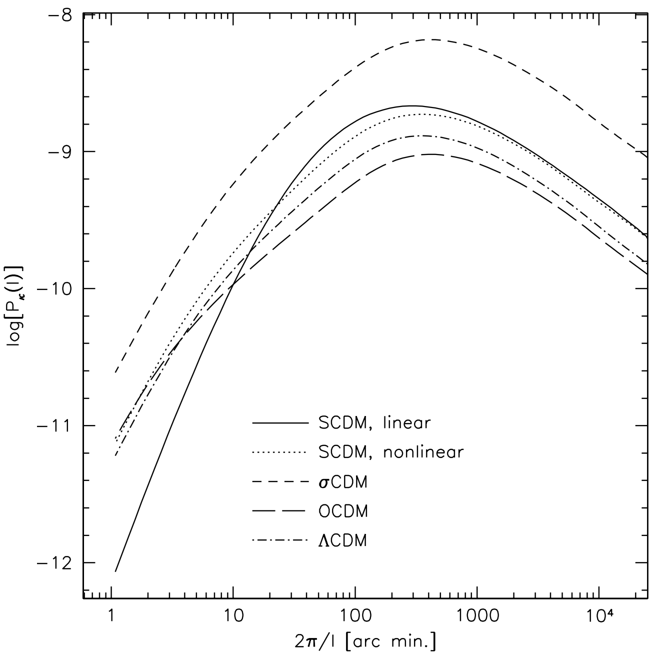

FIG. 15.—Five effective-convergence power spectra $P_\kappa(l)$ are shown as functions of the angular scale $2\pi l^{-1}$, expressed in arc minutes. All sources were assumed to lie at $z_s = 1$. The five curves represent the four realisations of the CDM cosmogony listed in Tab. 1, all with non-linearly evolving density-perturbation power spectra P_δ, plus the SCDM model with linearly evolving P_δ. Solid curve (1): Linearly evolving SCDM model; dotted curve (2): non-linearly evolving SCDM; short-dashed curve (3): non-linearly evolving σCDM; dashed-dotted and long-dashed curves (4 and 5): non-linearly evolving OCDM and ΛCDM, respectively.

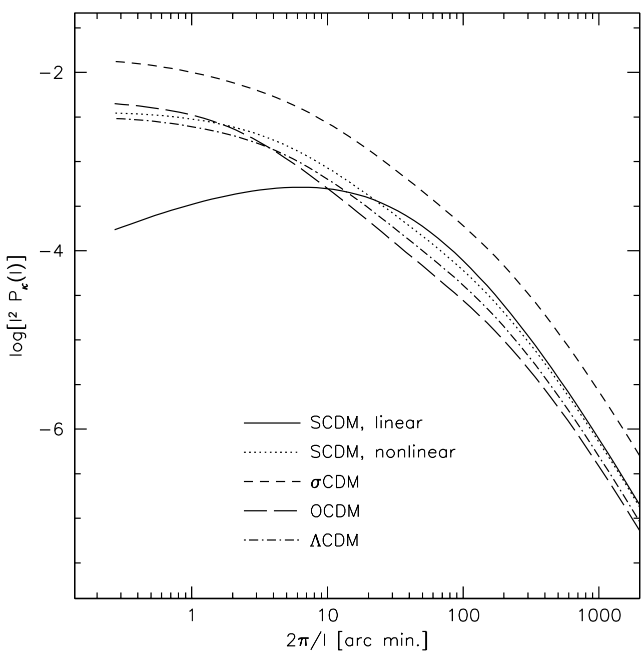

FIG. 16.—Different representation of the curves in Fig. 15. We plot here $l^2 P_\kappa(l)$, representing the total power in the effective convergence per logarithmic l interval. See the caption of Fig. 15 for the meaning of the different line types. The figure demonstrates that the total power increases monotonically towards small angular scales when non-linear evolution is taken into account (i.e. with the exception of the solid curve). On angular scales still smaller than $\approx 1'$, the curves level off and decrease very slowly. This shows that weak lensing by cosmological mass distributions is mostly sensitive to structures smaller than $\approx 10'$.

cels, and we find

$$P_\kappa(l) = \frac{9H_0^4\Omega_0^2}{4c^4} \int_0^{w_H} \mathrm{d}w\, g^2[a(w)]\, \bar{W}^2(w)\, P_\delta^0\left(\frac{l}{f_K(w)}\right) . \tag{6.23}$$

Here, $P_\delta^0(k)$ is the density-contrast power spectrum linearly extrapolated to the present epoch.

In an Einstein-de Sitter universe, the growth function $g(a)$ is unity since P_δ grows like the squared scale factor. In that special case, the expression for the power spectrum of $\bar{\kappa}_{\rm eff}$ further reduces to

$$P_\kappa(l) = \frac{9H_0^4}{4c^4} \int_0^{w_H} \mathrm{d}w\, \bar{W}^2(w)\, P_\delta^0\left(\frac{l}{w}\right) , \tag{6.24}$$

and the weight function $\bar{W}$ simplifies to

$$\bar{W}(w) = \int_w^{w_H} \mathrm{d}w'\, G(w') \left(1 - \frac{w}{w'}\right) . \tag{6.25}$$

In some situations, the distance distribution of the sources can be approximated by a delta peak at some distance w_s, $G(w) = \delta_D(w - w_s)$. A typical example is weak lensing of the Cosmic Microwave Background, where the source is the surface of last scattering at redshift $z_s \approx 1000$. Under such circumstances,

$$\bar{W}(w) = \left(1 - \frac{w}{w_s}\right) \mathrm{H}(w_s - w) , \tag{6.26}$$

where the Heaviside step function $\mathrm{H}(x)$ expresses the fact that sources at w_s are only lensed by mass distributions at smaller distance w. For this specific case, the effective-convergence power spectrum reads

$$P_\kappa(l) = \frac{9H_0^4}{4c^4}\, w_s \int_0^1 \mathrm{d}y\, (1-y)^2\, P_\delta^0\left(\frac{l}{w_s y}\right) , \tag{6.27}$$

where $y = w/w_s$ is the distance ratio between lenses and sources. This equation illustrates that all density-perturbation modes whose wave numbers are larger than $k_{\rm min} = w_s^{-1} l$ contribute to $P_\kappa(l)$, or whose wavelengths are smaller than $\lambda_{\rm max} = w_s\theta$. For example, the power spectrum of weak lensing on angular scales of $\theta \approx 10'$ on sources at redshifts $z_s \approx 2$ originates from all density perturbations smaller than $\approx 7\,h^{-1}$ Mpc. This result immediately illustrates the limitations of the foregoing approximations. Density perturbations on scales smaller than a few Mpc become non-linear even at moderate redshifts, and the assumption of linear evolution breaks down.

6.5. *Magnification and Shear*

In analogy to the Jacobian matrix $\mathcal{A}$ of the conventional lens equation (3.11, page 17), we now form the matrix

$$\mathcal{A}(\vec{\theta}, w) = I - \frac{\partial\vec{\alpha}(\vec{\theta}, w)}{\partial\vec{\theta}} = \frac{1}{f_K(w)} \frac{\partial\vec{x}(\vec{\theta}, w)}{\partial\vec{\theta}} . \tag{6.28}$$

The magnification is the inverse of the determinant of $\mathcal{A}$ (see eq. 3.14, page 17). To first order in the perturbations, we obtain for the magnification of a source at distance w seen in direction $\vec{\theta}$

$$\mu(\vec{\theta}, w) \quad = \quad \frac{1}{\det\mathcal{A}(\vec{\theta}, w)} \approx 1 + \nabla_\theta \cdot \vec{\alpha}(\vec{\theta}, w) = 1 + 2\kappa_{\rm eff}(\vec{\theta}, w)$$

$$\equiv \quad 1+\delta\mu(\vec{\theta},w)\ . \tag{6.29}$$

In the weak-lensing approximation, the magnification fluctuation $\delta\mu$ is simply twice the effective convergence $\kappa_{\rm eff}$, just as in the thin-lens approximation.

We emphasise again that the approximations made imply that the matrix $\mathcal{A}$ is symmetric. In general, when higher-order terms in the Newtonian potential are considered, $\mathcal{A}$ attains an asymmetric contribution. Jain et al. (2000) used ray-tracing simulations through the density distribution of the Universe computed in very high resolution N-body simulations to show that the symmetry of $\mathcal{A}$ is satisfied to very high accuracy. Only for those light rays which happen to propagate close to more than one strong deflector can the deviation from symmetry be appreciable. Further estimates of the validity of the various approximations have been carried out analytically by Bernardeau et al. (1997) and Schneider et al. (1998a).

Therefore, as in the single lens-plane situation, the anisotropic deformation, or shear, of a light bundle is determined by the trace-free part of the matrix $\mathcal{A}$ (cf. eq. 3.11, page 17). As explained there, the shear makes elliptical images from circular sources. Let a and b be the major and minor axes of the image ellipse of a circular source, respectively, then the ellipticity is

$$|\chi| = \frac{a^2-b^2}{a^2+b^2} \approx 2|\gamma|\ , \tag{6.30}$$

where the latter approximation is valid for weak lensing, $|\gamma| \ll 1$; cf. eq. (4.18). The quantity 2γ was sometimes called *polarisation* in the literature (Blandford et al. 1991, Miralda-Escudé 1991, Kaiser 1992).

In the limit of weak lensing which is relevant here, the two-point statistical properties of $\delta\mu$ and of 2γ are identical (e.g. Blandford et al. 1991). To see this, we first note that the first derivatives of the deflection angle occurring in eqs. (6.29) can be written as second derivatives of an effective deflection potential ψ which is defined in terms of the effective surface mass density $\kappa_{\rm eff}$ in the same way as in the single lens-plane case; see (3.9, page 17). We then imagine that $\delta\mu$ and γ are Fourier transformed, whereupon the derivatives with respect to θ_i are replaced by multiplications with components of the wave vector $\vec{l}$ conjugate to $\vec{\theta}$. In Fourier space, the expressions for the averaged quantities $\langle\delta\mu^2\rangle$ and $4\langle|\gamma|^2\rangle$ differ only by the combinations of l_1 and l_2 which appear under the average. We have

$$\begin{array}{lll} \left(l_1^2+l_2^2\right)^2 = |\vec{l}|^4 & \text{for} & \langle\delta\mu^2\rangle \\ \left(l_1^2-l_2^2\right)^2 + 4\,l_1^2 l_2^2 = |\vec{l}|^4 & \text{for} & 4\langle|\gamma|^2\rangle = 4\langle\gamma_1^2+\gamma_2^2\rangle \end{array}\ , \tag{6.31}$$

and hence the two-point statistical properties of $\delta\mu$ and 2γ agree identically. Therefore, the power spectra of effective convergence and shear agree,

$$\langle\hat{\kappa}_{\rm eff}(\vec{l})\hat{\kappa}^*_{\rm eff}(\vec{l}')\rangle = \langle\hat{\gamma}(\vec{l})\hat{\gamma}^*(\vec{l}')\rangle \quad \Rightarrow \quad P_\kappa(l) = P_\gamma(l)\ . \tag{6.32}$$

Thus we can concentrate on the statistics of either the magnification fluctuations or the shear only. Since $\delta\mu = 2\kappa_{\rm eff}$, the magnification power spectrum P_μ is $4P_\kappa$, and we can immediately employ the convergence power spectrum P_κ.

6.6. *Second-Order Statistical Measures*

We aim at the statistical properties of the magnification fluctuation and the shear. In particular, we are interested in the amplitude of these quantities and their angular coherence. Both can be described by their angular auto-correlation functions, or other second-order statistical measures that will turn out to be more practical later. As long as the density fluctuation field δ remains Gaussian, the probability distributions of $\delta\mu$ and γ are also Gaussians with mean zero, and two-point statistical measures are sufficient for their complete statistical description. When non-linear evolution of the density contrast sets in, non-Gaussianity develops, and higher-order statistical measures become important.

6.6.1. *Angular Auto-Correlation Function*

The angular autocorrelation function $\xi_q(\phi)$ of some isotropic quantity $q(\vec{\theta})$ is the Fourier transform of the power spectrum $P_q(l)$ of $q(\vec{\theta})$. In particular, the auto-correlation function of the magnification fluctuation, $\xi_\mu(\phi)$, is related to the effective-convergence power spectrum $P_\kappa(l)$ through

$$\begin{aligned} \xi_\mu(\phi) &= \langle\delta\mu(\vec{\theta})\delta\mu(\vec{\theta}+\vec{\phi})\rangle = 4\,\langle\kappa_{\rm eff}(\vec{\theta})\kappa_{\rm eff}(\vec{\theta}+\vec{\phi})\rangle \\ &= 4\,\langle\gamma(\vec{\theta})\gamma^*(\vec{\theta}+\vec{\phi})\rangle \\ &= 4\int\frac{{\rm d}^2 l}{(2\pi)^2}\,P_\kappa(l)\,\exp(-{\rm i}\vec{l}\cdot\vec{\phi}) \\ &= 4\int_0^\infty\frac{l{\rm d}l}{2\pi}\,P_\kappa(l)\,{\rm J}_0(l\phi)\ , \end{aligned} \tag{6.33}$$

where $\vec{\phi}$ is a vector with norm ϕ. The factor four in front of the integral accounts for the fact that $\delta\mu = 2\kappa_{\rm eff}$ in the weak-lensing approximation. For the last equality in (6.33), we integrated over the angle enclosed by $\vec{l}$ and $\vec{\phi}$, leading to the zeroth-order Bessel function of the first kind, ${\rm J}_0(x)$. Equation (6.33) shows that the magnification (or shear) auto-correlation function is an integral over the power spectrum of the effective convergence $\kappa_{\rm eff}$, filtered by the Bessel function ${\rm J}_0(x)$. Since the latter is a broad-band filter, the magnification auto-correlation function is not well suited for extracting information on P_κ. It would be desirable to replace $\xi_\mu(\phi)$ by another measurable quantity which involves a narrow-band filter.

Nonetheless, inserting eq. (6.22) into eq. (6.33), we obtain the expression for the magnification auto-correlation function,

$$\begin{aligned} \xi_\mu(\phi) = \; & \frac{9H_0^4\Omega_0^2}{c^4}\int_0^{w_{\rm H}}{\rm d}w\,f_K^2(w)\,\bar{W}^2(w,w)\,a^{-2}(w) \\ & \times \int_0^\infty\frac{k{\rm d}k}{2\pi}\,P_\delta(k,w)\,{\rm J}_0[f_K(w)k\phi]\ . \end{aligned} \tag{6.34}$$

The magnification autocorrelation function therefore turns out to be an integral over the density-fluctuation power spectrum weighted by a k–space window function which selects the contributing density perturbation modes.

The correlation function of the image ellipticity (or the shear) is then $\langle\varepsilon\varepsilon^*\rangle(\phi) = \xi_\gamma(\phi) = \xi_\mu(\phi)/4$. Since the ellipticity has two components, one can define and calculate the corresponding correlations functions as well: Any pair of galaxy images defines the direction φ of their separation vector. With respect to this direction, one can define in complete analogy to (5.31) the tangential and cross-components of the ellipticities, $\varepsilon_{\rm t}$ and $\varepsilon_\times = -\Im(\varepsilon\exp(-2{\rm i}\varphi))$, respectively. One then finds (Kaiser 1992) that

$$\left(\begin{array}{c}\langle\varepsilon_{\rm t}\varepsilon_{\rm t}\rangle(\phi)\\ \langle\varepsilon_\times\varepsilon_\times\rangle(\phi)\end{array}\right) = \frac{1}{2}\int_0^\infty\frac{l{\rm d}l}{2\pi}\,P_\kappa(l)\left(\begin{array}{c}{\rm J}_0(l\phi)+{\rm J}_4(l\phi)\\ {\rm J}_0(l\phi)-{\rm J}_4(l\phi)\end{array}\right)\ , \tag{6.35}$$

and $\langle\varepsilon_{\rm t}\varepsilon_\times\rangle(\phi) = 0$. The latter expression can be used to estimate systematic errors on a given data set from which the correlation functions are calculated.

6.6.2. Special Cases and Qualitative Expectations

In order to gain some insight into the expected behaviour of the magnification auto-correlation function $\xi_\mu(\phi)$, we now make a number of simplifying assumptions. Let us first specialise to linear density evolution in an Einstein-de Sitter universe, and assume sources are at a single distance w_s. Equation (6.34) then immediately simplifies to

$$\xi_\mu(\phi) = \frac{9H_0^4}{c^4}\,w_s^3\int_0^1 \mathrm{d}y\,y^2(1-y)^2\int_0^\infty \frac{k\mathrm{d}k}{2\pi}\,P_\delta^0(k)\,\mathrm{J}_0(wyk\phi)\;, \tag{6.36}$$

with $y \equiv w_s^{-1}w$.

We now introduce two model spectra $P_\delta^0(k)$, one of which has an exponential cut-off above some wave number k_0, while the other falls off like k^{-3} for $k > k_0$. For small k, both spectra increase like k. They approximately describe two extreme cases of popular cosmogonies, the HDM and the CDM model. We choose the functional forms

$$P_{\delta,\mathrm{HDM}}^0 = A\,k\,\exp\left(-\frac{k}{k_0}\right)\;,\quad P_{\delta,\mathrm{CDM}}^0 = A\,k\,\frac{9k_0^4}{(k^2+3k_0^2)^2}\;, \tag{6.37}$$

where A is the normalising amplitude of the power spectra. The numerical coefficients in the CDM model spectrum are chosen such that both spectra peak at the same wave number $k = k_0$. Inserting these model spectra into eq. (6.36), performing the k integration, and expanding the result in a power series in ϕ, we obtain (Bartelmann 1995b)

$$\begin{aligned}
\xi_{\mu,\mathrm{HDM}}(\phi) &= \frac{3A'}{10\pi}\,(w_sk_0)^3 - \frac{9A'}{35\pi}\,(w_sk_0)^5\,\phi^2 + O(\phi^4)\;,\\
\xi_{\mu,\mathrm{CDM}}(\phi) &= \frac{9\sqrt{3}A'}{80}\,(w_sk_0)^3 - \frac{27A'}{40\pi}\,(w_sk_0)^4\,\phi + O(\phi^2)\;,
\end{aligned} \tag{6.38}$$

where $A' = (H_0c^{-1})^4A$. We see from eq. (6.38) that the magnification correlation function for the HDM spectrum is flat to first order in ϕ, while it decreases linearly with ϕ for the CDM spectrum. This demonstrates that the shape of the magnification autocorrelation function $\xi_\mu(\phi)$ reflects the shape of the dark-matter power spectrum. Motivated by the result of a large number of cosmological studies showing that HDM models have the severe problem of structure on small scales forming at times much later than observed (see e.g. Peacock 1999), we now neglect the HDM model and focus on the CDM power spectrum only.

We can then expect $\xi_\mu(\phi)$ to increase linearly with ϕ as ϕ goes to zero. Although we assumed linear evolution of the power spectrum to achieve this result, this qualitative behaviour remains valid when non-linear evolution is assumed, because for large wave numbers k, the non-linear CDM power spectra also asymptotically fall off $\propto k^{-3}$ for large k.

Although the model spectra (6.37) are of limited validity, we can extract some useful information from the small-angle approximations given in eq. (6.38). First, the correlation amplitude $\xi_\mu(0)$ scales with the comoving distance to the sources w_s as w_s^3. In the Einstein-de Sitter case, for which eq. (6.38) was derived, $w_s = (2c/H_0)\left[1-(1+z_s)^{-1/2}\right]$. For low source redshifts, $z_s \ll 1$, $w_s \approx (c/H_0)\,z_s$, so that $\xi_\mu(0) \propto z_s^3$. For $z_s \gg 1$, $w_s \to (2c/H_0)$, and $\xi_\mu(0)$ becomes independent of source redshift. For intermediate source redshifts, progress can be made by defining $\zeta_s \equiv \ln(z_s)$ and expanding $\ln w[\exp(\zeta_s)]$ in a power series in ζ_s. The result is an approximate power-law expression, $w(z_s) \propto z_s^\varepsilon$, valid in the vicinity of the zero point of the expansion. The exponent ε changes from ≈ 0.6 at $z_s \approx 1$ to ≈ 0.38 at $z_s \approx 3$.

Second, typical source distances are of order 2 Gpc. Since k_0 is the wave number corresponding to the horizon size when relativistic and non-relativistic matter had equal densities, $k_0^{-1} = d_\mathrm{H}(a_\mathrm{eq}) \approx 12\,(\Omega_0h^2)^{-1}$ Mpc. Therefore, $w_sk_0 \approx 150$. Typically, the spectral amplitude A' ranges between 10^{-8}–10^{-9}. A rough estimate for the correlation amplitude $\xi_\mu(0)$ thus ranges between 10^{-2}–10^{-3} for 'typical' source redshifts $z_s \gtrsim 1$.

Third, an estimate for the angular scale ϕ_0 of the magnification correlation is obtained by determining the angle where $\xi_\mu(\phi)$ has dropped to half its maximum. From the small-angle approximation (6.38), we find $\phi_0 = \pi\sqrt{3}(12\,w_sk_0)^{-1}$. Inserting as before $w_sk_0 \approx 150$, we obtain $\phi_0 \approx 10'$, *decreasing* with increasing source redshift.

Summarising, we expect $\xi_\mu(\phi)$ in a CDM universe to

1. start at 10^{-2}–10^{-3} at $\phi = 0$ for source redshifts $z_s \sim 1$;
2. decrease linearly for small ϕ on an angular scale of $\phi_0 \approx 10'$; and
3. increase with source redshift roughly as $\propto z_s^{0.6}$ around $z_s = 1$.

6.6.3. Realistic Cases

After this digression, we now return to realistic CDM power spectra normalised to fit observational constraints. Some representative results are shown in Fig. 17 for the model parameter sets listed in Tab. 1.

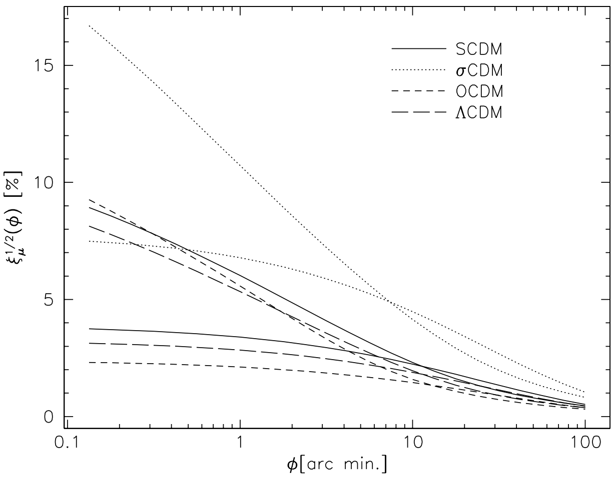

FIG. 17.—Four pairs of magnification auto-correlation functions are shown for the cosmological model parameter sets listed in Tab. 1, and for an assumed source redshift $z_s = 1$. For each pair, plotted with the same line type, the curve with lower amplitude at small angular scale was calculated assuming linear, and the other one non-linear density evolution. Solid curves: SCDM; dotted curves: σCDM; short-dashed curve: OCDM; and long-dashed curve: ΛCDM. Non-linear evolution increases the amplitude of $\xi_\mu^{1/2}(\phi)$ on small angular scales by factors of three to four. The results for the cluster-normalised models differ fairly little. At $\phi \approx 1'$, $\xi_\mu^{1/2}(\phi) \approx 6\%$ for non-linear density evolution. For the Einstein-de Sitter models, the difference between cluster- and $\sigma_8 = 1$ normalisation amounts to about a factor of two in $\xi_\mu^{1/2}(\phi)$.

The figure shows that typical values for $\xi_\mu^{1/2}(\phi)$ in cluster-normalised CDM models with non-linear density evolution are $\approx 6\%$ at $\phi \approx 1'$, quite independent of the cosmological model. The effects of non-linear evolution are considerable. Non-linear evolution increases the $\xi_\mu^{1/2}$ by factors of three to four. The uncertainty in the normalisation is illustrated by the two curves for the Einstein-de Sitter model, one of which was calculated with the cluster-, the other one with the $\sigma_8 = 1$ normalisation, which yields about a factor of two larger results for $\xi_\mu^{1/2}$. For the other cosmological models (OCDM and ΛCDM), the effects of different normalisations (cluster vs. COBE) are substantially smaller.

6.6.4. *Application: Magnification Fluctuations*

At zero lag, the magnification autocorrelation function reads

$$\xi_\mu(0) = \left\langle \left[\mu(\vec\theta) - 1\right]^2 \right\rangle \equiv \left\langle \delta\mu^2 \right\rangle , \qquad (6.39)$$

which is the variance of the magnification fluctuation $\delta\mu$. Consequently, the *rms* magnification fluctuation is

$$\delta\mu_{\rm rms} = \left\langle \delta\mu^2 \right\rangle^{1/2} = \xi_\mu^{1/2}(0) . \qquad (6.40)$$

Figure 18 shows $\delta\mu_{\rm rms}$ as a function of source redshift for four different realisations of the CDM cosmogony. For cluster-normalised CDM models, the *rms* magnification fluctuation is of order $\delta\mu_{\rm rms} \approx 20\%$ for sources at $z_{\rm s} \approx 2$, and increases to $\delta\mu_{\rm rms} \approx 25\%$ for $z_{\rm s} \approx 3$. The strongest effect occurs for open CDM (OCDM) because there non-linear evolution sets in at the highest redshifts.

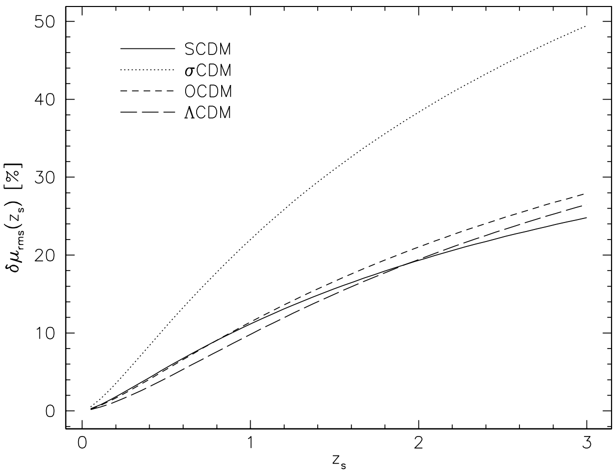

FIG. 18.—The *rms* magnification fluctuation $\delta\mu_{\rm rms}$ is shown as a function of source redshift $z_{\rm s}$ for non-linearly evolving density fluctuations in the four different realisations of the CDM cosmogony detailed in Tab. 1. Solid curve: SCDM; dotted curve: σCDM; short-dashed curve: OCDM; and long-dashed curve: ΛCDM. Except for the σCDM model, typical *rms* magnification fluctuations are of order 20% at $z_{\rm s} = 2$, and 25% for $z_{\rm s} = 3$.

The results shown in Fig. 18 indicate that for any cosmological source, gravitational lensing causes a statistical uncertainty of its brightness. In magnitudes, a typical effect at $z_{\rm s} \approx 2$ is $\delta m \approx 2.5 \times \log(1.2) \approx 0.2$. This can be important for e.g. high-redshift supernovae of type Ia, which are used as cosmological standard candles. Their intrinsic magnitude scatter is of order $\delta m \approx 0.1 - 0.2$ magnitudes (e.g. Phillips 1993; Riess et al. 1995, 1996; Hamuy et al. 1996). Therefore, the lensing-induced brightness fluctuation is comparable to the intrinsic uncertainty at redshifts $z_{\rm s} \gtrsim 2$ (Frieman 1996; Wambsganss et al. 1997; Holz 1998; Metcalf & Silk 1999).

Since the magnification probability can be highly skewed, the *most probable* observed flux of a high-redshift supernova can deviate from the *mean* flux at given redshift, even if the intrinsic luminosity distribution is symmetric. This means that particular care needs to be taken in the analysis of future large SN surveys. However, if SNe Ia are quasi standard candles also at high redshifts, with an intrinsic scatter of $\Delta L = 4\pi D_{\rm lum}^2(z)\Delta S(z)$ around the mean luminosity $L_0 = 4\pi D_{\rm lum}^2(z) S_0(z)$, then it is possible to obtain *volume-limited samples* (in contrast to flux-limited samples) of them.

If, for a given redshift, the sensitivity limit is chosen to be $S_{\rm min} \lesssim \mu_{\rm min}\,(S_0 - 3\Delta S)$, one can be sure to find all SNe Ia at the redshift considered. Here, $\mu_{\rm min}$ is the minimum magnification of a source at the considered redshift. Since no source can be more de-magnified than one that is placed behind a hypothetical empty cone (see Dyer & Roder 1973 and the discussion in Sect. 4.5 of Schneider et al. 1992), $\mu_{\rm min}$ is not much smaller than unity. Flux conservation (e.g. Weinberg 1976) implies that the mean magnification of all sources at given redshift is unity, $\langle\mu(z)\rangle = 1$, and so the expectation value of the observed flux at given redshift is the unlensed flux, $\langle S(z)\rangle = S_0(z)$. It should be pointed out here that a similar relation for the magnitudes does *not* hold, since magnitude is a logarithmic measure of the flux, and so $\langle m(z)\rangle \neq m_0(z)$. This led to some confusing conclusions in the literature claiming that lensing introduces a bias in cosmological parameter estimates from lensing, but this is not true: One just has to work in terms of fluxes rather than magnitudes.

However, a broad magnification probability distribution increases the confidence contours for Ω_0 and Ω_Λ (e.g. Holz 1998). If the probability distribution was known, more sensitive estimators of the cosmological model than the mean flux at given redshift could be constructed. Furthermore, if the intrinsic luminosity distribution of the SNe was known, the normalisation of the power spectrum as a function of Ω_0 and Ω_Λ could be inferred from the broadened observed flux distribution (Metcalf 1999). If part of the dark matter is in the form of compact objects with mass $\gtrsim 10^{-2} M_\odot$, these objects can individually magnify a SN (Schneider & Wagoner 1987), additionally broadening the magnification probability distribution and thus enabling the nature of dark matter to be tested through SN observations (Metcalf & Silk 1999, Seljak & Holz 1999).

6.6.5. *Shear in Apertures*

We mentioned below eq. (6.33) that measures of cosmic magnification or shear other than the angular auto-correlation function which filter the effective-convergence power spectrum P_κ with a function narrower than the Bessel function $\mathrm{J}_0(x)$ would be desirable. In practice, a convenient measure would be the variance of the effective convergence within a circular aperture of radius θ. Within such an aperture, the averaged effective convergence and shear are

$$\kappa_{\rm av}(\theta) = \int_0^\theta \frac{\mathrm{d}^2\phi}{\pi\theta^2}\,\bar\kappa_{\rm eff}(\vec\phi) , \quad \gamma_{\rm av}(\theta) = \int_0^\theta \frac{\mathrm{d}^2\phi}{\pi\theta^2}\,\gamma(\vec\phi) , \qquad (6.41)$$

and their variance is

$$\langle\kappa_{\rm av}^2\rangle(\theta) = \int_0^\theta \frac{\mathrm{d}^2\phi}{\pi\theta^2}\int_0^\theta \frac{\mathrm{d}^2\phi'}{\pi\theta^2}\left\langle \bar\kappa_{\rm eff}(\vec\phi)\bar\kappa_{\rm eff}(\vec\phi')\right\rangle = \langle|\gamma_{\rm av}|^2\rangle(\theta) . \qquad (6.42)$$

The remaining average is the effective-convergence auto-correlation function $\xi_\kappa(|\vec\phi - \vec\phi'|)$, which can be expressed in terms of the power spectrum P_κ. The final equality follows from $\xi_\kappa = \xi_\gamma$. Inserting (6.42) and performing the angular integrals yields

$$\langle\kappa_{\rm av}^2\rangle(\theta) = 2\pi\int_0^\infty l\,{\rm d}l\,P_\kappa(l)\,\left[\frac{{\rm J}_1(l\theta)}{\pi l\theta}\right]^2 = \langle|\gamma_{\rm av}|^2\rangle(\theta)\;, \quad (6.43)$$

where ${\rm J}_1(x)$ is the first-order Bessel function of the first kind. Results for the *rms* shear in apertures of varying size are shown in Fig. 19 (cf. Blandford et al. 1991, Kaiser 1992, Jain & Seljak 1997).

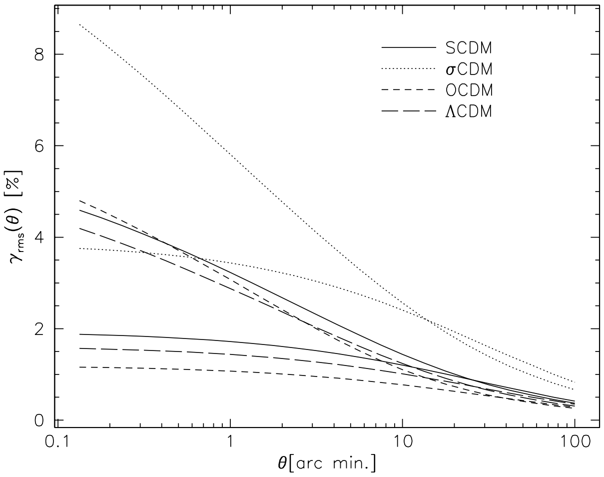

FIG. 19.—The *rms* shear $\gamma_{\rm rms}(\theta)$ in circular apertures of radius θ is plotted as a function of θ for the four different realisations of the CDM cosmogony detailed in Tab. 1, where all sources are assumed to be at redshift $z_{\rm s}=1$. A pair of curves is plotted for each realisation, where for each pair the curve with lower amplitude at small θ is for linearly, the other one for non-linearly evolving density fluctuations. Solid curves: SCDM; dotted curves: σCDM; short-dashed curves: OCDM; and long-dashed curves: ΛCDM. For the cluster-normalised models, typical *rms* shear values are $\approx 3\%$ for $\theta\approx 1'$. Non-linear evolution increases the amplitude by about a factor of two at $\theta\approx 1'$ over linear evolution.

6.6.6. Aperture Mass

Another measure for the effects of weak lensing, the *aperture mass* $M_{\rm ap}(\theta)$ (cf. Sect. 5.3.1.), was introduced for cosmic shear by Schneider et al. (1998a) as

$$M_{\rm ap}(\theta) = \int_0^\theta {\rm d}^2\phi\,U(\phi)\,\bar\kappa_{\rm eff}(\vec\phi)\;, \quad (6.44)$$

where the weight function $U(\phi)$ satisfies the criterion

$$\int_0^\theta \phi{\rm d}\phi\,U(\phi) = 0\;. \quad (6.45)$$

In other words, $U(\phi)$ is taken to be a *compensated* radial weight function across the aperture. For such weight functions, the aperture mass can be expressed in terms of the tangential component of the observable shear relative to the aperture centre,

$$M_{\rm ap}(\theta) = \int_0^\theta {\rm d}^2\phi\,Q(\phi)\,\gamma_{\rm t}(\vec\phi)\;, \quad (6.46)$$

where $Q(\phi)$ is related to $U(\phi)$ by (5.24). $M_{\rm ap}$ is a scalar quantity directly measurable in terms of the shear. The variance of $M_{\rm ap}$ reads

$$\langle M_{\rm ap}^2\rangle(\theta) = 2\pi\int_0^\infty l\,{\rm d}l\,P_\kappa(l)\,\left[\int_0^\theta \phi{\rm d}\phi\,U(\phi)\,{\rm J}_0(l\phi)\right]^2\;. \quad (6.47)$$

Equations (6.43) and (6.47) provide alternative observable quantities which are related to the effective-convergence power spectrum P_κ through narrower filters than the auto-correlation function ξ_κ. The $M_{\rm ap}$ statistic in particular permits one to tune the filter function through different choices of $U(\phi)$ within the constraint (6.45). It is important that $M_{\rm ap}$ can also be expressed in terms of the shear [see eq. (5.27, page 37)], so that $M_{\rm ap}$ can directly be obtained from the observed galaxy ellipticities.

Schneider et al. (1998a) suggested a family of radial filter functions $U(\phi)$, the simplest of which is

$$U(\phi) = \frac{9}{\pi\theta^2}(1-x^2)\left(\frac{1}{3}-x^2\right)\;,\quad Q(\phi) = \frac{6}{\pi\theta^2}x^2(1-x^2)\;, \quad (6.48)$$

where $x\theta=\phi$. With this choice, the variance $\langle M_{\rm ap}^2\rangle(\theta)$ becomes

$$\langle M_{\rm ap}^2\rangle(\theta) = 2\pi\int_0^\infty l\,{\rm d}l\,P_\kappa(l)\,J^2(l\theta)\;, \quad (6.49)$$

with the filter function

$$J(\eta) = \frac{12}{\pi\eta^2}{\rm J}_4(\eta)\;, \quad (6.50)$$

where ${\rm J}_4(\eta)$ is the fourth-order Bessel function of the first kind. Examples for the *rms* aperture mass, $M_{\rm ap,rms}(\theta) = \langle M_{\rm ap}^2\rangle^{1/2}(\theta)$, are shown in Fig. 20.

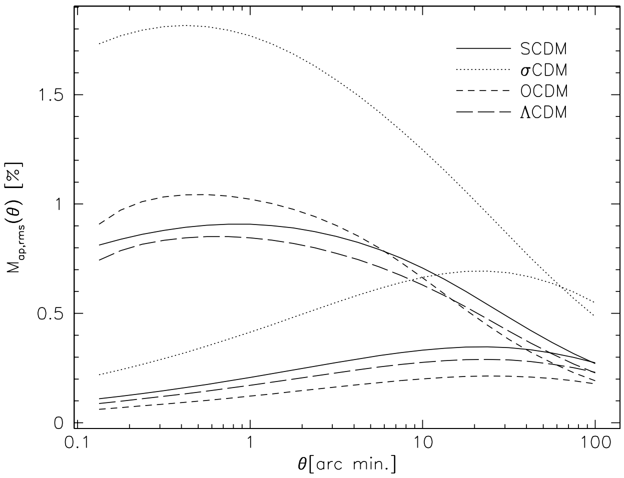

FIG. 20.—The *rms* aperture mass, $M_{\rm ap,rms}(\theta)$, is shown in dependence of aperture radius θ for the four different realisations of the CDM cosmogony detailed in Tab. 1 where all sources are assumed to be at redshift $z_{\rm s}=1$. For each realisation, a pair of curves is plotted; one curve with lower amplitude for linear, and the second curve for non-linear density evolution. Solid curves: SCDM; dotted curves: σCDM; short-dashed curves: OCDM; and long-dashed curves: ΛCDM. Non-linear evolution has a pronounced effect: The amplitude is approximately doubled, and the peak shifts from degree- to arc-minute scales.

The curves look substantially different from those shown in Figs. 17 and 19. Unlike there, the aperture mass does not increase monotonically as $\theta\to 0$, but reaches a maximum at finite θ and drops for smaller angles. When non-linear evolution of the density fluctuations is assumed, the maximum occurs at much smaller θ than for linear evolution: Linear evolution predicts the peak at angles of order one degree, non-linear evolution around $1'$! The amplitude of $M_{\rm ap,rms}(\theta)$ reaches $\approx 1\%$ for cluster-normalised models, quite independent of the cosmological parameters.

Some insight into the expected amplitude and shape of $\langle M_{\rm ap}^2\rangle(\theta)$ can be gained by noting that $J^2(\eta)$ is well approximated by a Gaussian,

$$J^2(\eta) \approx A \exp\left[-\frac{(\eta-\eta_0)^2}{2\sigma^2}\right] , \tag{6.51}$$

with mean $\eta_0 \approx 4.11$, amplitude $A \approx 4.52\times10^{-3}$, and width $\sigma \approx 1.24$. At aperture radii of $\theta \approx 1'$, the peak $\eta_0 \approx 4.11$ corresponds to angular scales of $2\pi l^{-1} \approx 1.6'$, where the total power $l^2 P_\kappa(l)$ in the effective convergence is close to its broad maximum (cf. Fig. 16). The filter function $J^2(\eta)$ is therefore fairly narrow. Its relative width corresponds to an l range of $\delta l/l \approx \sigma/\eta_0 \sim 0.3$. Thus, the contributing range of modes l in the integral (6.49) is very small. Crudely approximating the Gaussian by a delta distribution,

$$J^2(\eta) \approx A\sqrt{2\pi}\sigma\,\delta_{\rm D}(\eta-\eta_0) , \tag{6.52}$$

we are led to

$$\begin{aligned}\langle M_{\rm ap}^2\rangle &\approx \langle\tilde M_{\rm ap}^2\rangle \equiv \frac{(2\pi)^{3/2}A\sigma}{\eta_0}\left(\frac{\eta_0}{\theta}\right)^2 P_\kappa\left(\frac{\eta_0}{\theta}\right)\\ &\approx 2.15\times10^{-2}\, l_0^2 P_\kappa(l_0) ,\end{aligned} \tag{6.53}$$

with $l_0 \equiv \eta_0\theta^{-1}$. Hence, the mean-square aperture mass is expected to directly yield the total power in the effective-convergence power spectrum, scaled down by a factor of $\approx 2.15\times10^{-2}$. We saw in Fig. 16 that $l^2P_\kappa(l) \approx 3\times10^{-3}$ for $2\pi l^{-1} \approx 1'$ in cluster-normalised CDM models, so that

$$\langle M_{\rm ap}^2\rangle^{1/2} \approx 0.8\% \quad \text{at} \quad \theta \approx 1' \tag{6.54}$$

for sources at redshift unity. We compare $M_{\rm ap,rms}(\theta)$ and the approximation $\tilde M_{\rm ap,rms}(\theta)$ in Fig. 21. Obviously, the approximation is excellent for $\theta \gtrsim 10'$, but even for smaller aperture radii of $\sim 1'$ the relative deviation is less than $\approx 5\%$. At this point, the prime virtue of the narrow filter function $J(\eta)$ shows up most prominently. Up to relatively small errors of a few per cent, the *rms* aperture mass very accurately reflects the effective-convergence power spectrum $P_\kappa(l)$. Observations of $M_{\rm ap,rms}(\theta)$ are therefore most suitable to obtain information on the matter power spectrum (cf. Bartelmann & Schneider 1999).

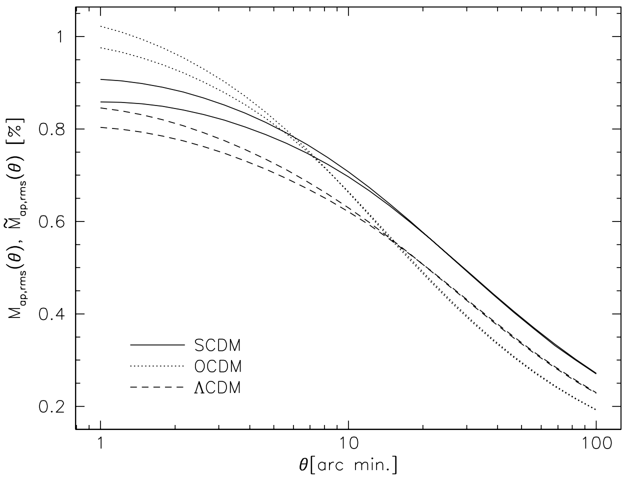

FIG. 21.—The *rms* aperture mass $M_{\rm ap,rms}(\theta)$ is shown together with the approximation $\tilde M_{\rm ap,rms}(\theta)$ of eq. (6.53). The three curves correspond to the three cluster-normalised cosmological models (SCDM, OCDM and ΛCDM) introduced in Tab. 1 for non-linearly evolving matter perturbations. All sources were assumed to be at redshift $z_{\rm s}=1$. Clearly, the *rms* aperture mass is very accurately approximated by $\tilde M_{\rm ap,rms}$ on angular scales $\theta \gtrsim 10'$, and even for smaller aperture sizes of order $\sim 1'$ the deviation between the curves is smaller than $\approx 5\%$. The observable *rms* aperture mass therefore provides a very direct measure for the effective-convergence power spectrum $P_\kappa(l)$.

6.6.7. *Power Spectrum and Filter Functions*

The three statistical measures discussed above, the magnification (or, equivalently, the shear) auto-correlation function ξ_μ, the mean-square shear in apertures $\langle\gamma^2\rangle$, and the mean-square aperture mass $\langle M_{\rm ap}^2\rangle$, are related to the effective-convergence power spectrum P_κ in very similar ways. According to eqs. (6.33), (6.43), and (6.49), they can all be written in the form

$$Q(\theta) = 2\pi\int_0^\infty l\,{\rm d}l\,P_\kappa(l)\,F(l\theta) , \tag{6.55}$$

where the filter functions $F(\eta)$ are given by

$$F(\eta) = \begin{cases} \dfrac{{\rm J}_0(\eta)}{\pi^2} & \text{for} \quad Q=\xi_\mu \\ \left[\dfrac{{\rm J}_1(\eta)}{\pi\eta}\right]^2 & \text{for} \quad Q=\langle\gamma_{\rm av}^2\rangle \\ \left[\dfrac{12\,{\rm J}_4(\eta)}{\pi\eta^2}\right]^2 & \text{for} \quad Q=\langle M_{\rm ap}^2\rangle \end{cases} . \tag{6.56}$$

Figure 22 shows these three filter functions as functions of $\eta = l\theta$. Firstly, the curves illustrate that the amplitude of ξ_μ is largest (owing to the factor of four relative to the definition of ξ_γ), and that of $\langle M_{\rm ap}^2\rangle$ is smallest because the amplitudes of the filter functions themselves decrease. Secondly, it becomes evident that, for given θ, the range of l modes of the effective-convergence power spectrum $P_\kappa(l)$ convolved into the weak-lensing estimator is largest for ξ_μ and smallest for $\langle M_{\rm ap}^2\rangle$. Thirdly, the envelope of the filter functions for large η decreases most slowly for ξ_μ and most rapidly for $\langle M_{\rm ap}^2\rangle$. Although the aperture mass has the smallest signal amplitude, it is a much better probe for the effective-convergence power spectrum $P_\kappa(l)$ than the other measures because it picks up the smallest range of l modes and most strongly suppresses the l modes smaller or larger than its peak location.

We can therefore conclude that, while the strongest weak-lensing signal is picked up by the magnification auto-correlation function ξ_μ, the aperture mass is the weak-lensing estimator most suitable for extracting information on the effective-convergence power spectrum.

6.6.8. *Signal-to-Noise Estimate of Aperture-Mass Measurements*

The question then arises whether the aperture mass can be measured with sufficient significance in upcoming wide-field imaging surveys. In practice, $M_{\rm ap}$ is derived from observations of image distortions of faint background galaxies, using eq. (5.27, page 37) and replacing the integral by a sum over galaxy ellipticities. If we consider $N_{\rm ap}$ independent apertures with N_i galaxies in the i-th aperture, an unbiased estimator of $\langle M_{\rm ap}^2\rangle$ is

$$\mathcal{M} = \frac{(\pi\theta^2)^2}{N_{\rm ap}}\sum_{i=1}^{N_{\rm ap}}\frac{1}{N_i(N_i-1)}\sum_{j\neq k}^{N_i} Q_{ij}Q_{ik}\epsilon_{{\rm t},ij}\,\epsilon_{{\rm t},ik} , \tag{6.57}$$

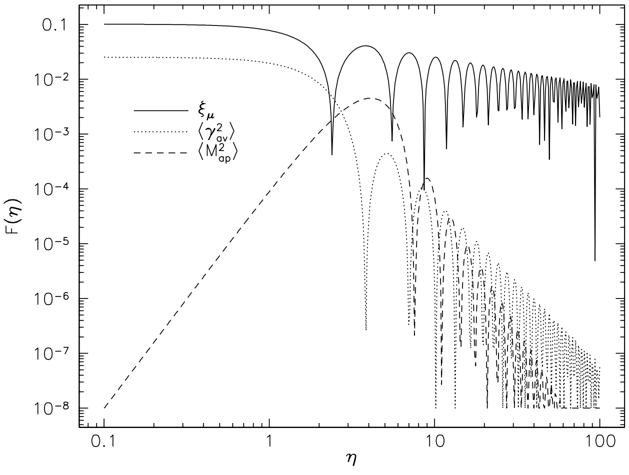

FIG. 22.—The three filter functions $F(\eta)$ defined in eq. (6.56) are shown as functions of $\eta = l\theta$. They occur in the expressions for the magnification auto-correlation function, ξ_μ (solid curve), the mean-square shear in apertures, $\langle\gamma^2\rangle$ (dotted curve), and the mean-square aperture mass, $\langle M_{\rm ap}^2\rangle$ (dashed curve).

where Q_{ij} is the value of the weight function at the position of the j-th galaxy in the i-th aperture, and $\varepsilon_{{\rm t},ij}$ is defined accordingly.

The noise properties of this estimator were investigated in Schneider et al. (1998a). One source of noise comes from the fact that galaxies are not intrinsically circular, but rather have an intrinsic ellipticity distribution. A second contribution to the noise is due to the random galaxy positions, and a third one to cosmic (or sampling) variance. Under the assumptions that the number of galaxies N_i in the apertures is large, $N_i \gg 1$, it turns out that the second of these contributions can be neglected compared to the other two. For this case, and assuming for simplicity that all N_i are equal, $N_i \equiv N$, the signal-to-noise of the estimator $\mathcal{M}$ becomes

$$\frac{\rm S}{\rm N} \equiv \frac{\langle M_{\rm ap}^2\rangle}{\sigma(\mathcal{M})} = N_{\rm ap}^{1/2}\left[\mu_4 + \left(\sqrt{2} + \frac{6\sigma_\varepsilon^2}{5\sqrt{2}N\langle M_{\rm ap}^2\rangle}\right)^2\right]^{-1}, \tag{6.58}$$

where $\sigma_\varepsilon \approx 0.2$ (e.g. Hudson et al. 1998) is the dispersion of the intrinsic galaxy ellipticities, and $\mu_4 = \langle M_{\rm ap}^4\rangle/\langle M_{\rm ap}^2\rangle^2 - 3$ is the curtosis of $M_{\rm ap}$, which vanishes for a Gaussian distribution. The two terms of (6.58) in parentheses represent the noise contributions from Gaussian sampling variance and the intrinsic ellipticity distribution, respectively, and μ_4 accounts for sampling variance in excess of that for a Gaussian distribution. On angular scales of a few arc minutes and smaller, the intrinsic ellipticities dominate the noise, while the cosmic variance dominates on larger scales.

Another convenient and useful property of the aperture mass $M_{\rm ap}$ follows from its filter function being narrow, namely that $M_{\rm ap}$ is a well localised measure of cosmic weak lensing. This implies that $M_{\rm ap}$ measurements in neighbouring apertures are almost uncorrelated even if the aperture centres are very close (Schneider et al. 1998a). It is therefore possible to gain a large number of (almost) independent $M_{\rm ap}$ measurements from a single large data field by covering the field densely with apertures. This is a significant advantage over the other two measures for weak lensing discussed above, whose broad filter functions introduce considerable correlation between neighbouring measurements, implying that for their measurement imaging data on widely separated fields are needed to ensure statistical independence. Therefore, a meaningful strategy to measure cosmic shear consists in taking a large data field, covering it densely with apertures of varying radius θ, and determining $\langle M_{\rm ap}^2\rangle$ in them via the ellipticities of galaxy images. Figure 23 shows an example for the signal-to-noise ratio of such a measurement that can be expected as a function of aperture radius θ.

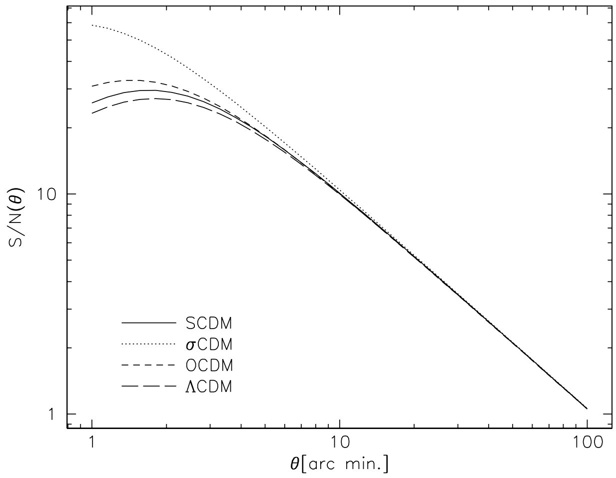

FIG. 23.—The signal-to-noise ratio $\rm S/N(\theta)$ of measurements of mean-square aperture masses $\langle M_{\rm ap}^2\rangle$ is plotted as a function of aperture radius θ for an experimental setup as described in the text. The curtosis was set to zero here. The four curves are for the four different realisations of the CDM cosmogony listed in Tab. 1. Solid curve: SCDM; dotted curve: σCDM; short-dashed curve: OCDM; and long-dashed curve: ΛCDM. Quite independently of the cosmological parameters, the signal-to-noise ratio S/N reaches values of > 10 on scales of $\approx 1' - 2'$.

Computing the curves in Fig. 23, we assumed that a data field of size $5^\circ \times 5^\circ$ is available which is densely covered by apertures of radius θ, hence the number of (almost) independent apertures is $N_{\rm ap} = (300'/2\theta)^2$. The number density of galaxies was taken as $30\,{\rm arcmin}^{-2}$, and the intrinsic ellipticity dispersion was assumed to be $\sigma_\varepsilon = 0.2$. Evidently, high signal-to-noise ratios of > 10 are reached on angular scales of $\approx 1'$ in cluster-normalised universes quite independent of the cosmological parameters. The decline of S/N for large θ is due to the decreasing number of independent apertures on the data field, whereas the decline for small θ is due to the decrease of the signal $\langle M_{\rm ap}^2\rangle$, as seen in Fig. 20. We also note that for calculating the curves in Fig. 23, we have put $\mu_4 = 0$. This is likely to be an overly optimistic assumption for small angular scales where the density field is highly non-linear. Unfortunately, μ_4 cannot easily be estimated analytically. It was numerically derived from ray-tracing through N-body simulations of large-scale matter distributions by Reblinsky et al. (1999). The curtosis exceeds unity even on scales as large as $10'$, demonstrating the highly non-Gaussian nature of the non-linearly developed density perturbations.

Although the aperture mass is a very convenient measure of cosmic shear and provides a localised estimate of the projected power spectrum $P_\kappa(l)$ [see (6.53)], it is by no means clear that it is an optimal measure for the projected power spectrum. Kaiser (1998) considered the case of a square-shaped data field and employed the Fourier-transformed Kaiser & Squires inversion formula, eq. (5.3, page 32). The Fourier transform of the shear is then replaced by a sum over galaxy ellipticities ε_i, so that $\hat\kappa_{\rm eff}(\vec l)$ is expressed directly in terms of the ε_i. The square $|\hat\kappa_{\rm eff}(\vec l)|^2$ yields an estimate for the power spectrum which allows a simple determination of the noise coming from the intrinsic ellipticity distribution. As Kaiser (1998) pointed out that, while this noise is very small for angular scales much smaller than the size of the data field, the sampling variance is much larger, so that different sampling strategies should be explored. For example, he suggests to use a sparse sampling strategy. Seljak (1998) developed an estimator for the power spectrum which achieves minimum variance in the case of a Gaussian field. Since the power spectrum $P_\kappa(l)$ deviates significantly from its linear prediction on angular scales below one degree, one expects that the field attains significant non-Gaussian features on smaller angular scales, so that this estimator does no longer need to have minimum variance.

6.7. *Higher-Order Statistical Measures*

6.7.1. *The Skewness*

As the density perturbation field δ grows with time, it develops non-Gaussian features. In particular, δ is bounded by -1 from below and unbounded from above, and therefore the distribution of δ is progressively skewed while evolution proceeds. The same then applies to quantities like the effective convergence $\kappa_{\rm eff}$ derived from δ (cf. Jain & Seljak 1997; Bernardeau et al. 1997; Schneider et al. 1998a). Skewness of the effective convergence can be quantified by means of the three-point correlator of $\kappa_{\rm eff}$. In order to compute that, we use expression (6.18), Fourier transform it, and also express the density contrast δ in terms of its Fourier transform. Additionally, we employ the same approximation used in deriving Limber's equation in Fourier space, namely that correlations of the density contrast *along* the line-of-sight are negligibly small. After carrying out this lengthy but straightforward procedure, the three-point correlator of the Fourier transform of $\kappa_{\rm eff}$ reads (suppressing the subscript 'eff' for brevity)

$$\begin{aligned}
&\langle\hat\kappa(\vec l_1)\hat\kappa(\vec l_2)\hat\kappa(\vec l_3)\rangle \qquad (6.59)\\
&= \frac{27H_0^6\Omega_0^3}{8c^6}\int_0^{w_{\rm H}}{\rm d}w\,\frac{\bar W^3(w)}{a^3(w)f_K^3(w)}\int_{-\infty}^{\infty}\frac{{\rm d}k_3}{2\pi}\exp({\rm i}k_3w)\\
&\times\left\langle\hat\delta\left(\frac{\vec l_1}{f_K(w)},k_3\right)\hat\delta\left(\frac{\vec l_2}{f_K(w)},0\right)\hat\delta\left(\frac{\vec l_3}{f_K(w)},0\right)\right\rangle .
\end{aligned}$$

Hats on symbols denote Fourier transforms. Note the fairly close analogy between (6.60) and (6.22): The three-point correlator of $\hat\kappa$ is a distance-weighted integral over the three-point correlator of the Fourier-transformed density contrast $\hat\delta$. The fact that the three-component k_3 of the wave vector $\vec k$ appears only in the first factor $\hat\delta$ reflects the approximation mentioned above, i.e. that correlations of δ along the line-of-sight are negligible.

Suppose now that the density contrast δ is expanded in a perturbation series, $\delta = \sum\delta^{(i)}$ such that $\delta^{(i)} = O([\delta^{(1)}]^i)$, and truncated after the second order. The three-point correlator of $\hat\delta^{(1)}$ vanishes because δ remains Gaussian to first perturbation order. The lowest-order, non-vanishing three-point correlator of δ can therefore symbolically be written $\langle\hat\delta^{(1)}\hat\delta^{(1)}\hat\delta^{(2)}\rangle$, plus two permutations of that expression. The second-order density perturbation is related to the first order through (Fry 1984; Goroff et al. 1986; Bouchet et al. 1992)

$$\hat\delta^{(2)}(\vec k,w) = D_+^2(w)\int\frac{{\rm d}^3k'}{(2\pi)^3}\,\hat\delta_0^{(1)}(\vec k')\hat\delta_0^{(1)}(\vec k-\vec k')\,F(\vec k',\vec k-\vec k')\;, \qquad (6.60)$$

where $\delta_0^{(1)}$ is the first-order density perturbation linearly extrapolated to the present epoch, and $D_+(w)$ is the linear growth factor, $D_+(w) = a(w)\,g[a(w)]$ with $g(a)$ defined in eq. (2.52) on page 9. The function $F(\vec x,\vec y)$ is given by

$$F(\vec x,\vec y) = \frac{5}{7}+\frac{1}{2}\left(\frac{1}{|\vec x|^2}+\frac{1}{|\vec y|^2}\right)\vec x\cdot\vec y+\frac{2}{7}\frac{(\vec x\cdot\vec y)^2}{|\vec x|^2|\vec y|^2}\;. \qquad (6.61)$$

Relation (6.60) implies that the lowest-order three-point correlator $\langle\hat\delta^{(1)}\hat\delta^{(1)}\hat\delta^{(2)}\rangle$ involves four-point correlators of $\hat\delta^{(1)}$. For Gaussian fields like $\delta^{(1)}$, four-point correlators can be decomposed into sums of products of two-point correlators, which can be expressed in terms of the linearly extrapolated density power spectrum $P_\delta^{(0)}$. This leads to

$$\begin{aligned}
&\langle\hat\delta^{(1)}(\vec k_1)\hat\delta^{(1)}(\vec k_2)\hat\delta^{(2)}(\vec k_3)\rangle \qquad (6.62)\\
&= 2\,(2\pi)^3\,D_+^4(w)\,P_\delta^{(0)}(k_1)P_\delta^{(0)}(k_2)\\
&\times\delta_{\rm D}(\vec k_1+\vec k_2+\vec k_3)\,F(\vec k_1,\vec k_2)\;.
\end{aligned}$$

The complete lowest-order three-point correlator of $\hat\delta$ is a sum of three terms, namely the left-hand side of (6.63) and two permutations thereof. Each permutation yields the same result, so that the complete correlator is three times the right-hand side of (6.63). We can now work our way back, inserting the three-point density correlator into eq. (6.60) and Fourier-transforming the result with respect to $\vec l_{1,2,3}$. The three-point correlator of the effective convergence so obtained can then in a final step be used to compute the third moment of the aperture mass. The result is (Schneider et al. 1998a)

$$\begin{aligned}
\langle M_{\rm ap}^3(\theta)\rangle \;&=\; \frac{81H_0^6\Omega_0^3}{8\pi c^6}\int_0^{w_{\rm H}}{\rm d}w\,\frac{\bar W^3(w)D_+^4(w)}{a^3(w)f_K(w)}\\
&\times\;\int{\rm d}^2l_1\,P_\delta^{(0)}\left(\frac{l_1}{f_K(w)}\right)J^2(l_1\theta)\\
&\times\;\int{\rm d}^2l_2\,P_\delta^{(0)}\left(\frac{l_2}{f_K(w)}\right)J^2(l_2\theta)\\
&\times\;J^2(|\vec l_1+\vec l_2|\theta)\,F(\vec l_1,\vec l_2)\;, \qquad (6.63)
\end{aligned}$$

with the filter function $J(\eta)$ defined in eq. (6.50). Commonly, third-order moments are expressed in terms of the skewness,

$$\mathcal{S}(\theta) \equiv \frac{\langle M_{\rm ap}^3(\theta)\rangle}{\langle M_{\rm ap}^2(\theta)\rangle^2}\;, \qquad (6.64)$$

where $\langle M_{\rm ap}^2(\theta)\rangle$ is calculated with the linearly evolved power spectrum. As seen earlier in eq. (6.49), $\langle M_{\rm ap}^2\rangle$ scales with the amplitude of the power spectrum, while $\langle M_{\rm ap}^3\rangle$ scales with the square of it. In this approximation, the skewness $\mathcal{S}(\theta)$ is therefore independent of the normalisation of the power spectrum, removing that major uncertainty and leaving cosmological parameters as primary degrees of freedom. For instance, the skewness $\mathcal{S}(\theta)$ is expected to scale approximately with Ω_0^{-1}. Figure 24 shows three examples.

As expected, lower values of Ω_0 yield larger skewness, and the skewness is reduced when Ω_Λ is increased keeping Ω_0 fixed. Despite the sensitivity of $\mathcal{S}(\theta)$ to the cosmological parameters, it should be noted that the source redshift distribution [entering through $\bar W(w)$] needs to be known sufficiently well before attempts can be made at constraining cosmological parameters through measurements of the aperture-mass skewness. However, photometric redshift estimates are expected to produce sufficiently well-constrained redshift distributions in the near future (Connolly et al. 1995; Gwyn & Hartwick 1996; Hogg et al. 1998).

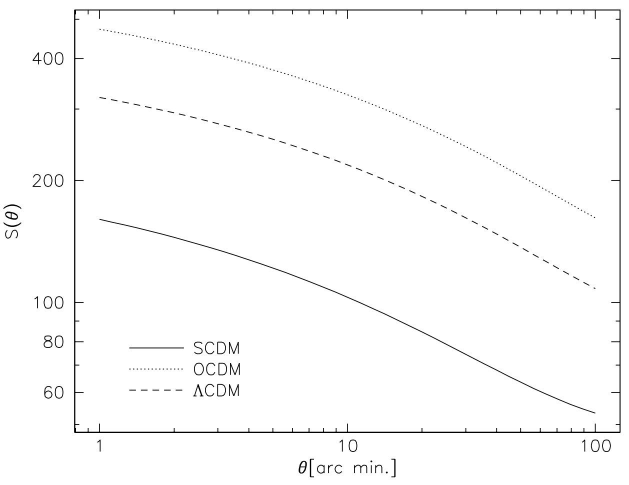

FIG. 24.—The skewness $\mathcal{S}(\theta)$ of the aperture mass $M_{\rm ap}(\theta)$ is shown as a function of aperture radius θ for three of the realisations of the cluster-normalised CDM cosmogony listed in Tab. 1: SCDM (solid curve); OCDM (dotted curve); and ΛCDM (dashed curve). The source redshift was assumed to be $z_s = 1$.

We have confined the discussion of the skewness to the aperture mass since $M_{\rm ap}$ is a scalar measure of the cosmic shear which can directly be expressed in terms of the observed image ellipticities. One can of course also consider the skewness directly in terms of κ, since κ can be obtained from the observed image ellipticities through a mass reconstruction algorithm as described in Sect. 5.. Analytical and numerical results for this skewness have been presented in, e.g., Bernardeau et al. (1997), van Waerbeke et al. (1999b), Jain et al. (2000) and Reblinsky et al. (1999). We shall discuss some of their results in Sect. 6.9.1..

As pointed out by Bernardeau (1998), the fact that the source galaxies are clustered in three-dimensional space, and therefore also in redshift space, generates an additional contribution to the skewness. This effect is more important than the contributions by the approximations made in the light propagation equations; in fact, Bernardeau (1998) estimated that the skewness can change by $\sim 25\%$ due to source clustering. Whereas the expectation values of second-order statistics of cosmic shear is unaffected by this clustering, the dispersion of any estimator increases. Of course, if the redshifts of the source galaxies are known, these effects can be avoided by suitably defining estimators for the quantities under consideration.

In the regime of small angular scales, where the relevant density contrast is highly nonlinear, different approximations apply for calculating higher-order statistical quantities. One of them is based on the so-called stable-clustering *ansatz*, which predicts a scaling relation for the n-point correlation function of the density contrast (Peebles 1980). Based on this assumption, and variants thereof, higher-order moments of cosmic-shear measures can be derived (e.g., Hui 1999a, Munshi & Coles 2000, Munshi & Jain 1999a), as well as approximations to the probability distribution for $\kappa_{\rm eff}$ itself and filtered (smoothed) versions thereof (Valageas 2000, Munshi & Jain 1999b, Valageas 1999). The resulting expressions, when compared to numerical simulations of light propagation through large-scale structures, are surprisingly accurate.

6.7.2. *Number density of (dark) halos*

In Sect. 5.3.1., we discussed the possibility to detect mass concentrations by their weak lensing effects on background galaxies by means of the aperture mass. The number density of mass concentrations that can be detected at a given threshold of $M_{\rm ap}$ depends on the cosmological model. Fixing the normalisation of the power spectrum so that the the local abundance of massive clusters is reproduced, the evolution of the density field proceeds differently in different cosmologies, and so the abundances will differ at redshifts $z \sim 0.3$ where the aperture-mass method is most sensitive.

The number density of halos above a given threshold of $M_{\rm ap}(\theta)$ can be estimated analytically, using two ingredients. First, the spatial number density of halos at redshift z with mass M can be described by the Press-Schechter theory (Press & Schechter 1974), which numerical simulations (Lacey & Cole 1993, Lacey & Cole 1994) have shown to be a fairly accurate approximation. Second, in a series of very large N-body simulations, Navarro et al. (1996a, 1997) found that dark matter halos have a universal density profile which can be described by two parameters, the halo mass and a characteristic scale length, which depends on the cosmological model and the redshift. Combining these two results from cosmology, Kruse & Schneider (1999b) calculated the number density of halos exceeding $M_{\rm ap}$. Using the signal-to-noise estimate eq. (6.58), a threshold value of $M_{\rm ap}$ can be directly translated into a signal-to-noise threshold S_c. For an assumed number density of $n = 30\,{\rm arcmin}^{-2}$ and an ellipticity dispersion $\sigma_\varepsilon = 0.2$, one finds $S_c \approx (\theta/1{\rm arcmin})(M_{\rm ap}(\theta)/0.016)$.

For the redshift distribution (2.69, page 12) with $\beta = 3/2$ and $z_0 = 1$, the number density of halos with $S_c \geq 5$ exceeds 10 per square degree for cluster-normalised cosmologies, across angular scales $1' \lesssim \theta \lesssim 10'$, and these halos have a broad redshift distribution which peaks at $z_d \sim 0.3$. This implies that a wide-field imaging survey should be able to detect a statistically interesting sample of medium redshift halos, thus allowing the definition of a *mass-selected* sample of halos. Such a sample will be of utmost interest for cosmology, since the halo abundance is considered to be one of the most sensitive cosmological probes (e.g., Eke et al. 1996, Bahcall & Fan 1998). Current attempts to apply this tool are hampered by the fact that halos are selected either by the X-ray properties or by their galaxy content. These properties are much more difficult to predict than the dark matter distribution of halos which can directly be determined from cosmological N-body simulations. Thus, these mass-selected halos will provide a much closer link to cosmological predictions than currently possible. Kruse & Schneider (1999b) estimated that an imaging survey of several square degrees will allow one to distinguish between the cosmological models given in Table 1, owing to the different number density of halos that they predict. Using the aperture-mass statistics, Erben et al. (2000) recently detected a highly significant matter concentration on two independent wide-field images centred on the galaxy cluster A 1942. This matter concentration $7'$ South of A 1942 is not associated with an overdensity of bright foreground galaxies, which sets strong lower limits on the mass-to-light ratio of this putative cluster.

6.8. *Cosmic Shear and Biasing*

Up to now, we have only considered the mass properties of the large-scale structure and tried to measure them with weak lensing techniques. An interesting question arises when the luminous constituents of the Universe are taken into account. Most importantly, the galaxies are supposed to be strongly tied to the distribution of dark matter. In fact, this assumption underlies all attempts to determine the power spectrum of cosmic density fluctuations from the observed distribution of galaxies. The relation between the galaxy and dark-matter distributions is parameterised by the so-called biasing factor b (Kaiser 1984), which is defined such that the relative fluctuations in the spatial number density of galaxies are b times the relative density fluctuations δ,

$$\frac{n(\vec{x})-\langle n\rangle}{\langle n\rangle}=b\,\delta(\vec{x})\;, \tag{6.65}$$

where $\langle n\rangle$ denotes the mean spatial number density of galaxies at the given redshift. The bias factor b is not really a single number, but generally depends on redshift, on the spatial scale, and on the galaxy type (see, e.g., Efstathiou 1996, Peacock 1997, Kauffmann et al. 1997, Coles et al. 1998). Typical values for the bias factor are assumed to be $b\sim 1-2$ at the current epoch, but can increase towards higher redshifts. The clustering properties of UV dropout galaxies (Steidel et al. 1998) indicate that b can be as large as 5 at redshifts $z\sim 3$, depending on the cosmology.

The projected surface mass density $\kappa_{\rm eff}(\vec{\theta})$ should therefore be correlated with the number density of (foreground) galaxies in that direction. Let $G_{\rm G}(w)$ be the distribution function of a suitably chosen population of galaxies in comoving distance (which can be readily converted to a redshift probability distribution). Then, assuming that b is independent of scale and redshift, the number density of the galaxies is

$$n_{\rm G}(\vec{\theta})=\langle n_{\rm G}\rangle\left[1+b\int{\rm d}w\,G_{\rm G}(w)\,\delta(f_K(w)\vec{\theta},w)\right]\;, \tag{6.66}$$

where $\langle n_{\rm G}\rangle$ is the mean number density of the galaxy population. The distribution function $G_{\rm G}(w)$ depends on the selection of galaxies. For example, for a flux-limited sample it may be of the form (2.69). Narrower distribution functions can be achieved by selecting galaxies in multi-colour space using photometric redshift techniques. The correlation function between $n_{\rm G}(\vec{\theta})$ and $\kappa_{\rm eff}(\vec{\theta})$ can directly be obtained from eq. (2.83) by identifying $q_1(w)=3H_0^2\Omega_0\bar{W}(w)f_K(w)/[2c^2a(w)]$ [see eq. (6.18)], and $q_2(w)=\langle n_{\rm G}\rangle bG_{\rm G}(w)$. It reads

$$\begin{aligned}\xi_{{\rm G}\kappa}(\theta) &\equiv \langle n_{\rm G}\kappa_{\rm eff}\rangle(\theta)\\ &= \frac{3H_0^2\Omega_0}{2c^2}\,b\langle n_{\rm G}\rangle\int{\rm d}w\,\frac{\bar{W}(w)f_K(w)}{a(w)}G_{\rm G}(w)\\ &\times \int\frac{{\rm d}k\,k}{2\pi}\,P_\delta(k,w)\,{\rm J}_0(f_K(w)\theta k)\;.\end{aligned} \tag{6.67}$$

Similar equations were derived by, e.g., Kaiser (1992), Bartelmann (1995b), Dolag & Bartelmann (1997), Sanz et al. (1997).

One way to study the correlation between foreground galaxies and the projected density field consists in correlating the aperture mass $M_{\rm ap}(\theta)$ with a similarly filtered galaxy number density, defined as

$$\mathcal{N}(\theta)=\int{\rm d}^2\vartheta\,U(|\vec{\vartheta}|)\,n_{\rm G}(\vec{\vartheta})\;, \tag{6.68}$$

with the same filter function U as in $M_{\rm ap}$. The correlation between $M_{\rm ap}(\theta)$ and $\mathcal{N}(\theta)$ then becomes

$$\begin{aligned}\xi(\theta) &\equiv \langle M_{\rm ap}(\theta)\mathcal{N}(\theta)\rangle\\ &= \int{\rm d}^2\vartheta\,U(|\vec{\vartheta}|)\int{\rm d}^2\vartheta'\,U(|\vec{\vartheta}'|)\xi_{{\rm G}\kappa}(|\vec{\vartheta}-\vec{\vartheta}'|) \quad (6.69)\\ &= 3\pi\left(\frac{H_0}{c}\right)^2\Omega_0 b\langle n_{\rm G}\rangle\int{\rm d}w\,\frac{\bar{W}(w)G_{\rm G}(w)}{a(w)f_K(w)}\\ &\times \int{\rm d}l\,l\,P_\delta\left(\frac{l}{f_K(w)},w\right)\,J^2(l\theta)\;,\end{aligned}$$

where we used eq. (2.83) for the correlation function $\xi_{{\rm G}\kappa}$ in the final step. The filter function J is defined in eq. (6.50). Note that this correlation function filters out the power spectrum P_δ at redshifts where the foreground galaxies are situated. Thus, by selecting galaxy populations with narrow redshift distribution, one can study the cosmological evolution of the power spectrum or, more accurately, the product of the power spectrum and the bias factor.

The convenient property of this correlation function is that one can define an unbiased estimator for ξ in terms of observables. If $N_{\rm b}$ galaxies are found in an aperture of radius θ at positions $\vec{\vartheta}_i$ with tangential ellipticity $\varepsilon_{{\rm t}i}$, and $N_{\rm f}$ foreground galaxies at positions $\vec{\varphi}_i$, then

$$\tilde{\xi}(\theta)=\frac{\pi\theta^2}{N_{\rm b}}\sum_{i=1}^{N_{\rm b}}Q(|\vec{\vartheta}_i|)\,\varepsilon_{{\rm t}i}\sum_{k=1}^{N_{\rm f}}U(|\varphi_k|) \tag{6.70}$$

is an unbiased estimator for $\xi(\theta)$. Schneider (1998) calculated the noise properties of this estimator, concentrating on an Einstein-de Sitter model and a linearly evolving power spectrum which can locally be approximated by a power law in k. A more general and thorough treatment is given in van Waerbeke (1998), where various cosmological models and the non-linear power spectrum are considered. van Waerbeke (1998) assumed a broad redshift distribution for the background galaxies, but a relatively narrow redshift distribution for the foreground galaxies, with $\delta z_{\rm d}/z_{\rm d}\sim 0.3$. For an open model with $\Omega_0=0.3$, $\xi(\theta)$ declines much faster with θ than for flat models, implying that open models have relatively more power on small scales at intermediate redshift. This is a consequence of the behaviour of the growth factor $D_+(w)$; see Fig. 6 on page 9. For foreground redshifts $z_{\rm d}\gtrsim 0.2$, the signal-to-noise ratio of the estimator (6.70) for a single aperture is roughly constant for $\theta\gtrsim 5'$, and relatively independent of the exact value of $z_{\rm d}$ over a broad redshift interval, with a characteristic value of ~ 0.4.

van Waerbeke (1998) also considered the ratio

$$R\equiv\frac{\xi(\theta)}{\langle\mathcal{N}^2(\theta)\rangle} \tag{6.71}$$

and found that it is nearly independent of θ. This result was shown in Schneider (1998) to hold for linearly evolving power spectra with power-law shape, but surprisingly it also holds for the fully non-linear power spectrum. Indeed, varying θ between $1'$ and $100'$, R varies by less than 2% for the models considered in van Waerbeke (1998). This is an extremely important result, in that any observed variation of R with angular scale indicates a corresponding scale dependence of the bias factor b. A direct observation of this variation would provide valuable constraints on the models for the formation and evolution of galaxies.

We point out that the ratio R depends, in the linear regime, on the combination Ω_0/b, independent of the normalisation of the power spectrum. This is to be compared with the combination $\Omega_0^{0.6}/b$ determined by peculiar motions of galaxies (e.g., Strauss & Willick 1995 and references therein). Since these combinations of the two parameters differ, one might hope that they can be derived separately by combining them.

6.9. Numerical Approach to Cosmic Shear, Cosmological Parameter Estimates, and Observations

6.9.1. Cosmic Shear Predictions from Cosmological Simulations

So far, we have treated the lensing effect of the large-scale structure with analytic means. This was possible because of two assumptions. First, we considered only the lowest-order lensing effect, by employing the Born approximation and neglecting lens-lens coupling in going from eq. (6.10) to eq. (6.12). Second, we used the prescription for the non-linear power spectrum as given by Peacock & Dodds (1996), assuming that it is a sufficiently accurate approximation. Both of these approximations may become less accurate on small angular scales. Providing a two-point quantity, the analytic approximation of P_κ is applicable only for two-point statistical measures of cosmic shear. In addition, the error introduced with these approximations cannot be controlled, i.e., we cannot attach 'error bars' to the analytic results.

A practical way to avoid these approximations is to study the propagation of light in a model universe which is generated by cosmological structure-formation simulations. They typically provide the three-dimensional mass distribution at different redshifts in a cube whose side-length is much smaller than the Hubble radius. The mass distribution along a line-of-sight can be generated by combining adjacent cubes from a sequence of redshifts. The cubes at different redshifts should either be taken from different realisations of the initial conditions, or, if this requires too much computing time, they should be translated and rotated such as to avoid periodicity along the line-of-sight. The mass distribution in each cube can then be projected along the line-of-sight, yielding a surface mass density distribution at that redshift. Finally, by employing the multiple lens-plane equations, which are a discretisation of the propagation equation (6.10; Seitz et al. 1994), shear and magnification can be calculated along light rays within a cone whose size is determined by the side length of the numerical cube. This approach was followed by many authors (e.g., Jaroszynski et al. 1990, Jaroszynski 1991, Bartelmann & Schneider 1991, Blandford et al. 1991, Waxman & Miralda-Escudé 1995), but the rapid development of *N*-body simulations of the cosmological dark matter distribution render the more recent studies particularly useful (Wambsganss et al. 1998, van Waerbeke et al. 1999b, Jain et al. 2000).

As mentioned below eq. (6.30), the Jacobian matrix $\mathcal{A}$ is generally asymmetric when the propagation equation is not simplified to (6.12). Therefore, the degree of asymmetry of $\mathcal{A}$ provides one test for the accuracy of this approximation. Jain et al. (2000) found that the power spectrum of the asymmetric component is at least three orders of magnitude smaller than that of $\kappa_{\rm eff}$. For a second test, we have seen that the power spectrum of $\kappa_{\rm eff}$ should equal that of the shear in the frame of our approximations. This analytic prediction is very accurately satisfied in the numerical simulations.

Jain et al. (2000) and Reblinsky et al. (1999) found that analytic predictions of the dispersions of κ and $M_{\rm ap}$ respectively, are very accurate when compared to numerical results. For both cosmic shear measures, however, the analytic predictions of the skewness are not satisfactory on angular scales below $\sim 10'$. This discrepancy reflects the limited accuracy of the second-order Eulerian perturbation theory employed in deriving the analytic results. Hui (1999b) showed that the accuracy of the analytic predictions can be much increased by using a prescription for the highly-nonlinear three-point correlation function of the cosmic density contrast, as developed by Scoccimarro & Frieman (1999). On larger angular scales, the predictions from perturbation theory as described in Sect. 6.7.1. are accurate, as shown by Gaztanaga & Bernardeau (1998).

The signal-to-noise ratio of the dispersion of the cosmic shear, given explicitly for $M_{\rm ap}$ in eq. (6.58), is determined by the intrinsic ellipticity dispersion of galaxies and the sampling variance, expressed in terms of the curtosis. As shown by van Waerbeke et al. (1999b), Reblinsky et al. (1999), and White & Hu (1999), this curtosis is remarkably large. For instance, the curtosis of the aperture mass exceeds unity even on scales larger than $10'$, revealing non-Gaussianity on such large scales. Unfortunately, this large sampling variance implies not only that the area over which cosmic shear needs to be measured to achieve a given accuracy for its dispersion must be considerably larger than estimated for a Gaussian density field, but also that numerical estimates of cosmic shear quantities need to cover large solid angles for an accurate numerical determination of the relevant quantities.

From such numerical simulations, one can not only determine moments of the shear distribution, but also consider its full probability distribution. For example, the predictions for the number density of dark matter halos that can be detected through highly significant peaks of $M_{\rm ap}$ – see Sect. 6.7.2. – have been found by Reblinsky et al. (1999) to be fairly accurate, perhaps surprisingly so, given the assumptions entering the analytic results. Similarly, the extreme tail (say more than 5 standard deviations from the mean) of the probability distribution for $M_{\rm ap}$, calculated analytically in Kruse & Schneider (1999a), does agree with the numerical results; it decreases exponentially.

6.9.2. *Cosmological Parameter Estimates*

Since the cosmic shear described in this section directly probes the total matter content of the universe, i.e., without any reference to the relation between mass and luminosity, it provides an ideal tool to investigate the large-scale structure of the cosmological density field. Assuming the dominance of cold dark matter, the statistical properties of the cosmic mass distribution are determined by a few parameters, the most important of which are Ω_0, Ω_Λ, the shape parameter of the power spectrum, Γ, and the normalisation of the power spectrum expressed in terms of σ_8. For each set of these parameters, the corresponding cosmic shear signals can be predicted, and a comparison with observations then constrains the cosmological parameters.

Furthermore, since weak lensing probes the shape of the projected power spectrum, modifications of the CDM power spectrum by a contribution from hot dark matter (such as massive neutrinos) may be measurable; e.g., Cooray (1999a) estimated that a deep weak-lensing survey of 100 square degrees may yield a lower limit on the neutrino mass of 3.5 eV.

Several approaches to this parameter estimation have been discussed in the literature. For example, van Waerbeke et al. (1999b) used numerical simulations to generate synthetic cosmic shear data, fixing the normalisation of the density fluctuations to $\sigma_8\,\Omega_0 = 0.6$, which is essentially the normalisation by cluster abundance. A moderately wide and deep weak-lensing survey, covering 25 square degrees and reaching a number density of 30 galaxies per arcmin^2 with characteristic redshift $z_{\rm s} \sim 1$, will enable the distinction between an Einstein-de Sitter model and an open universe with $\Omega_0 = 0.3$ at the 6-σ level, though each of these models is degenerate in the Ω_0 vs. Ω_Λ plane. For this conclusion, only the skewness of the reconstructed effective surface mass density or the aperture mass was used. Kruse & Schneider (1999a) instead considered the highly non-Gaussian tail of the aperture mass statistics to constrain cosmological parameters, whereas Kruse & Schneider (1999b) considered the abundance of highly significant peaks of $M_{\rm ap}$ as a probe of the cosmological models. The peak statistics of reconstructed surface density maps (Jain & van Waerbeke 2000) also provides a valuable means to distinguish between various cosmological models.

Future work will also involve additional information on the redshifts of the background galaxies. Hu (1999) pointed out that splitting up the galaxy sample into several redshift bins substantially increases the ability to constrain cosmological parameters. He considered the power spectrum of the projected density and found that the accuracy of the corresponding cosmological parameters improves by a factor of ~ 7 for Ω_Λ, and by a factor of ~ 3 for Ω_0, estimated for a median redshift of unity.

All of the quoted work concentrated mainly on one particular measure of cosmic shear. One goal of future theoretical investigations will certainly be the construction of a method which combines the various measures into a 'global' statistics, designed to minimise the volume of parameter space allowed by the data of future observational weak lensing surveys. Future, larger-scale numerical simulations will guide the search for such a statistics and allow one to make accurate predictions.

In addition to a pure cosmic shear investigation, cosmic shear constraints can be used in conjunction with other measures of cosmological parameters. One impressive example has been given by Hu & Tegmark (1999), who showed that even a relatively small weak lensing survey could dramatically improve the accuracy of cosmological parameters measured by future Cosmic Microwave Background missions.

6.9.3. *Observations*

One of the first attempts to measure cosmic shear was reported in Mould et al. (1994), where the mean shear was investigated across a field of $9\overset{'}{.}6 \times 9\overset{'}{.}6$, observed with the Hale 5-meter Telescope. The image is very deep and has good quality (i.e., a seeing of $0\overset{''}{.}87$ FWHM). It is the same data as used by Brainerd et al. (1996) for the first detection of galaxy-galaxy lensing (see Sect. 8.). The mean ellipticity of the 4363 galaxies within a circle of $4\overset{'}{.}8$ radius with magnitudes $23 \le r \le 26$ was found to be $(0.5 \pm 0.5)\%$. A later, less conservative reanalysis of these data by Villumsen (unpublished), where an attempt was made to account for the seeing effects, yielded a 3-σ detection of a non-vanishing mean ellipticity.

Following the suggestion that the observed large-angle QSO-galaxy associations are due to weak lensing by the large-scale structure in which the foreground galaxies are embedded (see Sect. 7.), Fort et al. (1996) searched for shear around five luminous radio quasars. In one of the fields, the number density of stars was so high that no reasonable shear measurement on faint background galaxies could be performed.[14] In the remaining four QSO fields, they found a shear signal on a scale of $\sim 1'$ for three of the QSOs (those which were observed with SUSI, which has a field-of-view of $\sim 2\overset{'}{.}2$), and on a somewhat larger angular scale for the fourth QSO. Taken at face value, these observations support the suggestion of magnification bias caused by the large-scale structure. A reanalysis of the three SUSI fields by Schneider et al. (1998b), considering the *rms* shear over the fields, produced a positive value for $\langle|\gamma|^2\rangle$ at the 99% significance level, as determined by numerous simulations randomising the orientation angles of the galaxy ellipticities. The amplitude of the *rms* shear, when corrected for the dilution by seeing, is of the same magnitude as expected from cluster-normalised models. However, if the magnification bias hypothesis is true, these three lines-of-sight are not randomly selected, and therefore this measurement is of no cosmological use.

Of course, one or a few narrow-angle fields cannot be useful for a measurement of cosmic shear, owing to cosmic variance. Therefore, a meaningful measurement of cosmic shear must either include many small fields, or must be obtained from a wide-field survey. Using the first strategy, several projects are under way: The Hubble Space Telescope has been carrying out so-called parallel surveys, where one or more of the instruments not used for primary observations are switched on to obtain data of a field located a few arc minutes away from the primary pointing. Over the past few years, a considerable database of such parallel data sets has accumulated. Two teams are currently analysing parallel data sets taken with WFPC2 and STIS, respectively (see Seitz et al. 1998a, Rhodes et al. 1999). In addition, a cosmic-shear survey is currently under way, in which randomly selected areas of the sky are mapped with the FORS instrument ($\sim 6\overset{'}{.}7 \times 6\overset{'}{.}7$) on the VLT. Some of these areas include the fields from the STIS parallel survey.

The alternative approach is to map big areas and measure the cosmic shear on a wide range of scales. The wide-field cameras currently being developed and installed are ideally suited for this purpose, and several groups are actively engaged in this work (see the proceedings of the Boston lens conference, July 1999).

Very recently, four groups have independently and almost simultaneously reported statistically significant detections of cosmic shear. In alphabetic order: Bacon et al. (2000) used 14 independent fields of size $8' \times 16'$ obtained with the WHT to measure the *rms* shear in squares of $8' \times 8'$. Kaiser et al. (2000) used six independent images taken with the UH8K camera on CFHT, each $30' \times 30'$ in size, to measure the cosmic shear on scales between $2'$ and $30'$. van Waerbeke et al. (2000) observed eight independent fields with the UH8K and UH12K ($30' \times 45'$) cameras at CFHT and measured the *rms* shear on scales below $3.5'$ since they avoided measurements in apertures crossing chip edges. Finally, Wittman et al. (2000) took three independent fields of size $43' \times 43'$ with the BTC at CTIO to measure the two-point correlation function of galaxy ellipticities on scales between $2'$ and $30'$. All four groups discuss their statistical and systematic uncertainties in detail and employ various tests to convincingly demonstrate the physical reality of the signal. In particular, they show that remaining systematics most probably contribute to the shear signal at a level below 1%, i.e. much less than the measured signal on scales $\lesssim 10'$–$15'$. The results of these groups are presented in Fig. 25. The yet unpublished result by Maoli et al. is not included. It was obtained from 45 images taken with the FORS1 instrument ($6.7' \times 6.7'$) on UT1 of VLT. Evidently, the results of the various groups are in excellent agreement despite the data being taken with different optical filters, different cameras, different telescopes, and reduced with different data analysis techniques. This provides additional evidence for the reality of the cosmic-shear signal. The significance of the results extends up to $5.5\,\sigma$, dependent of course on the total size of the fields used for the respective analyses. Since, except for the VLT data, the number of independent fields used for these studies is small, the error is entirely dominated by cosmic variance.

These impressive results prove the power of cosmic-shear measurements as a novel tool for probing the statistical properties of large-scale structures on small scales and at late times in the universe. In the near future, such measurements will become comparably important, and will provide complementary cosmological information to that obtained from CMB experiments.

There is nothing special about weak lensing being carried out predominantly in the optical wavelength regime, except that the optical sky is full of faint extended sources, whereas the radio sky is relatively empty. The FIRST radio survey covers at present about 4200 square degrees and contains 4×10^5 sources, i.e., the number density is smaller by about a factor ~ 1000 than in deep optical images. However, this radio survey covers a much larger solid angle than current or foreseeable *deep* optical surveys. As discussed in Refregier et al. (1998), this survey may yield a significant measurement of the two-point correlation function of image ellipticities on angular scales $\gtrsim 10'$. On smaller angular scales, sources with intrinsic double-lobe structure cannot be separated from individual independent sources. The Square Kilometer Array (van Haarlem & van der Hulst 1999) currently being discussed will yield such a tremendous increase in sensitivity for cm-wavelength radio astronomy that the radio sky will then be as crowded as the current optical sky. Finally, the recently commissioned Sloan telescope will map a quarter of the

[14] This field was subsequently used to demonstrate the superb image quality of the SUSI instrument on the ESO NTT.

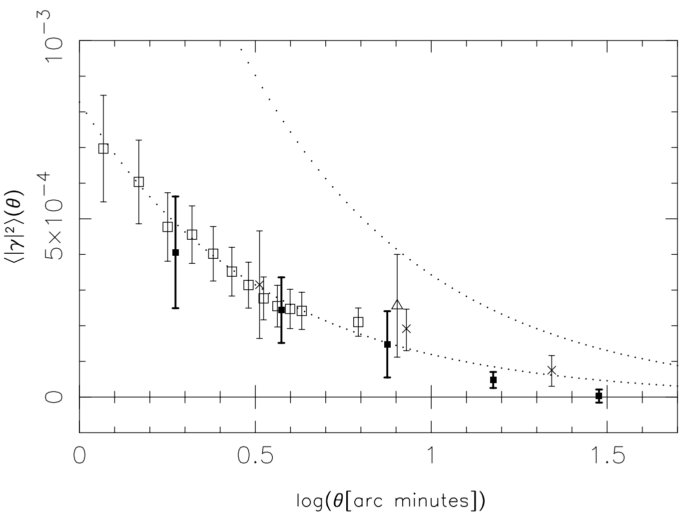

FIG. 25.—Compilation of the results of four different measurements of the cosmic-shear dispersion (with the two-point shear correlation function of Wittman et al. 2000 transformed into an equivalent dispersion for comparison). Open triangle: Bacon et al. (2000); filled squares: Kaiser et al. (2000); open squares: van Waerbeke et al. (2000); crosses: Wittman et al. (2000). The error bars include both statistical errors and cosmic variance. Points from the same group at different angular scales are not statistically independent. The dotted curves are predictions for a cluster-normalised ΛCDM model with effective source redshifts of $z_s = 1$ (lower curve) and $z_s = 2$ (upper curve), taken from Jain & Seljak (1997). (adapted from Kaiser et al. 2000)

sky in five colours. Although the imaging survey will be much shallower than current weak-lensing imaging, the huge area surveyed can compensate for the reduced galaxy number density and their smaller mean redshift Stebbins et al. (1996). Indeed, first weak-lensing results were already reported at the Boston lensing conference (July 1999) from commissioning data of the telescope (see also Fischer et al. 1999).

7. QSO MAGNIFICATION BIAS AND LARGE-SCALE STRUCTURE

7.1. *Introduction*

Magnification by gravitational lenses is a purely geometrical phenomenon. The solid angle spanned by the source is enlarged, or equivalently, gravitational focusing directs a larger fraction of the energy radiated by the source to the observer. Sources that would have been too faint without magnification can therefore be seen in a flux-limited sample. However, these sources are now distributed over a larger patch of the sky because the solid angle is stretched by the lens, so that the number density of the sources on the sky is reduced. The net effect on the number density depends on how many sources are added to the sample because they appear brighter. If the number density of sources increases steeply with decreasing flux, many more sources appear due to a given magnification, and the simultaneous dilution can be compensated or outweighed.

This magnification bias was described in Sect. 4.4.1. (page 25) and quantified in eq. (4.38). As introduced there, let $\mu(\vec\theta)$ denote the magnification into direction $\vec\theta$ on the sky, and $n_0(>S)$ the intrinsic counts of sources with observed flux exceeding S. In the limit of weak lensing, $\mu(\vec\theta) \gtrsim 1$, and the flux will not change by a large factor, so that it is sufficient to know the behaviour of $n_0(>S)$ in a small neighbourhood of S. Without loss of generality, we can assume the number-count function to be a power law in that neighbourhood, $n_0(>S) \propto S^{-\alpha}$. We can safely ignore any redshift dependence of the intrinsic source counts here because we aim at lensing effects of moderate-redshift mass distributions on high-redshift sources. Equation (4.43, page 26) then applies, which relates the cumulative source counts $n(>S,\vec\theta)$ observed in direction $\vec\theta$ to the intrinsic source counts,

$$n(>S,\vec\theta) = \mu^{\alpha-1}(\vec\theta)\, n_0(>S)\;. \tag{7.1}$$

Hence, if $\alpha > 1$, the observed number density of objects is increased by lensing, and reduced if $\alpha < 1$. This effect is called *magnification bias* or *magnification anti-bias* (e.g. Schneider et al. 1992).

The intrinsic number-count function of QSOs is well fit by a broken power law with a slope of $\alpha \sim 0.64$ for QSOs fainter than $\sim$ 19th blue magnitude, and a steeper slope of $\alpha \sim 2.52$ for brighter QSOs (Boyle et al. 1988; Hartwick & Schade 1990; Pei 1995). Faint QSOs are therefore anti-biased by lensing, and bright QSOs are biased. In the neighbourhood of gravitational lenses, the number density of bright QSOs is thus expected to be higher than average, in other words, more bright QSOs should be observed close to foreground lenses than expected without lensing. According to eq. (7.1), the overdensity factor is

$$q(\vec\theta) = \frac{n(>S,\vec\theta)}{n_0(>S)} = \mu^{\alpha-1}(\vec\theta)\;. \tag{7.2}$$

If the lenses are individual galaxies, the magnification $\mu(\vec\theta)$ drops rapidly with increasing distance from the lens. The natural scale for the angular separation is the Einstein radius, which is of order an arc second for galaxies. Therefore, individual galaxies are expected to increase the number density of bright QSOs only in a region of radius a few arc seconds around them.

Fugmann (1990) reported an observation which apparently contradicts this expectation. He correlated bright, radio-loud QSOs at moderate and high redshifts with galaxies from the Lick catalogue (Seldner et al. 1977) and found that there is a significant overdensity of galaxies around the QSOs of some of his sub-samples. This is intriguing because the Lick catalogue contains the counts of galaxies brighter than $\sim$ 19th magnitude in square-shaped cells with $10'$ side length. Galaxies of $\lesssim$ 19th magnitude are typically at much lower redshifts than the QSOs, $z \lesssim 0.1-0.2$, so that the QSOs with redshifts $z \gtrsim 0.5-1$ are in the distant background of the galaxies, with the two samples separated by hundreds of megaparsecs. Physical correlations between the QSOs and the galaxies are clearly ruled out. Can the observed overdensity be expected from gravitational lensing? By construction, the angular resolution of the Lick catalogue is of order $10'$, exceeding the Einstein radii of individual galaxies by more than two orders of magnitude. The result that Lick galaxies are correlated with bright QSOs can thus neither be explained by physical correlations nor by gravitational lensing due to individual galaxies.

On the other hand, the angular scale of $\sim 10'$ is on the right order of magnitude for lensing by large-scale structures. The question therefore arises whether the magnification due to lensing by large-scale structures is sufficient to cause a magnification bias in flux-limited QSO samples which is large enough to explain the observed QSO-galaxy correlation. The idea is that QSOs are then expected to appear more abundantly behind matter overdensities. More galaxies are expected where the matter density is higher than on average, and so the galaxies would act as tracers for the dark material responsible for the lensing magnification. This could then cause foreground galaxies to be overdense around background QSOs. This exciting possibility clearly deserves detailed investigation.

Even earlier than Fugmann, Tyson (1986) had inferred that galaxies apparently underwent strong luminosity evolution from a detection of significant galaxy overdensities on scales of $30''$ around 42 QSOs with redshifts $1 \leq z \leq 1.5$, assuming that the excess galaxies were at the QSO redshifts. In the light of later observations and theoretical studies, he probably was the first to detect weak-lensing induced associations of distant sources with foreground galaxies.

7.2. *Expected Magnification Bias from Cosmological Density Perturbations*

To estimate the magnitude of the effect, we now calculate the angular cross-correlation function $\xi_{QG}(\phi)$ between background QSOs and foreground galaxies expected from weak lensing due to large-scale structures (Bartelmann 1995b; Dolag & Bartelmann 1997; Sanz et al. 1997). We employ a simple picture for the relation between the number density of galaxies and the density contrast of dark matter, the linear biasing scheme (e.g. Kaiser 1984; Bardeen et al. 1986; White et al. 1987). Within this picture, and assuming weak lensing, we shall immediately see that the desired correlation function ξ_{QG} is proportional to the cross-correlation function $\xi_{\mu\delta}$ between magnification μ and density contrast δ. The latter correlation can straightforwardly be computed with the techniques developed previously.

7.2.1. *QSO-Galaxy Correlation Function*

The angular cross-correlation function $\xi_{QG}(\phi)$ between galaxies and QSOs is defined by

$$\xi_{GQ}(\phi) = \frac{1}{\langle n_Q\rangle\langle n_G\rangle} \left\langle \left[n_Q(\vec\theta) - \langle n_Q\rangle\right] \left[n_G(\vec\theta+\vec\phi) - \langle n_G\rangle\right] \right\rangle \,, \tag{7.3}$$

where $\langle n_{Q,G}\rangle$ are the mean number densities of QSOs and galaxies averaged over the whole sky. Assuming isotropy, $\xi_{QG}(\phi)$ does not depend on the direction of the lag angle $\vec\phi$. All number densities depend on flux (or galaxy magnitude), but we leave out the corresponding arguments for brevity.

We saw in eq. (7.1) in the introduction that $n_Q(\vec\theta) = \mu^{\alpha-1}(\vec\theta)\,\langle n_Q\rangle$. Since the magnification expected from large-scale structures is small, $\mu = 1 + \delta\mu$ with $|\delta\mu| \ll 1$, we can expand $\mu^{\alpha-1} \approx 1 + (\alpha-1)\delta\mu$. Hence, we can approximate

$$\frac{n_Q(\vec\theta) - \langle n_Q\rangle}{\langle n_Q\rangle} \approx (\alpha-1)\,\delta\mu(\vec\theta) \,, \tag{7.4}$$

so that the relative fluctuation of the QSO number density is proportional to the magnification fluctuation, and the factor of proportionality quantifies the magnification bias. Again, for $\alpha = 1$, lensing has no effect on the number density.

The linear biasing model for the fluctuations in the galaxy density asserts that the relative fluctuations in the galaxy number counts are proportional to the density contrast δ,

$$\frac{n_G(\vec\theta) - \langle n_G\rangle}{\langle n_G\rangle} = b\,\bar\delta(\vec\theta) \,, \tag{7.5}$$

where $\bar\delta(\vec\theta)$ is the line-of-sight integrated density contrast, weighted by the galaxy redshift distribution, i.e. the w-integral in eq. (6.66), page 57. The proportionality factor b is the effective biasing factor appropriately averaged over the line-of-sight. Typical values for the biasing factor are assumed to be $b \gtrsim 1-2$. Both the relative fluctuations in the galaxy number density and the density contrast are bounded by -1 from below, so that the right-hand side should be replaced by $\max[b\bar\delta(\vec\theta), -1]$ in places where $\bar\delta(\vec\theta) < -b^{-1}$. For simplicity we use (7.5), keeping this limitation in mind.

Using eqs. (7.4) and (7.5), the QSO-galaxy cross-correlation function (7.3) becomes

$$\xi_{QG}(\phi) = (\alpha-1)\,b\,\langle\delta\mu(\vec\theta)\bar\delta(\vec\theta+\vec\phi)\rangle \,. \tag{7.6}$$

Hence, it is proportional to the cross-correlation function $\xi_{\mu\delta}$ between magnification and density contrast, and the proportionality factor is given by the steepness of the intrinsic QSO number counts and the bias factor (Bartelmann 1995b). As expected from the discussion of the magnification bias, the magnification bias is ineffective for $\alpha = 1$, and QSOs and galaxies are anti-correlated for $\alpha < 1$. Furthermore, if the number density of galaxies does not reflect the dark-matter fluctuations, b would vanish, and the correlation would disappear. In order to find the QSO-galaxy cross-correlation function, we therefore have to evaluate the angular cross-correlation function between magnification and density contrast.

7.2.2. *Magnification-Density Correlation Function*

We have seen in Sect. 6. that the magnification fluctuation is twice the effective convergence $\delta\mu(\vec\theta) = 2\kappa_{\rm eff}(\vec\theta)$ in the limit of weak lensing, see eq. (6.29, page 49). The latter is given by eq. (6.19, page 47), in which the average over the source-distance distribution has already been performed. Therefore, we can immediately write down the source-distance averaged magnification fluctuation as

$$\delta\bar\mu(\vec\theta) = \frac{3H_0^2\Omega_0}{c^2} \int_0^{w_H} dw\, \bar W_Q(w)\, f_K(w)\, \frac{\delta[f_K(w)\vec\theta, w]}{a(w)} \,. \tag{7.7}$$

Here, $\bar W_Q(w)$ is the modified QSO weight function

$$\bar W_Q(w) \equiv \int_w^{w_H} dw'\, G_Q(w')\, \frac{f_K(w'-w)}{f_K(w')} \,, \tag{7.8}$$

and $G_Q(w)$ is the normalised QSO distance distribution.

Both the average density contrast $\bar\delta$ and the average magnification fluctuation $\delta\bar\mu$ are weighted projections of the density fluctuations along the line-of-sight, which is assumed to be a homogeneous and isotropic random field. As in the derivation of the effective-convergence power spectrum in Sect. 6., we can once more employ Limber's equation in Fourier space to find the cross power spectrum $P_{\mu\delta}(l)$ for projected magnification and density contrast,

$$P_{\mu\delta}(l) = \frac{3H_0^2\Omega_0}{c^2} \int_0^{w_H} dw\, \frac{\bar W_Q(w)\, G_G(w)}{a(w)\, f_K(w)}\, P_\delta\left(\frac{l}{f_K(w)}\right) \,. \tag{7.9}$$

The cross-correlation function between magnification and density contrast is obtained from eq. (7.9) via Fourier transformation, which can be carried out and simplified to yield

$$\begin{aligned} \xi_{\mu\delta}(\phi) = \; & \frac{3H_0^2\Omega_0}{c^2} \int_0^{w_H} dw'\, f_K(w')\, \bar W_Q(w')\, G_G(w')\, a^{-1}(w') \\ \times \; & \int_0^\infty \frac{k dk}{2\pi}\, P_\delta(k, w')\, J_0[f_K(w') k\phi] \,. \end{aligned} \tag{7.10}$$

Quite obviously, there is a strong similarity between this equation and that for the magnification autocorrelation function, eq. (6.34, page 49). We note that eq. (7.10) automatically accounts for galaxy autocorrelations through the matter power spectrum $P_\delta(k)$.

We point out that the dependence of the QSO-galaxy correlation function scales like $\xi_{QG} \propto b\Omega_0 P_\delta(k_{\rm eff}$, where $k_{\rm eff}$ is the comoving wave number determined by the peak of the redshift distribution of the foreground galaxies and the angular separation ϕ considered. On the other hand, the auto-correlation function of the foreground galaxies behaves like $\xi_{GG} \propto b^2 P_\delta(k_{\rm eff})$, which implies that the ratio $\xi_{QG}/\xi_{GG} \propto \Omega/b$, the same dependence as already stressed earlier (Sect. 6.8., page 56). Again, this ratio is nearly independent of the normalisation of the power spectrum, and therefore a convenient measure of the ratio Ω/b (Benítez & Sanz 1999).

7.2.3. Distance Distributions and Weight Functions

The QSO and galaxy weight functions $G_{Q,G}(w)$ are normalised representations of their respective redshift distributions, where the redshift needs to be transformed to comoving distance w.

The redshift distribution of QSOs has frequently been measured and parameterised. Using the functional form and the parameters determined by Pei (1995), the modified QSO weight function $\bar{W}_Q(w)$ has the shape illustrated in the top panel of Fig. 26. It is necessary for our present purposes to be able to impose a lower redshift limit on the QSO sample. Since we want to study lensing-induced correlations between background QSOs and foreground galaxies, there must be a way to exclude QSOs physically associated with galaxy overdensities. This is observationally achieved by choosing a lower QSO redshift cut-off high enough to suppress any redshift overlap between the QSO and galaxy samples. This procedure must be reproduced in theoretical calculations of the QSO-galaxy cross-correlation function. This can be achieved by cutting off the observed redshift distribution G_Q below some redshift z_0, re-normalising it, and putting the result into eq. (7.8) to find $\bar{W}_Q$. The five curves shown in the top panel of Fig. 26 are for cut-off redshifts z_0 increasing from 0.0 (solid curve) to 2.0 in steps of 0.5. Obviously, the peak in $\bar{W}_Q$ shifts to larger w for increasing z_0.

Galaxy redshift distributions G_G can be obtained by extrapolating local galaxy samples to higher redshifts, adopting a constant comoving number density and a Schechter-type luminosity function. For the present purposes, this is a safe procedure because the galaxies to be correlated with the QSOs *should* be at sufficiently lower redshifts than the QSOs to avoid overlap between the samples. Thus the extrapolation from the local galaxy population is well justified. In order to convert galaxy luminosities to observed magnitudes, k-corrections need to be taken into account. Conveniently, the resulting weight functions should be parameterised by the brightness cut-off of the galaxy sample, in practice by the maximum galaxy magnitude m_0 (i.e. the minimum luminosity) required for a galaxy to enter the sample. The five representative curves for $G_G(w)$ in the lower panel of Fig. 26 are for m_0 increasing from 18.5 to 22.5 (solid curve) in steps of one magnitude. R-band magnitudes are assumed. For increasing cut-off magnitude m_0, i.e. for fainter galaxy samples, the distributions broaden, as expected. The correlation amplitude as a function of m_0 peaks if m_0 is chosen such that the median distance to the galaxies is roughly half the distance to the bulk of the QSO population considered.

7.2.4. Simplifications

It turns out in practice that the exact shapes of the QSO and galaxy weight functions $\bar{W}_Q(w)$ and $G_G(w)$ are of minor importance for the results. Allowing inaccuracies of order 10%, we

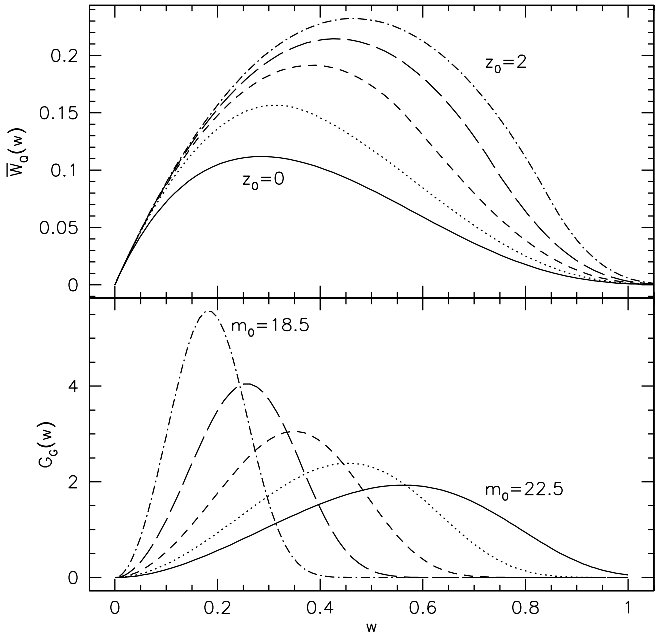

FIG. 26.—QSO and galaxy weight functions, $\bar{W}_Q(w)$ and $G_G(w)$, respectively. Top panel: $\bar{W}_Q(w)$ for five different choices of the lower cut-off redshift z_0 imposed on the QSO sample; z_0 increases from 0.0 (solid curve) to 2.0 in steps of 0.5. The peak in $\bar{W}_Q(w)$ shifts to larger distances for increasing z_0. Bottom panel: $G_G(w)$ for five different galaxy magnitude limits m_0, increasing from 18.5 to 22.5 (solid curve) in steps of one magnitude. The peak in the galaxy distance distribution shifts towards larger distances with increasing m_0, i.e. with decreasing brightness of the galaxy sample.

can replace the functions $G_{Q,G}(w)$ by delta distributions centred on typical QSO and galaxy distances w_Q and $w_G < w_Q$. Then, from eq. (7.8),

$$\bar{W}_Q(w) = \frac{f_K(w_Q - w)}{f_K(w_Q)} \, \mathrm{H}(w_Q - w) \,, \tag{7.11}$$

where $\mathrm{H}(x)$ is the Heaviside step function, and the line-of-sight integration in eq. (7.7) becomes trivial. It is obvious that matter fluctuations at redshifts higher than the QSO redshift do not contribute to the cross-correlation function $\xi_{\mu\delta}(\phi)$: Inserting (7.11) together with $G_G = \delta(w - w_G)$ into eq. (7.10), we find $\xi_{\mu\delta}(\phi) = 0$ if $w_G > w_Q$, as it should be.

The expression for the magnification-density cross-correlation function further simplifies if we specialise to a model universe with zero spatial curvature, $K = 0$, such that $f_K(w) = w$. Then,

$$\bar{W}_Q(w) = \left(1 - \frac{w}{w_Q}\right) \mathrm{H}(w_Q - w) \,, \tag{7.12}$$

and the cross-correlation function $\xi_{\mu\delta}(\phi)$ reduces to

$$\begin{aligned} \xi_{\mu\delta}(\phi) \quad &= \quad \frac{3H_0^2\Omega_0}{c^2} \, \frac{w_G}{a(w_G)} \left(1 - \frac{w_G}{w_Q}\right) \\ &\times \quad \int_0^\infty \frac{k\mathrm{d}k}{2\pi} \, P_\delta(k, w_G) \, \mathrm{J}_0(w_G k\phi) \end{aligned} \tag{7.13}$$

for $w_Q > w_G$, and $\xi_{\mu\delta}(\phi) = 0$ otherwise.

7.3. *Theoretical Expectations*

7.3.1. *Qualitative Behaviour*

Before we evaluate the magnification-density cross-correlation function fully numerically, we can gain some insight into its expected behaviour by inserting the CDM and HDM model spectra defined in eq. (6.37, page 50) into eq. (7.10) and expanding the result into a power series in ϕ (Bartelmann 1995b). As in the case of the magnification auto-correlation function before, the two model spectra produce qualitatively different results. To first order in ϕ, $\xi_{\mu\delta}(\phi)$ decreases linearly with increasing ϕ for CDM, while it is flat for HDM. The reason for this different appearance is the lack of small-scale power in HDM, and the abundance thereof in CDM. The two curves shown in Fig. 27 illustrate this for an Einstein-de Sitter universe with Hubble constant $h = 0.5$. The underlying density-perturbation power spectra were normalised by the local abundance of rich clusters, and linear density evolution was assumed.

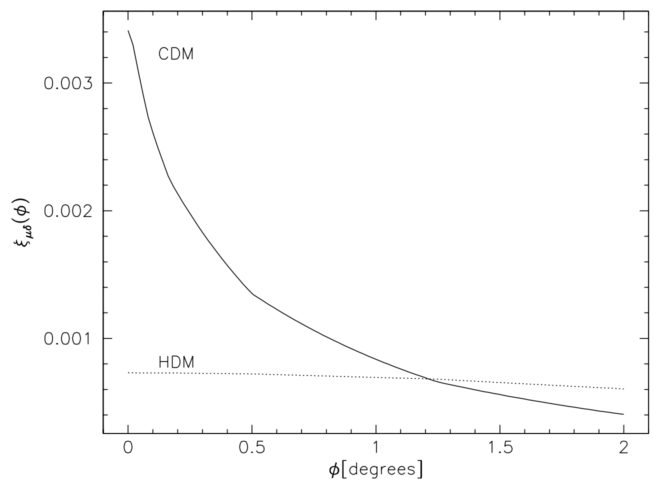

FIG. 27.—Cross-correlation functions between magnification and density contrast, $\xi_{\mu\delta}(\phi)$, are shown for an Einstein-de Sitter universe with $h = 0.5$, adopting CDM (solid curve) and HDM (dotted curve) density fluctuation spectra. Both spectra are normalised to the local cluster abundance, and linear density evolution is assumed. The lower cut-off redshift of the QSOs is $z_0 = 0.3$, the galaxy magnitude limit is $m_0 = 20.5$. In agreement with the expectation derived from the CDM and HDM model spectra (6.37, page 50), the CDM cross-correlation function decreases linearly with increasing ϕ for small ϕ, while it is flat to first order in ϕ for HDM. The small-scale matter fluctuations in CDM compared to HDM cause $\xi_{\mu\delta}(\phi)$ to increase more steeply as $\phi \to 0$.

The *linear* correlation amplitude, $\xi_{\mu\delta}(0)$, for CDM is of order 3×10^{-3}, and about a factor of five smaller for HDM. The magnification-density cross-correlation function for CDM drops to half its peak value within a few times 10 arc minutes. This, and the monotonic increase of $\xi_{\mu\delta}$ towards small ϕ, indicate that density perturbations on angular scales below $10'$ contribute predominantly to $\xi_{\mu\delta}$. At typical lens redshifts, such angular scales correspond to physical scales up to a few Mpc. Evidently therefore, the non-linear evolution of the density perturbations needs to be taken into account, and its effect is expected to be substantial.

7.3.2. *Results*

Figure 28 confirms this expectation; it shows magnification-density cross-correlation functions for the four cosmological models detailed in Tab. 1 on page 45. Two curves are shown for each model, one for linear and the other for non-linear density evolution. The two curves of each pair are easily distinguished because non-linear evolution increases the cross-correlation amplitude at small ϕ by about an order of magnitude above linear evolution, quite independent of the cosmological model. At the same time, the angular cross-correlation scale is reduced to a few arc minutes. At angular scales $\lesssim 30'$, the non-linear cross-correlation functions are above the linear results, falling below at larger scales. The correlation functions for the three cluster-normalised models (SCDM, OCDM and ΛCDM; see Tab. 1 on page 45) are very similar in shape and amplitude. The curve for the σCDM model lies above the other curves by a factor of about five, but for low-density universes, the influence of different power-spectrum normalisations are much less prominent.

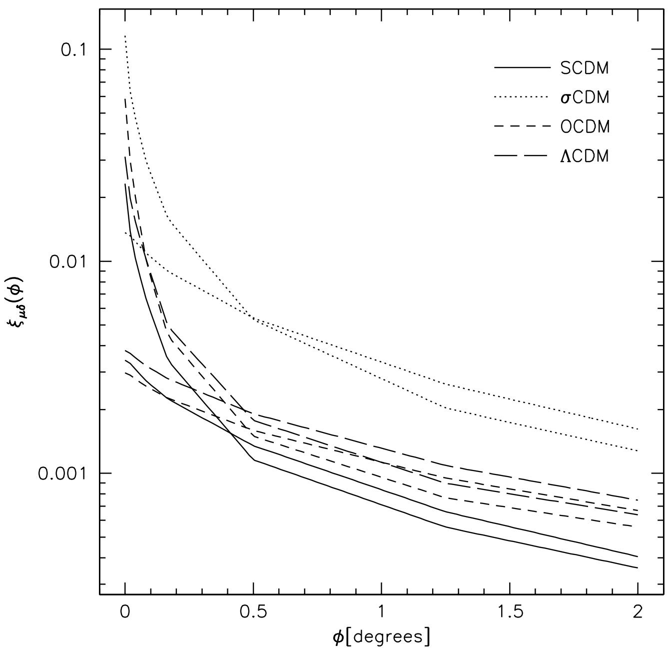

FIG. 28.—Angular magnification-density cross-correlation functions $\xi_{\mu\delta}(\phi)$ are shown for the four cosmological models specified in Table 1 on page 45. Two curves are shown for each cosmological model; those with the higher (lower) amplitude at $\phi = 0$ were calculated with the non-linearly (linearly) evolving density-perturbation power spectra, respectively. The models are: SCDM (solid curves), σCDM (dotted curves), OCDM (short-dashed curves), and ΛCDM (long-dashed curves). Obviously, non-linear evolution has a substantial effect. It increases the correlation amplitude by about an order of magnitude. The Einstein-de Sitter model normalised to $\sigma_8 = 1$ has a significantly larger cross-correlation amplitude than the cluster-normalised Einstein-de Sitter model. For the low-density models, the difference is much smaller. The curves for the cluster-normalised models are very similar, quite independent of cosmological parameters.

The main results to be extracted from Fig. 28 are that the amplitude of the magnification-density cross-correlation function, $\xi_{\mu\delta}(0)$, reaches approximately 5×10^{-2}, and that $\xi_{\mu\delta}$ drops by an order of magnitude within about $20'$. This behaviour is quite independent of the cosmological parameters if the density-fluctuation power spectrum is normalised by the local abundance of rich galaxy clusters. More detailed results can be found in Dolag & Bartelmann (1997) and Sanz et al. (1997).

7.3.3. Signal-to-Noise Estimate

The QSO-galaxy correlation function $\xi_{QG}(\phi)$ is larger than $\xi_{\mu\delta}(\phi)$ by the factor $(\alpha-1)b$. The value of the bias factor b is yet unclear, but it appears reasonable to assume that it is between 1 and 2. For optically selected QSOs, $\alpha \approx 2.5$, so that $(\alpha-1)b \approx 2-3$. Combining this with the correlation amplitude for CDM read off from Fig. 28, we can expect $\xi_{QG}(0) \lesssim 0.1$.

Given the meaning of $\xi_{QG}(\phi)$, the probability to find a foreground galaxy close to a background QSO is increased by a factor of $[1+\xi_{QG}(\phi)] \lesssim 1.1$ above random. In a small solid angle $d^2\omega$ around a randomly selected background QSO, we thus expect to find

$$N_G \approx [1+\xi_{QG}(0)]\,\langle n_G\rangle\, d^2\omega \equiv [1+\xi_{QG}(0)]\,\langle N_G\rangle \tag{7.14}$$

galaxies, where $\langle N_G\rangle$ is the average number of galaxies within a solid angle of $d^2\omega$. In a sample of N_Q fields around randomly selected QSOs, the signal-to-noise ratio for the detection of a galaxy overdensity is then

$$\frac{S}{N} \approx \frac{N_Q(N_G-\langle N_G\rangle)}{(N_Q\langle N_G\rangle)^{1/2}} = (N_Q\langle N_G\rangle)^{1/2}\,\xi_{QG}(0)\,. \tag{7.15}$$

Typical surface number densities of reasonably bright galaxies are of order $n_G \sim 10$ per square arc minute. Therefore, there should be of order $\langle N_G\rangle \sim 30$ galaxies within a randomly selected disk of one arc minute radius, in which the QSO-galaxy cross correlation is sufficiently strong. If we require a certain minimum signal-to-noise ratio such that $S/N \geq (S/N)_0$, the number of QSO fields to be observed in order to meet this criterion is

$$\begin{aligned} N_Q &\geq \left(\frac{S}{N}\right)_0^2 \xi_{QG}^{-2}(0)\,\langle N_G\rangle^{-1} \\ &= \left(\frac{S}{N}\right)_0^2 [(\alpha-1)\,b]^{-2}\,\xi_{\mu\delta}^{-2}(0)\,\langle N_G\rangle^{-1} \\ &= 20\left[\frac{(S/N)_0}{5}\right]^2 \left[\frac{(\alpha-1)\,b}{4}\right]^{-2} \left(\frac{\xi_{\mu\delta}(0)}{0.05}\right)^{-2} \left(\frac{30}{\langle N_G\rangle}\right)\,, \end{aligned} \tag{7.16}$$

where we have inserted typical numbers in the last step. This estimate demonstrates that gravitational lensing by non-linearly evolving large-scale structures in cluster-normalised CDM can produce correlations between background QSOs and foreground galaxies at the $5\,\sigma$ level on arc minute scales in samples of $\gtrsim 20$ QSOs. The angular scale of the correlations is expected to be of order 1 to 10 arc minutes. Equation (7.17) makes it explicit that more QSO fields need to be observed in order to establish the significance of the QSO-galaxy correlations if (i) the QSO number count function is shallow (α close to unity), and (ii) the galaxy bias factor b is small. In particular, no correlations are expected if $\alpha = 1$, because then the dilution of the sources and the increase in QSO number exactly cancel. Numerical simulations (Bartelmann 1995b) confirm the estimate (7.17).

Fugmann's (1990) observation was also tested in a numerical model universe based on the adhesion approximation to structure formation (Bartelmann & Schneider 1992). This model universe was populated with QSOs and galaxies, and QSO-galaxy correlations on angular scales on the order of $\sim 10'$ were investigated using Spearman's rank-order correlation test (Bartelmann & Schneider 1993a). Light propagation in the model universe was described with the multiple lens-plane approximation of gravitational lensing. In agreement with the analytical estimate presented above, it was found that lensing by large-scale structures can indeed account for the observed correlations between high-redshift QSOs and low-redshift galaxies, provided the QSO number-count function is steep. Lensing by individual galaxies was confirmed to be entirely negligible.

7.3.4. Multiple-Waveband Magnification Bias

The magnification bias quantified by the number-count slope α can be substantially increased if QSOs are selected in two or more mutually uncorrelated wave bands rather than one (Borgeest et al. 1991). To see why, suppose that optically bright *and* radio-loud QSOs were selected, and that their fluxes in the two wave bands are uncorrelated. Let $S_{1,2}$ be the flux thresholds in the optical and in the radio regimes, respectively, and $n_{1,2}$ the corresponding number densities of either optically bright or radio-loud QSOs on the sky. As in the introduction, we assume that $n_{1,2}$ can be written as power laws in $S_{1,2}$, with exponents $\alpha_{1,2}$.

In a small solid angle $d^2\omega$, the probability to find an optically bright *or* radio-loud QSO is then $p_i(S_i) = n_i(S_i)\,d^2\omega$, and the joint probability to find an optically bright *and* radio-loud QSO is the product of the individual probabilities, or

$$\begin{aligned} p(S_1,S_2) &= p_1(S_1)\,p_2(S_2) \\ &= [n_1(S_1)\,n_2(S_2)]\,d^2\omega \propto S_1^{-\alpha_1} S_2^{-\alpha_2}\,d^2\omega\,, \end{aligned} \tag{7.17}$$

provided there is no correlation between the fluxes $S_{1,2}$ so that the two probabilities are independent. Suppose now that lensing produces a magnification factor μ across $d^2\omega$. The joint probability is then changed to

$$p'(S_1,S_2) \propto \left(\frac{\mu}{S_1}\right)^{\alpha_1} \left(\frac{\mu}{S_2}\right)^{\alpha_2} \frac{d^2\omega}{\mu} = \mu^{\alpha_1+\alpha_2-1}\,p(S_1,S_2)\,. \tag{7.18}$$

Therefore, the magnification bias in the optically bright *and* radio-loud QSO sample is as efficient as if the number-count function had a slope of $\alpha = \alpha_1+\alpha_2$.

More generally, the effective number-count slope for the magnification bias in a QSO sample that is flux limited in m mutually uncorrelated wave bands is

$$\alpha = \sum_{i=1}^{m} \alpha_i\,, \tag{7.19}$$

where α_i are the number-count slopes in the individual wave bands. Then, the QSO-galaxy cross-correlation function is

$$\xi_{QG}(\phi) = \left(\sum_{i=1}^{m} \alpha_i - 1\right) b\,\xi_{\mu\delta}(\phi)\,, \tag{7.20}$$

and can therefore be noticeably larger than for a QSO sample which is flux limited in one wave band only.

7.4. Observational Results

After this theoretical investigation, we turn to observations of QSO-galaxy cross-correlations on large angular scales. The existence of QSO-galaxy correlations was tested and verified in several studies using some very different QSO- and galaxy samples.

Bartelmann & Schneider (1993b) repeated Fugmann's analysis with a well-defined sample of background QSOs, namely the optically identified QSOs from the 1-Jansky catalogue (Kühr et al. 1981; Stickel et al. 1993; Stickel & Kühr 1993). Optically identified QSOs with measured redshifts need to be bright enough for detection and spectroscopy, hence the chosen sample is implicitly also constrained by an optical flux limit. Optical and radio QSO fluxes are generally not strongly correlated, so that the sample is affected by a double-waveband magnification bias, which can further be strengthened by explicitly imposing an optical flux (or magnitude) limit.

Although detailed results differ from Fugmann's, the presence of the correlation is confirmed at the 98% confidence level for QSOs with redshifts ≥ 0.75 and brighter than 18th magnitude. The number of QSOs matching these criteria is 56. The correlation significance decreases both for lower- and higher-redshift QSO samples, and also for optically fainter ones. This is in accordance with an explanation in terms of a (double-waveband) magnification bias due to gravitational lensing. For low-redshift QSOs, lensing is not efficient enough to produce the correlations. For high-redshift QSOs, the most efficient lenses are at higher redshifts than the galaxies, so that the *observed* galaxies are uncorrelated with the structures which magnify the QSOs. Hence, the correlation is expected to disappear for increasing QSO redshifts. For an optically unconstrained QSO sample, the effective slope of the number-count function is smaller, reducing the strength of the magnification bias and therefore also the significance of the correlation.

With a similar correlation technique, correlations between the 1-Jansky QSO sample and IRAS galaxies (Bartelmann & Schneider 1994) and diffuse X–ray emission (Bartelmann et al. 1994; see also Cooray 1999b) were investigated, leading to qualitatively similar results. IRAS galaxies are correlated with optically bright, high-redshift $z \geq 1.5$ 1-Jansky sources at the 99.8% confidence level. The higher QSO redshift for which the correlation becomes significant can be understood if the IRAS galaxy sample is deeper than the Lick galaxy sample, so that the structures responsible for the lensing can be traced to higher redshift.

Bartsch et al. (1997) re-analysed the correlation between IRAS galaxies and 1-Jansky QSOs using a more advanced statistical technique which can be optimised to the correlation function expected from lensing by large-scale structures. In agreement with Bartelmann & Schneider (1994), they found significant correlations between the QSOs and the IRAS galaxies on angular scales of $\sim 5'$, but the correlation amplitude is higher than expected from large-scale structure lensing, assuming linear evolution of the density-perturbation power spectrum. Including non-linear evolution, however, the results by Bartsch et al. (1997) can well be reproduced (Dolag & Bartelmann 1997).

X–ray photons from the ROSAT *All-Sky Survey* (e.g. Voges 1992) are correlated with optically bright 1-Jansky sources both at low ($0.5 \leq z \leq 1.0$) and at high redshifts ($1.5 \leq z \leq 2.0$), but there is no significant correlation with QSOs in the intermediate redshift regime. A plausible explanation for this is that the correlation of X–ray photons with low-redshift 1-Jansky QSOs is due to hot gas which is physically associated with the QSOs, e.g. which resides in the host clusters of these QSOs. Increasing the source redshift, the flux from these clusters falls below the detection threshold of the *All-Sky Survey*, hence the correlation disappears. Upon further increasing the QSO redshift, lensing by large-scale structures becomes efficient, and the X–ray photons trace hot gas in the lenses.

Rodrigues-Williams & Hogan (1994) found a highly significant correlation between optically-selected, high-redshift QSOs and Zwicky clusters. Their cluster sample was fairly bright, which indicates that the clusters are in the foreground of the QSOs. This rules out that the clusters are physically associated with the QSOs and thus exert environmental effects on them which might lead to the observed association. Rodrigues-Williams & Hogan discussed lensing as the most probable reason for the correlations, although simple mass models for the clusters yield lower magnifications than required to explain the significance of the effect. Seitz & Schneider (1995b) repeated their analysis with the 1-Jansky sample of QSOs. They found agreement with Rodrigues-Williams & Hogan's result for intermediate-redshift ($z \sim 1$) QSOs, but failed to detect significant correlations for higher-redshift sources. In addition, a significant under-density of low-redshift QSOs close to Zwicky clusters was found, for which environmental effects like dust absorption are the most likely explanation. A variability-selected QSO sample was correlated with Zwicky clusters by Rodrigues-Williams & Hawkins (1995). They detected a significant correlation between QSOs with $0.4 \leq z \leq 2.2$ with foreground Zwicky clusters (with $\langle z \rangle \sim 0.15$) and interpreted it in terms of gravitational lensing. Again, the implied average QSO magnification is substantially larger than that inferred from simple lens models for clusters with velocity dispersions of $\sim 10^3\ \mathrm{km\,s^{-1}}$. Wu & Han (1995) searched for associations between distant 1-Jansky and 2-Jansky QSOs and foreground Abell clusters. They found no correlations with the 1-Jansky sources, and a marginally significant correlation with 2-Jansky sources. They argue that lensing by individual clusters is insufficient if cluster velocity dispersions are of order $10^3\ \mathrm{km\,s^{-1}}$, and that lensing by large-scale structures provides a viable explanation.

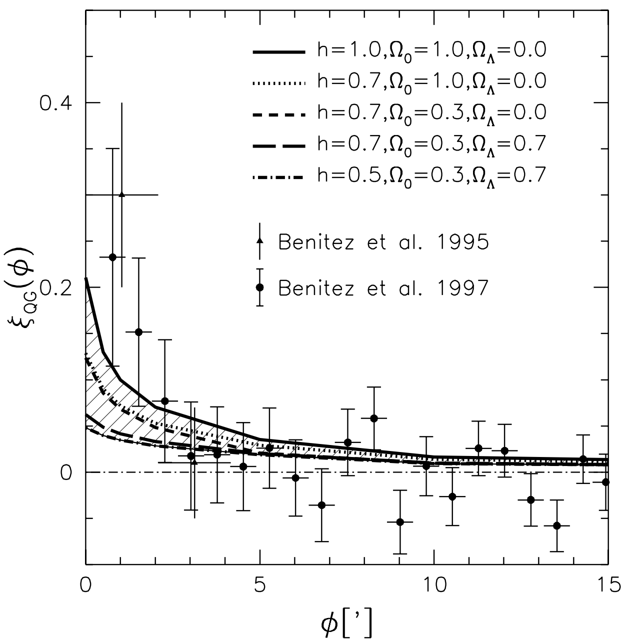

FIG. 29.—QSO-galaxy cross-correlation measurements are plotted together with theoretical cross-correlation functions $\xi_{\mathrm{QG}}(\phi)$ for various cosmological models as indicated by line type. The CDM density-perturbation power spectrum was cluster-normalised, and non-linear evolution was taken into account. The figure shows that the measurements fall above the theoretical predictions at small angular scales, $\phi \lesssim 2'$. This excess can be attributed to gravitational lensing by individual galaxy clusters (see the text for more detail). The theoretical curves depend on the Hubble constant h through the shape parameter $\Gamma = \Omega_0 h$, which determines the peak location of the power spectrum.

Benítez & Martínez-González (1995) found an excess of red galaxies from the APM catalog with moderate-redshift ($z \sim 1$) 1-Jansky QSOs on angular scales $< 5'$ at the 99.1% significance level. Their colour selection ensures that the galaxies are most likely at redshifts $0.2 \leq z \leq 0.4$, well in the foreground of the QSOs. The amplitude and angular scale of the excess is compatible with its originating from lensing by large-scale structures.

The measurements by Benítez & Martínez-González (1995) are plotted together with various theoretical QSO-galaxy cross-correlation functions in Fig. 29, which clearly shows that the QSO-galaxy cross-correlation measurements agree quite well with the cross-correlation functions $\xi_{QG}(\phi)$, but they fall above the range of theoretical predictions at small angular scales, $\phi \lesssim 2'$. This can be attributed to the magnification bias due to gravitational lensing by individual clusters. Being based on the weak-lensing approximation, our approach breaks down when the magnification becomes comparable to unity, $\mu \gtrsim 1.5$, say. This amount of magnification occurs for QSOs closer than $\sim$ 3 Einstein radii to cluster cores. Depending on cosmological parameters, QSO and galaxy redshifts, $\sim$ 3 Einstein radii correspond to $\sim 1' - 2'$. Hence, we *expect* the theoretical expectations from lensing by large-scale structures alone to fall below the observations on angular scales $\phi \lesssim 1' - 2'$.

Norman & Impey (1999) took wide-field R-band images centred on a subsample of 1-Jansky QSOs with redshifts between 1 and 2. They searched for an excess of galaxies in the magnitude range $19.5 < R < 21$ on angular scales of $\gtrsim 10'$ around these QSOs and found a correlation at the 99% significance level. The redshift distribution of the galaxies is likely to peak around $z \sim 0.2$. The angular cross-correlation function between the QSOs and the galaxies agrees well with the theoretical expectations, although the error bars are fairly large.

All these results indicate that there are correlations between background QSOs and foreground 'light', with light either in the optical, the infrared, or the (soft) X–ray wave bands. The angular scale of the correlations is compatible with that expected from lensing by large-scale structures, and the amplitude is either consistent with that explanation or somewhat larger. Wu & Fang (1996) discussed whether the autocorrelation of clusters modelled as singular isothermal spheres can produce sufficient magnification to explain this result. They found that this is not the case, and argued that large-scale structures must contribute substantially.

If lensing is indeed responsible for the correlations detected, other signatures of lensing should be found in the vicinity of distant QSOs. Indeed, Fort et al. (1996) searched for the shear induced by weak lensing in the fields of five luminous QSOs with $z \approx 1$ and found coherent shear signals in four of them (see also Schneider et al. 1998b). In addition, they detected galaxy groups in three of their fields. Earlier, Bonnet et al. (1993) had found evidence for coherent weak shear in the field of the potentially multiply-imaged QSO 2345+007, which was later identified with a distant cluster (Mellier et al. 1994; Fischer et al. 1994).

Bower & Smail (1997) searched for weak-lensing signals in fields around eight luminous radio sources at redshifts ~ 1. They confirmed the coherent shear detected earlier by Fort et al. (1996) around one of the sources (3C336 at $z = 0.927$), but failed to find signatures of weak lensing in the combined remaining seven fields.

A cautionary note was recently added to this discussion by Williams & Irwin (1998) and Norman & Williams (1999). Cross-correlating LBQS and 1-Jansky quasars with APM galaxies, they claimed significant galaxy overdensities around QSOs on angular scales of order one degree. As discussed above, lensing by currently favoured models of large-scale structures is not able to explain such large correlation scales. Thus, if these results hold up, they would provide evidence that there is a fundamental difficulty with the current models of large-scale structure formation.

7.5. Magnification bias of galaxies

The investigation of the angular correlation between QSOs and foreground galaxies was motivated by observational evidence of this effect, as described in the previous subsection. However, the magnification bias generates a similar correlation function between foreground galaxies and different classes of background sources, provided the latter have a slope of the cumulative sources counts different from unity. QSOs are particularly convenient due to their steep number counts and their high redshift. Moessner et al. (1998) and Moessner & Jain (1998) studied the angular correlation between two different populations of galaxies. If, for example, the two populations of galaxies were selected by their apparent magnitude, the fainter one will on average be more distant than the brighter one; therefore, matter traced by the brighter galaxies magnifies the fainter population of galaxies. Unfortunately, owing to the broad redshift distribution of galaxies at fixed apparent magnitude, there will be a significant overlap in redshift between these two populations. Since galaxies are auto-correlated, this intrinsic clustering contribution is likely to swamp any lensing-induced correlation. Note that, owing to the high-redshift cut used for the QSO samples considered in the previous subsection, this intrinsic correlation is of little or no importance there.

However, if the foreground and background populations can be better separated, the lensing effect may be stronger than the intrinsic correlation. For example, by using photometric redshift estimates, the two galaxy populations may be nicely separated in their redshift distribution. In that case, the cross-correlation function will take the form

$$\xi_{12}(\phi) = (\alpha_2 - 1) b_1 \xi_{\mu\delta}(\phi) + \xi'(\phi) , \tag{7.21}$$

where the first term is the contribution due to the magnification and has the same form as that derived for the QSO-galaxy correlation in the the previous subsection, and ξ' is the intrinsic cross-correlation function coming from imperfect redshift separation of the two galaxy populations. Note that α_2 is the number-count slope of the background galaxies, and b_1 the bias factor of the foreground population. If $\xi_{\mu\delta}$ and ξ' have different functional forms with respect to ϕ, these two contributions to the cross-correlation function may be separable.

From early commissioning data of the Sloan Digital Sky Survey, covering 100 square degrees in five passbands, Jain et al. (1999) attempted to detect this magnification bias-induced cross-correlation between two galaxy populations. From their photometric redshift estimates for the galaxies, they define the foreground and background galaxy samples by $0 \le z_1 \le 0.15$ and $0.35 \le z_2 \le 0.45$, together with a magnitude cut at $r \le 20.5$. The large gap between the two redshift ranges accounts for the fact that photometric redshifts have an uncertainty of slightly less that $\Delta z = 0.1$, so that this conservative cut should minimise the overlap between the two populations. At the magnitude cut, the so-defined background sample exhibits an effective slope of $\alpha \sim 0.5$, so that lensing should produce an anti-correlation. In fact, Jain et al. (1999) found that ξ_{12} is negative for $\phi \gtrsim 1'$, but slightly positive for smaller angular separations. Note that this behaviour is expected from eq. (7.21), since the positive correlation at small angles is due to the prevalence of the intrinsic cross-correlation owing to the redshift overlap of the two samples. In order to strengthen their interpretation of this result, Jain et al. (1999) split their background sample into a red and a blue half. The number-count slope of these two subsamples of background galaxies at the magnitude cut is $\alpha \sim 0$ and $\alpha \sim 1$, respectively. Correspondingly, they find that ξ_{12} calculated with the blue subsample shows no sign of an anti-correlation at any angular separation, whereas the red subsample shows a stronger anti-correlation than for the total sample of background galaxies. Hence, it seems that the magnification bias of galaxies has been measured; given that the data on which this result is based constitutes only $\sim 1\%$ of the total imaging data the Sloan Survey will accumulate, it is clear that the correlation function ξ_{12} will be measurable with high precision out to large angular separations, providing a very convenient handle on Ω/b, and the scale dependence of b at redshifts $z \sim 0.1$.

7.6. Outlook

Cross correlations between distant QSOs and foreground galaxies on angular scales of about ten arc minutes have been observed, and they can be attributed to the magnification bias due to gravitational lensing by large-scale structures. Coherent shear patterns have been detected around QSOs which are significantly correlated with galaxies. The observations so far are in reasonable agreement with theoretical expectations, except for the higher observed signal in the innermost few arc minutes, and the claimed correlation signal on degree scales. While the excess cross-correlation on small scales can be understood by the lensing effects of individual galaxy clusters, correlations on degree scales pose a severe problem for the lensing explanation if they persist, because the lensing-induced cross-correlation quickly dies off beyond scales of approximately $10'$.

QSO-galaxy cross-correlations have the substantial advantage over other diagnostics of weak lensing by large-scale structures that they do not pose any severe observational problems. In particular, it is not necessary to measure either shapes or sizes of faint background galaxies accurately, because it is sufficient to detect and count comparatively bright foreground galaxies near QSOs. However, such counting requires homogeneous photometry, which is difficult to achieve in particular on photographic plates, and requires careful calibration.

Since the QSO-galaxy cross-correlation function involves filtering the density-perturbation power spectrum with a fairly broad function, the zeroth-order Bessel function $\mathrm{J}_0(x)$ [cf. eq. (7.10)], these correlations are not well suited for constraining the power spectrum. If the cluster normalisation is close to the correct one, the QSO-galaxy cross-correlation function is also fairly insensitive to cosmological parameters.

Rather, QSO-galaxy cross correlations are primarily important for measuring the bias parameter b. The rationale of future observations of QSO-galaxy correlations should therefore be to accurately measure the correlation amplitude on scales between a few and 10 arc minutes. On smaller scales, the influence of individual galaxy clusters sets in, and on larger scales, the correlation signal is expected to be weak. Once it becomes possible to reliably constrain the density-fluctuation power spectrum, such observations can then be used to quantify the bias parameter, and thereby provide most valuable information for theories of galaxy formation. A possible dependence of the bias parameter on scale and redshift can also be extracted.

Sufficiently large data fields for this purpose will soon become available, in particular through wide-field surveys like the 2dF Survey (Colless 1998) and the Sloan Digital Sky Survey (Gunn & Knapp 1993, Loveday & Pier 1998). It therefore appears feasible that within a few years weak lensing by large-scale structures will be able to quantify the relation between the distributions of galaxies and the dark matter.

8. Galaxy-Galaxy Lensing

8.1. Introduction

Whereas the weak lensing techniques described in Sect. 5. are adequate to map the projected matter distribution of galaxy clusters, individual galaxies are not sufficiently massive to show up in the distortion of the images of background galaxies. From the signal-to-noise ratio (4.55, page 27) we see that individual isothermal halos with a velocity dispersion in excess of $\sim 600\,\mathrm{km\,s^{-1}}$ can be detected at a high significance level with the currently achievable number densities of faint galaxy images. Galaxies have halos of much lower velocity dispersion: The velocity dispersion of an L_* elliptical galaxy is $\sim 220\,\mathrm{km\,s^{-1}}$, that of an L_* spiral $\sim 145\,\mathrm{km\,s^{-1}}$.

However, if one is not interested in the mass properties of individual galaxies, but instead in the statistical properties of massive halos of a population of galaxies, the weak lensing effects of several such galaxies can statistically be superposed. For example, if one considers N_f identical foreground galaxies, the signal-to-noise ratio of the combined weak lensing effect increases as $N_\mathrm{f}^{1/2}$, so that for a typical velocity dispersion for spiral galaxies of $\sigma_v \sim 160\,\mathrm{km\,s^{-1}}$, a few hundred foreground galaxies are sufficient to detect the distortion they induce on the background galaxy images.

Of course, detection alone does not yield new insight into the mass properties of galaxy halos. A quantitative analysis of the lensing signal must account for the fact that 'identical' foreground galaxies cannot be observed. Therefore, the mass properties of galaxies have to be parameterised in order to allow the joint analysis of the foreground galaxy population. In particular, one is interested in the velocity dispersion of a typical (L_*, say) galaxy. Furthermore, the rotation curves of (spiral) galaxies which have been observed out to $\sim 30h^{-1}$ kpc show no hint of a truncation of the dark halo out to this distance. Owing to the lack of dynamical tracers, with the exception of satellite galaxies (Zaritsky & White 1994), a direct observation of the extent of the dark halo towards large radii is not feasible with conventional methods. The method described in this section uses the light bundles of background galaxies as dynamical tracers, which are available at all distances from the galaxies' centres, and are therefore able, at least in principle, to probe the size (or the truncation radius) of the halos. Methods for a quantitative analysis of galaxy halos will be described in Sect. 8.2..

The first attempt at detecting this galaxy-galaxy lensing effect was reported by Tyson et al. (1984), but the use of photographic plates and the relatively poor seeing prevented them from observing a galaxy-galaxy lensing signal. The first detection was reported by Brainerd et al. (1996), and as will be described in Sect. 8.3., several further observational results have been derived.

Gravitational light deflection can also be used to study the dark matter halos of galaxies in clusters. The potential influence of the environment on the halo properties of galaxies can provide a strong hint on the formation and lifetimes of clusters. One might expect that galaxy halos are tidally stripped in clusters and therefore physically smaller than those of field galaxies. In Sect. 8.4., we consider galaxy-galaxy lensing in clusters, and report on some first results.

Intermediate in mass between clusters and galaxies are groups of galaxies. With a characteristic velocity dispersion of ~ 300 km/s, they are also not massive enough to be detected individually with weak lensing techniques. For them, the foregoing remarks also apply: As galaxies, groups can be statistically superposed to investigate the statistical properties of their mass profile. Hoekstra et al. (1999) describe a first application of this technique, finding a highly significant shear signal in a sample of 59 groups detected by spectroscopic methods, which yields an average velocity dispersion of ~ 320 km/s and a mass-to-light ratio of $\sim 250h^{-1}$.

8.2. *The Theory of Galaxy-Galaxy Lensing*

A light bundle from a distant galaxy is affected by the tidal field of many foreground galaxies. Therefore, in order to describe the image distortion, the whole population of foreground galaxies has to be taken into account. But first we shall consider the simple case that the image shape is affected (mainly) by a single foreground galaxy. Throughout this section we assume that the shear is weak, so that we can replace (4.12, page 22) by

$$\varepsilon^{(s)} = \varepsilon - \gamma\,. \tag{8.1}$$

Consider an axi-symmetric mass distribution for the foreground galaxy, and background images at separation θ from its centre. The expectation value of the image ellipticity then is the shear at θ, which is oriented tangentially. If $p(\varepsilon)$ and $p^{(s)}(\varepsilon^{(s)})$ denote the probability distributions of the image and source ellipticities, then according to (8.1),

$$p(\varepsilon) = p^{(s)}(\varepsilon-\gamma) = p^{(s)}(\varepsilon) - \gamma_\alpha \frac{\partial}{\partial\varepsilon_\alpha} p^{(s)}(\varepsilon)\,, \tag{8.2}$$

where the second equality applies for $|\gamma| \ll 1$. If φ is the angle between the major axis of the image ellipse and the line connecting source and lens centre, one finds the probability distribution of φ by integrating (8.2) over the modulus of ε,

$$p(\varphi) = \int \mathrm{d}|\varepsilon|\,|\varepsilon|\,p(\varepsilon) = \frac{1}{2\pi} - \gamma_\mathrm{t}\cos(2\varphi)\frac{1}{2\pi}\int \mathrm{d}|\varepsilon|\;p^{(s)}(\varepsilon)\,, \tag{8.3}$$

where φ ranges within $[0,2\pi]$. Owing to the symmetry of the problem, we can restrict φ to within 0 and $\pi/2$, so that the probability distribution becomes

$$p(\varphi) = \frac{2}{\pi}\left[1 - \gamma_\mathrm{t}\left\langle\frac{1}{\varepsilon^{(s)}}\right\rangle\cos(2\varphi)\right]\,, \tag{8.4}$$

i.e., the probability distribution is skewed towards values larger than $\pi/4$, showing preferentially a tangential alignment.

Lensing by additional foreground galaxies close to the line-of-sight to the background galaxy does not substantially change the probability distribution (8.4). First of all, since we assume weak lensing throughout, the effective shear acting on a light bundle can well be approximated by the sum of the shear contributions from the individual foreground galaxies. This follows either from the linearity of the propagation equation in the mass distribution, or from the lowest-order approximation of multiple-deflection gravitational lensing (e.g., Blandford & Narayan 1986; Seitz & Schneider 1992). Second, the additional lensing galaxies are placed at random angles around the line-of-sight, so that the expectation value of their combined shear averages to zero. Whereas they slightly increase the dispersion of the observed image ellipticities, this increase is negligible since the dispersion of the intrinsic ellipticity distribution is by far the dominant effect. However, if the lens galaxy under consideration is part of a galaxy concentration, such as a cluster, the surrounding galaxies are not isotropically distributed, and the foregoing argument is invalid. We shall consider galaxy-galaxy lensing in clusters in Sect. 8.4., and assume here that the galaxies are generally isolated.

For an ensemble of foreground-background pairs of galaxies, the probability distribution for the angle φ simply reads

$$p(\varphi) = \frac{2}{\pi}\left[1 - \langle\gamma_\mathrm{t}\rangle\left\langle\frac{1}{\varepsilon^{(s)}}\right\rangle\cos(2\varphi)\right]\,, \tag{8.5}$$

where $\langle\gamma_\mathrm{t}\rangle$ is the mean tangential shear of all pairs considered. The function $p(\varphi)$ is an observable. A significant deviation from a uniform distribution signals the presence of galaxy-galaxy lensing. To obtain quantitative information on the galaxy halos from the amplitude of the cosine term, one needs to know $\langle 1/\varepsilon^{(s)}\rangle$. It can directly be derived from observations because the weak shear assumed here does not significantly change this average between source and image ellipticities, from a parameterised relation between observable galaxy properties, and from the mean shear $\langle\gamma_\mathrm{t}\rangle$. Although in principle fine binning in galaxy properties (like colour, redshift, luminosity, morphology) and angular separation of foreground-background pairs is possible in order to probe the shear as a function of angular distance from a well-defined set of foreground galaxies and thus to obtain its radial mass profile without any parameterisation, this approach is currently unfeasible owing to the relatively small fields across which observations of sufficient image quality are available.

A convenient parameterisation of the mass profile is the truncated isothermal sphere with surface mass density

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi}\left(1 - \frac{\xi}{\sqrt{s^2+\xi^2}}\right)\,, \tag{8.6}$$

where s is the truncation radius. This is a special case of the mass distribution (3.20, page 18). Brainerd et al. (1996) showed that this mass profile corresponds to a physically realisable dark-matter particle distribution.[15] The velocity dispersion is assumed to scale with luminosity according to (2.68, page 12), which is supported by observations. A similar scaling of s with luminosity L or velocity dispersion σ_v is also assumed,

$$s = s_*\left(\frac{\sigma_v}{\sigma_{v,*}}\right)^2 = s_*\left(\frac{L}{L_*}\right)^{2/\alpha}\,, \tag{8.7}$$

where the choice of the exponent is largely arbitrary. The scaling in (8.7) is such that the ratio of truncation radius and Einstein radius at fixed redshift is independent of L. If, in addition, $\alpha = 4$, the total mass-to-light ratio is identical for all galaxies. The fiducial luminosity L_* may depend on redshift. For instance, if the galaxies evolve passively, their mass properties are unaffected, but aging of the stellar population cause them to become fainter with decreasing redshift. This effect may be important for very deep observations, such as the Hubble Deep Field (Hudson et al. 1998), in which the distribution of lens galaxies extends to high redshifts.

The luminosity L of a lens galaxy can be inferred from the observed flux and an assumed redshift. Since the scaling relation (2.68) applies to the luminosity measured in a particular waveband, the calculation of the luminosity from the apparent magnitude in a specified filter needs to account for the k-correction. If data are available in a single waveband only, an approximate average k-correction relation has to be chosen. For multi-colour data, the k-correction can be estimated for individual galaxies more reliably. In any case, one assumes a relation between luminosity, apparent magnitude, and redshift,

$$L = L(m,z)\,. \tag{8.8}$$

[15] It is physically realisable in the sense that there exists an isotropic, non-negative particle distribution function which gives rise to a spherical density distribution corresponding to (8.6).

The final aspect to be discussed here is the redshift of the galaxies. Given that a galaxy-galaxy analysis involves at least several hundred foreground galaxies, and even more background galaxies, one cannot expect that all of them have spectroscopically determined redshifts. In a more favourable situation, multi-colour data are given, from which a redshift estimate can be obtained, using the photometric redshift method (e.g., Connolly et al. 1995; Gwyn & Hartwick 1996; Hogg et al. 1998). These redshift estimates are characteristically accurate to $\Delta z \sim 0.1$, depending on the photometric accuracy and the number of filter bands in which photometric data are measured. For a single waveband only, one can still obtain a redshift estimate, but a quite unprecise one. One then has to use the redshift distribution of galaxies at that particular magnitude, obtained from spectroscopic or multi-colour redshift surveys in other fields. Hence, one assumes that the redshift probability distribution $p_z(z;m)$ as a function of magnitudes is known sufficiently accurately.

Suppose for a moment that all galaxy redshifts were known. Then, one can predict the effective shear for each galaxy, caused by all the other galaxies around it,

$$\gamma_i = \sum_j \gamma_{ij}(\vec{\theta}_i - \vec{\theta}_j, z_i, z_j, m_j) \; , \qquad (8.9)$$

where γ_{ij} is the shear produced by the j-th galaxy on the i-th galaxy image, which depends on the angular separation and the mass properties of the j-th galaxy. From its magnitude and redshift, the luminosity can be inferred from (8.8), which fixes σ_v and the halo size s through the scaling relations (2.68) and (8.7). Of course, for $z_i \le z_j$, $\gamma_{ij} = 0$. Although the sum in (8.9) should in principle extend over the whole sky, the lensing effect of all foreground galaxies with angular separation larger than some $\theta_{\rm max}$ will average to zero. Therefore, the sum can be restricted to separations $\le \theta_{\rm max}$. We shall discuss the value of $\theta_{\rm max}$ further below.

In the realistic case of unknown redshifts, but known probability distribution $p_z(z;m)$, the shear γ_i cannot be determined. However, by averaging (8.9) over $p_z(z;m)$, the mean and dispersion, $\langle\gamma_i\rangle$ and $\sigma_{\gamma,i}$, of the shear for the i-th galaxy can be calculated. Instead of performing the high-dimensional integration explicitly, this averaging can conveniently be done by a Monte-Carlo integration. One can generate multiple realisations of the redshift distribution by randomly drawing redshifts from the probability density $p_z(z;m)$. For each realisation, the γ_i can be calculated from (8.9). By averaging over the realisations, the mean $\langle\gamma_i\rangle$ and dispersion $\sigma_{\gamma,i}$ of γ_i can be estimated.

8.3. *Results*

The first attempt at detecting galaxy-galaxy lensing was made by Tyson et al. (1984). They analysed a deep photographic survey consisting of 35 prime-focus plates with the 4-meter Mayall Telescope at Kitt Peak. An area of 36 (arc min.)^2 on each plate was digitised. After object detection, $\sim 12{,}000$ 'foreground' and $\sim 47{,}000$ 'background' galaxies were selected by their magnitudes, such that the faintest object in the 'foreground' class was one magnitude brighter than the brightest 'background' galaxy. This approach assumes that the apparent magnitude of an object provides a good indication for its redshift, which seems to be valid, although the redshift distributions of 'foreground' and 'background' galaxies will substantially overlap. There were $\sim 28{,}000$ foreground-background pairs with $\Delta\theta \le 63''$ in their sample, but no significant tangential alignment could be measured. By comparing their observational results with Monte-Carlo simulations, Tyson et al. concluded that the characteristic velocity dispersion of a foreground galaxy in their sample must be smaller than about $120\,{\rm km\,s^{-1}}$. This limit was later revised upwards to $\sim 230\,{\rm km\,s^{-1}}$ by Kovner & Milgrom (1987) who noted that the assumption made in Tyson et al.'s analysis that all background galaxies are at infinite distance (i.e., $D_{\rm ds}/D_{\rm s} = 1$) was critical. This upper limit is fully compatible with our knowledge of galaxy masses.

This null-detection of galaxy-galaxy lensing in a very large sample of objects apparently discouraged other attempts for about a decade. After the first weak-lensing results on clusters became available, it was obvious that this method requires deep data with superb image quality. In particular, the non-linearity of photographic plates and mediocre seeing conditions are probably fatal to the detection of this effect, owing to its smallness. The shear at $5''$ from an L_* galaxy with $\sigma_v = 160\,{\rm km\,s^{-1}}$ is less than 5%, and pairs with smaller separations are very difficult to investigate as the bright galaxy will affect the ellipticity measurement of its close neighbour on ground-based images.

Using a single $9\rlap{.}'6 \times 9\rlap{.}'6$ blank field, with a total exposure time of nearly seven hours on the 5-meter Hale Telescope on Mount Palomar, Brainerd et al. (1996) reported the first detection of galaxy-galaxy lensing. Their co-added image had a seeing of $0\rlap{.}''87$ at FWHM, and the 97% completeness limit was $r = 26$. They considered 'foreground' galaxies in the magnitude range $20 \le r \le 23$, and several fainter bins for defining the 'background' population, and investigated the distribution function $p(\varphi)$ for pairs with separation $5'' \le \Delta\theta \le 34''$. The most significant deviation of $p(\varphi)$ from a flat distribution occurs for 'background' galaxies in the range $23 \le r \le 24$. For fainter (and thus smaller) galaxies, the accuracy of the shape determination deteriorates, as Brainerd et al. explicitly show. The number of 'foreground' galaxies, 'background' galaxies, and pairs, is $N_{\rm f} = 439$, $N_{\rm b} = 506$, and $N_{\rm pairs} = 3202$. The binned distribution for this 'background' sample is shown in Fig. 30, together with a fit according to (8.5). A Kolmogorov-Smirnov test rejects a uniform distribution of $p(\varphi)$ at the 99.9% level, thus providing the first detection of galaxy-galaxy lensing.

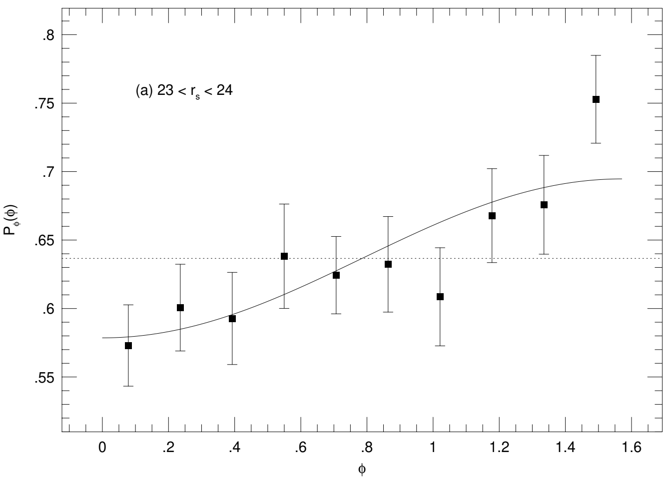

FIG. 30.—The probability distribution $p(\varphi)$ for the 3202 foreground-background pairs ($20 \le r \le 23$ and $23 \le r \le 24$, respectively) with $5'' \le \Delta\theta \le 34''$ in the sample used by Brainerd et al. (1996), together with the best fit according to (8.5). The observed distribution is incompatible with a flat distribution (dotted line) at a high confidence level of 99.9% (from Brainerd et al.).

Brainerd et al. performed a large number of tests to check for possible systematic errors, including null tests (e.g., replacing the positions of 'foreground' galaxies by random points, or stars), splitting the whole sample into various subsamples (e.g., inner part vs. outer part of the image, upper half vs. lower half etc.), and these tests were passed satisfactorily. Also a slight PSF anisotropy in the data, or contamination of the ellipticity measurement of faint galaxies by brighter neighbouring galaxies, cannot explain the observed relative alignment, as tested with extensive simulations, so that the detection must be considered real.

Brainerd et al. then quantitatively analysed their observed alignment, using the model outlined in Sect. 8.2., with $\alpha = 4$. The predictions of the model were inferred from Monte-Carlo simulations, in which galaxies were randomly distributed with the observed number density, and redshifts were assigned according to a probability distribution $p_z(z;m)$, for which they used a slight extrapolation from existing redshift surveys, together with a simple prescription for the k-correction in (8.8) to assign luminosities to the galaxies. The ellipticity for each background galaxy image was then obtained by randomly drawing an intrinsic ellipticity, adding shear according to (8.9). The simulated probability distribution $p(\varphi)$ was discretised into several bins in angular separation $\Delta\theta$, and compared to the observed orientation distribution, using χ^2-minimisation with respect to the model parameters $\sigma_{v,*}$ and s_*. The result of this analysis is shown in Fig. 31. The shape of the χ^2-contours is characteristic in that they form a valley which is relatively narrow in the $\sigma_{v,*}$-direction, but extends very far out into the s_*-direction. Thus, the velocity dispersion $\sigma_{v,*}$ can significantly be constrained with these observations, while only a lower limit on s_* can be derived. Formal 90% confidence limits on $\sigma_{v,*}$ are $\sim 100\,\mathrm{km\,s^{-1}}$ and $\sim 210\,\mathrm{km\,s^{-1}}$, with a best-fitting value of about $160\,\mathrm{km\,s^{-1}}$, whereas the 1- and 2-σ lower limits on s_* are $25\,h^{-1}$ kpc and $\sim 10\,h^{-1}$ kpc, respectively.

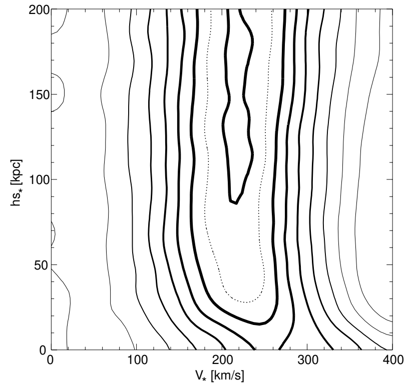

FIG. 31.—Contours of constant χ^2 in the V_*–hs_* parameter plane, where $V_* = \sqrt{2}\sigma_{v,*}$, obtained from a comparison of the observed tangential alignment $\langle\gamma_{\rm t}\rangle$ with the distribution found in Monte-Carlo simulations. The solid contours range from 0.8 (innermost) to 8 per degree of freedom; the dotted curve displays $\chi^2 = 1$ per degree of freedom. (from Brainerd et al. 1996).

Finally, Brainerd et al. studied the dependence of the lensing signal $\langle\gamma_{\rm t}\rangle$ on the colour of their 'background' sample, by splitting it into a red and a blue half. The lensing signal of the former is compatible with zero on all scales, while the blue sample reveals a strong signal which decreases with angular separation as expected. This result is in accordance with that discussed in Sect. 5.5.3., where the blue galaxies showed a stronger lensing signal as well, indicating that their redshift distribution extends to larger distances.

We have discussed the work of Brainerd et al. (1996) in some detail since it provided the first detection of galaxy-galaxy lensing, and since it was the only one obtained from the ground until recently. Also, their careful analysis exemplifies the difficulties in deriving a convincing result.

Griffiths et al. (1996) analysed the images from the *Hubble Space Telescope* Medium Deep Survey (MDS) in terms of galaxy-galaxy lensing. The MDS is an imaging survey, using parallel data obtained with the WFPC2 camera on-board HST. They identified 1600 'foreground' ($15 < I < 22$) and 14000 'background' ($22 < I < 26$) galaxies. Owing to the spatial resolution of the HST, a morphological classification of the foreground galaxies could be performed, and spiral and elliptical galaxies could separately be analysed. They considered the mean orientation angle $\langle\varphi\rangle = \pi/4 + \pi^{-1}\langle\gamma_{\rm t}\rangle\langle 1/|\varepsilon^{(\rm s)}|\rangle$ as a statistical variable, and scaled the truncation radius in their mass models in proportion to the half-light radius. They found that $\sigma_{v,*} = 220\,\mathrm{km\,s^{-1}}$ and $\sigma_{v,*} = 160\,\mathrm{km\,s^{-1}}$ are compatible with their shear data for elliptical and spiral galaxies, respectively. For their sample of elliptical foreground galaxies, they claim that the truncation radius must be more than ten times the half-light radius to fit their data, and that a de Vaucouleurs mass profile is excluded. Unfortunately, no significance levels are quoted.

A variant of the method for a quantitative analysis of galaxy-galaxy lensing was developed by Schneider & Rix (1997). Instead of a χ^2-analysis of $\langle\gamma_{\rm t}\rangle$ in angular separation bins, they suggested a maximum-likelihood analysis, using the individual galaxy images. In their Monte-Carlo approach, the galaxy positions (and magnitudes) are kept fixed, and only the redshifts of the galaxies are drawn from their respective probability distribution $p_z(z;m)$, as described at the end of Sect. 8.2.. The resulting log-likelihood function

$$\ell = -\sum_i \frac{|\varepsilon_i - \langle\gamma_i\rangle|^2}{\rho^2 + \sigma^2_{\gamma,i}} - \sum_i \ln\left[\pi(\rho^2 + \sigma^2_{\gamma,i})\right] , \tag{8.10}$$

where ρ is the dispersion of intrinsic ellipticity distribution, here assumed to be a Gaussian, can then be maximised with respect to the model parameters, e.g., $\sigma_{v,*}$ and s_*. Extensive simulations demonstrated that this approach, which utilises all of the information provided by observations, yields an unbiased estimate of these model parameters. Later, Erben (1997) showed that this remains valid even if the lens galaxies have elliptical projected mass profiles.

This method was applied to the deep multi-colour imaging data of the Hubble Deep Field (HDF; Williams et al. 1996) by Hudson et al. (1998), after Dell'Antonio & Tyson (1996) detected a galaxy-galaxy lensing signal in the HDF on an angular scale of $\lesssim 5''$. The availability of data in four wavebands allows an estimate of photometric redshifts, a method demonstrated to be quite reliable by spectroscopy of HDF galaxies (e.g., Hogg et al. 1998). The accurate redshift estimates, and the depth of the HDF, compensates for the small field-of-view of $\sim 5\,\mathrm{arcmin}^2$. A similar study of the HDF data was carried out by the Caltech group (see Blandford et al. 1998).

In order to avoid k-corrections, using the multi-colour photometric data to relate all magnitudes to the rest-frame B-band, Hudson et al. considered lens galaxies with redshift $z \lesssim 0.85$ only, leaving 208 galaxies. Only such source-lens pairs for which the estimated redshifts differ by at least 0.5 were included in the analysis, giving about 10^4 foreground-background pairs. They adopted the same parameterisation for the lens population as described in Sect. 8.2., except that the depth of the HDF suggests that the fiducial luminosity L_* should be allowed to depend on redshift, $L_* \propto (1+z)^\zeta$. Assuming no evolution, $\zeta = 0$, and a Tully-Fisher index of $1/\alpha = 0.35$, they found $\sigma_{v,*} = (160 \pm 30)\,\mathrm{km\,s^{-1}}$. Various control tests were performed to demonstrate the robustness of this result, and potential systematic effects were shown to be negligible.

As in the previous studies, halo sizes could not be significantly constrained. The lensing signal is dominated by spiral galaxies at a redshift of $z \sim 0.6$. Comparing the Tully-Fisher relation at this redshift to the local relation, the lensing results indicate that intermediate-redshift galaxies are fainter than local spirals by 1 ± 0.6 magnitudes in the B-band, at fixed circular velocity.

Hence, all results reported so far yield compatible values of $\sigma_{v,*}$, but do not allow upper bounds on the halo size to be set. The flatness of the likelihood surface in the s_*-direction shows that a measurement of s_* requires much larger samples than used before. We can understand the insensitivity to s_* in the published analyses at least qualitatively. The shear caused by a galaxy at a distance of, say, 100 kpc is very small, of order 1%. This implies that the difference in shear caused by galaxies with truncation radius of 20 kpc and $s = 100$ kpc is very small indeed. In addition, there are typically other galaxies closer to the line-of-sight to background galaxies which produce a larger shear, making it more difficult to probe the shear of widely separated foreground galaxies. Hence, to probe the halo size, many more foreground-background pairs must be considered. In addition, the angular scale $\theta_{\rm max}$ within which pairs are considered needs to be larger than the angular scale of the truncation radius at typical redshifts of the galaxies, and on the other hand, $\theta_{\rm max}$ should be much smaller than the size of the data field available. Hence, to probe large scales of the halo, wide-field imaging data are needed.

There is a related problem which needs to be understood in greater detail. Since galaxies are clustered, and probably (biased) tracers of an underlying dark matter distribution (e.g., most galaxies may live in groups), it is not evident whether the shear caused by a galaxy at a spatial separation of, say, 100 kpc is caused mainly by the dark matter halo of the galaxy itself, or rather by the dark-matter halo associated with the group. Here, numerical simulations of the dark matter may indicate to which degree these two effects can be separated, and observational strategies for this need to be developed.

In fact, the two points just mentioned were impressively illustrated by a galaxy-galaxy lensing analysis of early commissioning imaging data from the Sloan Digital Sky Survey (Fischer et al. 1999), covering 225 square degrees. The separation between foreground and background galaxies was based on apparent magnitude, with an estimated mean redshift of the foreground sample of $\langle z_{\rm d} \rangle \approx 0.17$. Fischer et al. (1999) used data in three optical filters for their analysis; the number of foreground (background) galaxies in each filter is ~ 28000 (1.4×10^6). The galaxy-galaxy lensing signal is seen out to $\sim 10'$ in all three filters, and the mean tangential shear in the annulus $10'' \le \theta \le 10'$ is $\approx 6 \times 10^{-4}$. With an assumed redshift distribution of foreground and background galaxies, the characteristic velocity dispersion could be estimated to be $\sigma_* = 170 \pm 20 \mathrm{km\,s}^{-1}$ at 95% confidence. Even at the large angular separation probed by this data set, no sign of a cut-off radius of the galaxy halos is seen, and a lower limit of $s_* \ge 275 h^{-1}$kpc can be derived. At such scales, the shear is probably no longer dominated by the foreground galaxy used as the origin for the definition of tangential shear, but by neighbouring galaxies and/or dark matter correlated with the galaxy. Therefore, the results of such a study may best be interpreted as a galaxy-mass correlation function (Kaiser 1992), which brings us back to the issue of biasing discussed in Sect. 6.8.. A preliminary analysis presented in Fischer et al. (1999) yields $\Omega_0/b \sim 0.3$, if a linear biasing factor b is assumed. At least as important as the quantitative results from the Sloan Survey is the fact that they demonstrate the enormous potential of this method – this analysis used about 2% of the imaging data the full Sloan Survey will provide, and did not yet utilise photometric redshift information which, as mentioned before, will increase the accuracy of the physical parameters derived.

8.4. *Galaxy-Galaxy Lensing in Galaxy Clusters*

An interesting extension of the work described above aims at the investigation of the dark-matter halo properties of galaxies within galaxy clusters. In the hierarchical model for structure formation, clusters grow by mergers of less massive halos, which by themselves formed by merging of even smaller substructures. Tidal forces in clusters, possible ram-pressure stripping by the intra-cluster medium, and close encounters during the formation process, may affect the halos of galaxies, most of which presumably formed at an early epoch. Therefore, it is unclear at present whether the halo properties of galaxies in clusters are similar to those of field galaxies.

Galaxy-galaxy lensing offers an exciting opportunity to probe the dark galaxy halos in clusters. There are several differences between the investigation of field and of cluster galaxies. First, the number of massive galaxies in a cluster is fairly small, so the statistics for a single cluster will be limited. This can be compensated by investigating several clusters simultaneously. Second, the image distortion is determined by the reduced shear, $g = \gamma/(1-\kappa)$. For field galaxies, where the shear and the surface mass density is small, one can set $g \approx \gamma$, but this approximation no longer holds for galaxies in clusters, where the cluster provides κ substantially above zero. This implies that one needs to know the mass distribution of the cluster before the statistical properties of the massive galaxy halos can be investigated. On the other hand, it magnifies the lensing signal from the galaxies, so that fewer cluster galaxies are needed to derive significant lensing results compared to field galaxies of similar mass. Third, most cluster galaxies are of early type, and thus their $\sigma_{v,*}$ – and consequently, their lensing effect – is expected to be larger than for typical field galaxies.

In fact, the lensing effect of individual cluster galaxies can even be seen from strong lensing. Modelling clusters with many strong-lensing constraints (e.g., several arcs, multiple images of background galaxies), the incorporation of individual cluster galaxies turns out to be necessary (e.g., Kassiola et al. 1992; Wallington et al. 1995; Kneib et al. 1996). However, the resulting constraints are relevant only for a few cluster galaxies which happen to be close to the strong-lensing features, and mainly concern the mass of these galaxies within $\sim 10 h^{-1}$ kpc.

The theory of galaxy-galaxy lensing in clusters was developed in Natarajan & Kneib (1997) and Geiger & Schneider (1998), using several different approaches. The simplest possibility is related to the aperture mass method discussed in Sect. 5.3.1.. Measuring the tangential shear within an annulus around each cluster galaxy, perhaps including a weight function, permits a measurement of the aperture mass, and thus to constrain the parameters of a mass model for the galaxies. Provided the scale of the aperture is sufficiently small, the tidal field of the cluster averages out to first order, and the local influence of the cluster occurs through the local surface mass density κ. In particular, the scale of the aperture should be small enough in order to exclude neighbouring cluster galaxies.

A more sophisticated analysis starts from a mass model of the cluster, as obtained by one of the reconstruction techniques discussed in Sect. 5., or by a parameterised mass model constructed from strong-lensing constraints. Then, parameterised galaxy models are added, again with a prescription similar to that of Sect. 8.2., and simultaneously the mass model of the cluster is multiplied by the relative mass fraction in the smoothly distributed cluster mass (compared to the total mass). In other words, the mass added by inserting galaxies into the cluster is subtracted from the smooth density profile. From the observed galaxy ellipticities, a likelihood function can be defined and maximised with respect to the parameters $(\sigma_{v,*}, s_*)$ of the galaxy model.

Natarajan et al. (1998) applied this method to WFPC2 images of the cluster AC 114 ($z_d = 0.31$). They concluded that most of the mass of a fiducial L_* cluster galaxy is contained in a radius of ~ 15 kpc, indicating that the halo size of galaxies in this cluster is smaller than that of field galaxies.

Using their HST mosaic image, Hoekstra et al. (2000) also detected galaxy-galaxy lensing in the high-redshift cluster MS1054−03 at $z = 0.83$. Avoiding the densest part of the cluster in selecting their foreground galaxies, they investigated the average tangential shear around them, after subtracting the shear from the cluster as determined from the mass reconstruction (see Sect. 5.3., page 36), also using the scaling (2.68). The galaxy-galaxy lensing signal is seen at the 99.8% confidence level. Using the redshift distribution of background galaxies as determined from photometric redshift estimates in the Hubble Deep Fields, their lensing signal yields $\sigma_* \approx 200 \pm 35\,\mathrm{km\,s^{-1}}$ for the cluster galaxies. Not unexpectedly, this value is larger than those obtained from field galaxies, since the cluster preferentially hosts early-type galaxies for which σ_* is known to be larger than for spirals. It is indeed encouraging that this method is able to measure the mass of high-redshift galaxies.

Once the mass contained in the cluster galaxies is a significant fraction of the total mass of the cluster, this method was found to break down, or give strongly biased results. Geiger & Schneider (1999) modified this approach by performing a maximum-likelihood cluster mass reconstruction for each parameter set of the cluster galaxies, allowing the determination of the best representation of the global underlying cluster component that is consistent with the presence of the cluster galaxies and the observed image ellipticities of background galaxies.

This method was then applied to the WFPC-2 image of the cluster Cl0939+4713, already described in Sect. 5.4.. The entropy-regularised maximum-likelihood mass reconstruction of the cluster is very similar to the one shown in Fig. 14 (page 39), except that the cluster centre is much better resolved, with a peak very close to the observed strong lensing features (Trager et al. 1997). Cluster galaxies were selected according to their magnitudes, and divided by morphology into two subsamples, viz. early-type galaxies and spirals. In Fig. 32 we show the likelihood contours in the s_*–$\sigma_{v,*}$ plane, for both subsets of cluster galaxies. Whereas there is no statistically significant detection of lensing by spiral galaxies, the lensing effect of early-type galaxies is clearly detected. Although no firm upper limit of the halo size s_* can be derived from this analysis owing to the small angular field of the image (the maximum of the likelihood function occurs at $8\,h^{-1}$ kpc, and a 1-σ upper limit would be $\sim 50\,h^{-1}$ kpc), the contours 'close' at smaller values of s_* compared to the results obtained from field galaxies. By statistically combining several cluster images, a significant upper limit on the halo size can be expected.

The maximum-likelihood estimate of σ_* for the early-type galaxies is $\sim 200\,\mathrm{km\,s^{-1}}$, in agreement with that found by Hoekstra et al. (2000).

It should be noted that the results presented above still contain some uncertainties, most notably the unknown redshift distribution of the background galaxies and the mass-sheet degeneracy, which becomes particularly severe owing to the small field-of-view of WFPC2. Changing the assumed redshift distribution and the scaling parameter λ in (5.10, page 34) shifts the likelihood contours in Fig. 32 up or down, i.e., the determination of $\sigma_{v,*}$ is affected. As for galaxy-galaxy lensing of field galaxies, the accuracy can be increased by using photometric redshift estimates. Similarly, the allowed range of the mass-sheet transformation can be constrained by combining these small-scale images with larger scale ground-based images, or, if possible, by using magnification information to break the degeneracy. Certainly, these improvements of the method will be a field of active research in the immediate future.

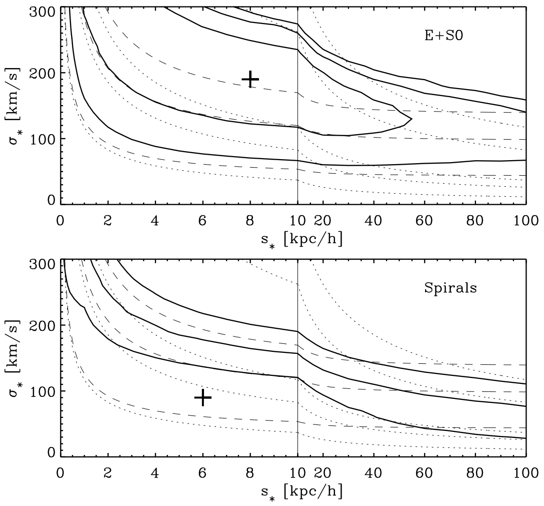

FIG. 32.—Results of applying the entropy-regularised maximum-likelihood method for galaxy-galaxy lensing to the WFPC2 image of the cluster Cl0939+4713. The upper and lower panels correspond to early-type and spiral galaxies, respectively. The solid lines are confidence contours at 68.3%, 95.4% and 99.7%, and the cross marks the maximum of the likelihood function. Dashed lines correspond to galaxy models with equal aperture mass $M_*(< 8h^{-1}\,\mathrm{kpc}) = (0.1, 0.5, 1.0) \times 10^{11}\,h^{-1}\,M_\odot$. Similarly, the dotted lines connect models of constant total mass for an L_*-galaxy, of $M_* = (0.1, 0.5, 1.0, 5.0, 10) \times 10^{11}\,h^{-1}\,M_\odot$, which corresponds to a mass fraction contained in galaxies of $(0.15, 0.75, 1.5, 7.5, 15)$%, respectively (from Geiger & Schneider 1999).

9. THE IMPACT OF WEAK GRAVITATIONAL LIGHT DEFLECTION ON THE MICROWAVE BACKGROUND RADIATION

9.1. *Introduction*

The Cosmic Microwave Background originated in the hot phase after the Big Bang, when photons were created in thermal equilibrium with electromagnetically interacting particles. While the Universe expanded and cooled, the photons remained in thermal equilibrium until the temperature was sufficiently low for electrons to combine with the newly formed nuclei of mainly hydrogen and helium. While the formation of atoms proceeded, the photons decoupled from the matter due to the rapidly decreasing abundance of charged matter. Approximately 300,000 years after the Big Bang, corresponding to a redshift of $z \approx 1{,}000$, the universe became transparent for the radiation, which retained the Planck spectrum it had acquired while it was in thermal equilibrium, and the temperature decreased in proportion with the scale factor as the Universe expanded. This relic radiation, cooled to $T = 2.73$ K, forms the Cosmic Microwave Background (hereafter CMB). Penzias & Wilson (1965) detected it as an "excess antenna temperature", and Fixsen et al. (1996) used the COBE-FIRAS instrument to prove its perfect black-body spectrum.

Had the Universe been ideally homogeneous and isotropic, the CMB would have the intensity of black-body radiation at 2.73 K in all directions on the sky, and would thus be featureless. Density perturbations in the early Universe, however, imprinted their signature on the CMB through various mechanisms, which are thoroughly summarised and discussed in Hu (1995). Photons in potential wells at the time of decoupling had to climb out, thus losing energy and becoming slightly cooler than the average

CMB. This effect, now called the *Sachs-Wolfe effect* was originally studied by Sachs & Wolfe (1967), who found that the temperature anisotropies in the CMB trace the potential fluctuations on the 'surface' of decoupling. CMB fluctuations were first detected by the COBE-DMR experiment (Smoot et al. 1992) and subsequently confirmed by numerous ground-based and balloon-borne experiments (see Smoot 1997 for a review).

The interplay between gravity and radiation pressure in perturbations of the cosmic 'fluid' before recombination gave rise to another important effect. Radiation pressure is only effective in perturbations smaller than the horizon. Upon entering the horizon, radiation pressure provides a restoring force against gravity, leading to acoustic oscillations in the tightly coupled fluid of photons and charged particles, which cease only when radiation pressure drops while radiation decouples. Therefore, for each physical perturbation scale, the acoustic oscillations set in at the same time, i.e. when the horizon size becomes equal the perturbation size, and they end at the same time, i.e. when radiation decouples. At fixed physical scale, these oscillations are therefore coherent, and they show up as distinct peaks (the so-called *Doppler peaks*) and troughs in the power spectrum of the CMB fluctuations. Perturbations large enough to enter the horizon after decoupling never experience these oscillations. Going through the CMB power spectrum from large to small scales, there should therefore be a 'first' Doppler peak at a location determined by the horizon scale at the time of decoupling.

A third important effect sets in on the smallest scales. If a density perturbation is small enough, radiation pressure can blow it apart because its self-gravity is too weak. This effect is comparable to the Jeans' criterion for the minimal mass required for a pressurised perturbation to collapse. It amounts to a suppression of small-scale fluctuations and is called *Silk damping*, leading to an exponential decline at the small-scale end of the CMB fluctuation power spectrum.

Other effects arise between the 'surface' of decoupling and the observer. Rees & Sciama (1968) pointed out that large non-linear density perturbations between the last-scattering surface and us can lead to a distinct effect if those fluctuations change while the photons traverse them. Falling into the potential wells, they experience a stronger blue-shift than climbing out of them because expansion makes the wells shallower in the meantime, thus giving rise to a net blue-shift of photons. Later, this effect was re-examined in the framework of the 'Swiss-Cheese' (Dyer 1976) and 'vacuole' (Nottale 1984) models of density perturbations in an expanding background space-time. The masses of such perturbations have to be very large for this effect to become larger than the Sunyaev-Zel'dovich effect[16] due to the hot gas contained in them; Dyer (1976) estimated that masses beyond $10^{19}\,M_\odot$ would be necessary, a value four to five orders of magnitude larger than that of typical galaxy clusters.

The gravitational lens effect of galaxy clusters moving transverse to the line-of-sight was investigated by Birkinshaw & Gull (1983) who found that a cluster with $\sim 10^{15}\,M_\odot$ and a transverse velocity of $\sim 6000\,\mathrm{km\,s^{-1}}$ should change the CMB temperature by $\sim 10^{-4}\,\mathrm{K}$. Later, Gurvits & I.G. (1986) re-investigated this effect and found it to be about an order of magnitude smaller.

Cosmic strings as another class of rapidly moving gravitational lenses were studied by Kaiser & Stebbins (1984) who discussed that they would give rise to step-like features in the CMB temperature pattern.

[16]The (thermal) Sunyaev-Zel'dovich effect is due to Compton-upscattering of CMB photons by thermal electrons in the hot plasma in galaxy clusters. Since the temperature of the electrons is much higher than that of the photons, CMB photons are effectively re-distributed towards higher energies. At frequencies lower than $\approx 272\,\mathrm{GHz}$, the CMB intensity is thus decreased towards galaxy clusters; in effect, they cast shadows on the surface of the CMB.

9.2. Weak Lensing of the CMB

The introduction shows that the CMB is expected to display distinct features in a hierarchical model of structure formation. The CMB power spectrum should be featureless on large scales, then exhibit pronounced Doppler peaks at scales smaller than the horizon at the time of decoupling, and an exponential decrease due to Silk damping at the small-scale end. We now turn to investigate whether and how gravitational lensing by large-scale structures can alter these features.

The literature on the subject is rich (see Blanchard & Schneider 1987, Cayón et al. 1993b, Cayón et al. 1993a, Cole & Efstathiou 1989, Fukugita et al. 1992, Kashlinsky 1988, Linder 1988, Linder 1990a, Linder 1990b, Martínez-González et al. 1990, Sasaki 1989, Tomita 1989, Watanabe & Tomita 1991), but different authors have sometimes arrived at contradicting conclusions. Perhaps the most elegant way of studying weak lensing of the CMB is the power-spectrum approach, which was most recently advocated by Seljak (1994, 1996).

We should like to start our discussion by clearly stating two facts concerning the effect of lensing on fluctuations in the Cosmic Microwave Background which clarify and resolve several apparently contradictory discussions and results in the literature.

1. *If the CMB was completely isotropic, gravitational lensing would have no effect whatsoever because it conserves surface brightness.* In this case, lensing would only magnify certain patches in the sky and de-magnify others, but since it would not alter the surface brightness in the magnified or de-magnified patches, the temperature remained unaffected. An analogy would be observers facing an infinitely extended homogeneously coloured wall, seeing some parts of it enlarged and others shrunk. Regardless of the magnification, they would see the same colour everywhere, and so they would notice nothing despite the magnification.

2. *It is not the absolute value of the light deflection due to lensing which matters, but the relative deflection of neighbouring light rays.* Imagine a model universe in which all light rays are isotropically deflected by the same arbitrary amount. The pattern of CMB anisotropies seen by an observer would then be coherently shifted relative to the intrinsic pattern, but remain unchanged otherwise. It is thus merely the *dispersion* of deflection angles what is relevant for the impact of lensing on the observed CMB fluctuation pattern.

9.3. CMB Temperature Fluctuations

In the absence of any lensing effects, we observe at the sky position $\vec\theta$ the intrinsic CMB temperature $T(\vec\theta)$. There are fluctuations $\Delta T(\vec\theta)$ in the CMB temperature about its average value $\langle T\rangle = 2.73\,\mathrm{K}$. We abbreviate the relative temperature fluctuations by

$$\frac{\Delta T(\vec\theta)}{\langle T\rangle} \equiv \tau(\vec\theta) \qquad (9.1)$$

in the following. They can statistically be described by their angular auto-correlation function

$$\xi_\mathrm{T}(\phi) = \left\langle \tau(\vec\theta)\,\tau(\vec\theta+\vec\phi) \right\rangle\,, \qquad (9.2)$$

with the average extending over all positions $\vec\theta$. Due to statistical isotropy, $\xi_\mathrm{T}(\phi)$ depends neither on the position θ nor on the direction of $\vec\phi$, but only on the absolute separation ϕ of the correlated points.

Commonly, CMB temperature fluctuations are also described in terms of the coefficients a_{lm} of an expansion into spherical harmonics,

$$\tau(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm} Y_l^m(\theta,\phi) \;, \tag{9.3}$$

and the averaged expansion coefficients constitute the angular power spectrum C_l of the CMB fluctuations,

$$C_l = \left\langle |a_{lm}|^2 \right\rangle \;. \tag{9.4}$$

It can then be shown that the correlation function $\xi_{\rm T}(\phi)$ is related to the power-spectrum coefficients C_l through

$$C_l = \int_0^{\pi} {\rm d}\phi \, \sin(\phi) \, P_l(\cos\phi) \, \xi_{\rm T}(\phi) \;, \tag{9.5}$$

with the Legendre functions $P_l(\cos\phi)$.

9.4. *Auto-Correlation Function of the Gravitationally Lensed CMB*

9.4.1. *Definitions*

If there are any density inhomogeneities along the line-of-sight towards the last-scattering surface at $z \approx 1{,}000$ (the 'source plane' of the CMB), a light ray starting into direction $\vec{\theta}$ at the observer will intercept the last-scattering surface at the deflected position

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) \;, \tag{9.6}$$

where $\vec{\alpha}(\vec{\theta})$ is the (position-dependent) deflection angle experienced by the light ray. We will therefore observe, at position $\vec{\theta}$, the temperature of the CMB at position $\vec{\beta}$, or

$$T(\vec{\beta}) \equiv T'(\vec{\theta}) = T[\vec{\theta} - \vec{\alpha}(\vec{\theta})] \;. \tag{9.7}$$

The intrinsic temperature autocorrelation function is thus changed by lensing to

$$\xi'_{\rm T}(\phi) = \left\langle \tau[\vec{\theta} - \alpha(\vec{\theta})] \, \tau[(\vec{\theta}+\vec{\phi}) - \vec{\alpha}(\vec{\theta}+\vec{\phi})] \right\rangle \;. \tag{9.8}$$

For simplicity of notation, we further abbreviate $\vec{\alpha}(\vec{\theta}) \equiv \vec{\alpha}$ and $\vec{\alpha}(\vec{\theta}+\vec{\phi}) \equiv \vec{\alpha}'$ in the following.

9.4.2. *Evaluation*

In this section we evaluate the modified correlation function (9.8) and quantify the lensing effects. For this purpose, it is convenient to decompose the relative temperature fluctuation $\tau(\vec{\theta})$ into Fourier modes,

$$\tau(\vec{\theta}) = \int \frac{{\rm d}^2\vec{l}}{(2\pi)^2} \, \hat{\tau}(\vec{l}) \, \exp({\rm i}\vec{l}\vec{\theta}) \;. \tag{9.9}$$

The expansion of $\tau(\vec{\theta})$ into Fourier modes rather than into spherical harmonics is permissible because we do not expect any weak-lensing effects on large angular scales, so that we can consider $T(\vec{\theta})$ on a plane locally tangential to the sky rather than on a sphere.

We insert the Fourier decomposition (9.9) into the expression for the correlation function (9.8) and perform the average. We need to average over ensembles and over the random angle between the wave vector $\vec{l}$ of the temperature modes and the angular separation $\vec{\phi}$ of the correlated points. The ensemble average corresponds to averaging over realisations of the CMB temperature fluctuations in a sample of universes or, since we focus on small scales, over a large number of disconnected regions on the sky. This average introduces the CMB fluctuation spectrum $P_{\rm T}(l)$, which is defined by

$$\left\langle \hat{\tau}(\vec{l}) \, \hat{\tau}^*(\vec{l}') \right\rangle \equiv (2\pi)^2 \, \delta^{(2)}(\vec{l} - \vec{l}') \, P_{\rm T}(l) \;. \tag{9.10}$$

Averaging over the angle between $\vec{l}$ and the position angle $\vec{\phi}$ gives rise to the zeroth-order Bessel function of the first kind, ${\rm J}_0(x)$. These manipulations leave eq. (9.8) in the form

$$\xi'_{\rm T}(\phi) = \int_0^{\infty} \frac{l{\rm d}l}{2\pi} P_{\rm T}(l) \left\langle \exp\left[{\rm i}\vec{l}\left(\vec{\alpha} - \vec{\alpha}'\right)\right] \right\rangle {\rm J}_0(l\phi) \;. \tag{9.11}$$

The average over the exponential in eq. (9.11) remains to be performed. To do so, we first expand the exponential into a power series,

$$\left\langle \exp({\rm i}\vec{l}\delta\vec{\alpha}) \right\rangle = \sum_{j=0}^{\infty} \frac{\langle({\rm i}\vec{l}\delta\vec{\alpha})^j\rangle}{j!} \;, \tag{9.12}$$

where $\delta\vec{\alpha} \equiv \vec{\alpha} - \vec{\alpha}'$ is the deflection-angle difference between neighbouring light rays with initial angular separation $\vec{\phi}$. We now assume that the deflection angles are Gaussian random fields. This is reasonable because (i) deflection angles are due to Gaussian random fluctuations in the density-contrast field as long as the fluctuations evolve linearly, and (ii) the assumption of linear evolution holds well for redshifts where most of the deflection towards the last-scattering surface occurs. Of course, this makes use of the commonly held view that the initial density fluctuations are of Gaussian nature. Under this condition, the odd moments in eq. (9.12) all vanish. It can then be shown that

$$\left\langle \exp({\rm i}\vec{l}\delta\vec{\alpha}) \right\rangle = \exp\left(-\frac{1}{2}l^2\sigma^2(\phi)\right) \tag{9.13}$$

holds exactly, where $\sigma^2(\phi)$ is the dispersion of one component of the deflection-angle,

$$\sigma^2(\phi) \equiv \frac{1}{2}\left\langle (\vec{\alpha} - \vec{\alpha}')^2 \right\rangle \;. \tag{9.14}$$

Even if the assumption that $\delta\vec{\alpha}$ is a Gaussian random field fails, eq. (9.13) still holds approximately. To see this, we note that the CMB power spectrum falls sharply on scales $l \gtrsim l_{\rm c} \approx (10'\,\Omega_0^{1/2})^{-1}$. The scale $l_{\rm c}$ is set by the width of the last-scattering surface at redshift $z \sim 1{,}000$. Smaller-scale fluctuations are efficiently damped by acoustic oscillations of the coupled photon-baryon fluid. Typical angular scales l^{-1} in the CMB fluctuations are therefore considerably larger than the difference between gravitational deflection angles of neighbouring rays, $\delta\vec{\alpha}$, so that $\vec{l}(\vec{\alpha} - \vec{\alpha}')$ is a small number. Hence, ignoring fourth-order terms in $\vec{l}\delta\vec{\alpha}$, the remaining exponential in (9.11) can be *approximated* by

$$\left\langle \exp({\rm i}\vec{l}\delta\vec{\alpha}) \right\rangle \approx 1 - \frac{1}{2}l^2\sigma^2(\phi) \approx \exp\left[-\frac{1}{2}l^2\sigma^2(\phi)\right] \;. \tag{9.15}$$

Therefore, the temperature auto-correlation function modified by gravitational lensing can safely be written,

$$\xi'_{\rm T}(\phi) = \int_0^{\infty} \frac{l{\rm d}l}{2\pi} P_{\rm T}(l) \, \exp\left[-\frac{1}{2}l^2\sigma^2(\phi)\right] {\rm J}_0(l\phi) \;. \tag{9.16}$$

This equation shows that the intrinsic temperature-fluctuation power spectrum is convolved with a Gaussian function in wave number l with dispersion $\sigma^{-1}(\phi)$. The effect of lensing on the CMB temperature fluctuations is thus to smooth fluctuations on angular scales of order or smaller than $\sigma(\phi)$.

9.4.3. Alternative Representations

Equation (9.16) relates the unlensed CMB power spectrum to the lensed temperature auto-correlation function. Noting that $P_{\rm T}(l)$ is the Fourier transform of $\xi_{\rm T}(\phi)$,

$$P_{\rm T}(l) = \int \mathrm{d}^2\phi\, \xi_{\rm T}(\phi)\, \exp(-\mathrm{i}\vec{l}\vec{\phi}) = 2\pi \int \phi \mathrm{d}\phi\, \xi_{\rm T}(\phi)\, \mathrm{J}_0(l\phi)\,, \tag{9.17}$$

we can substitute one for the other. Isotropy permitted us to perform the integration over the (random) angle between $\vec{l}$ and $\vec{\phi}$ in the last step of (9.17). Inserting (9.17) into (9.16) leads to

$$\xi'_{\rm T}(\phi) = \int \phi'\mathrm{d}\phi'\, \xi_{\rm T}(\phi')\, K(\phi,\phi')\,. \tag{9.18}$$

The kernel $K(\phi,\phi')$ is given by

$$\begin{aligned} K(\phi,\phi') &\equiv \int_0^\infty l\mathrm{d}l\, \mathrm{J}_0(l\phi)\, \mathrm{J}_0(l\phi')\, \exp\left[-\frac{1}{2}l^2\sigma^2(\phi)\right] \\ &= \frac{1}{\sigma^2(\phi)} \exp\left[-\frac{\phi^2+\phi'^2}{2\sigma^2(\phi)}\right] \mathrm{I}_0\left[\frac{\phi\phi'}{\sigma^2(\phi)}\right]\,, \end{aligned} \tag{9.19}$$

where $\mathrm{I}_0(x)$ is the modified zeroth-order Bessel function. Equation (6.663.2) of Gradshteyn & Ryzhik (1994) was used in the last step. As will be shown below, $\sigma(\phi) \ll \phi$, so that the argument of I_0 is generally a very large number. Noting that $\mathrm{I}_0(x) \approx (2\pi x)^{-1/2}\exp(x)$ for $x \to \infty$, we can write eq. (9.18) in the form

$$\xi'_{\rm T}(\phi) \approx \frac{1}{(2\pi\phi)^{1/2}\sigma(\phi)} \int \mathrm{d}\phi'\, \phi'^{1/2}\, \xi_{\rm T}(\phi')\, \exp\left[-\frac{(\phi-\phi')^2}{2\sigma^2(\phi)}\right]\,. \tag{9.20}$$

Like eq. (9.16), this expression shows that lensing smoothes the intrinsic temperature auto-correlation function $\xi_{\rm T}(\phi)$ over angular scales of $\sigma(\phi)$. Note in particular that, if $\sigma(\phi) \to 0$, the exponential in (9.20) tends towards a Dirac delta distribution,

$$\lim_{\sigma(\phi)\to 0} \frac{1}{\sqrt{2\pi}\sigma(\phi)} \exp\left[-\frac{(\phi-\phi')^2}{2\sigma^2(\phi)}\right] = \delta(\phi-\phi')\,, \tag{9.21}$$

so that the lensed and unlensed temperature auto-correlation functions agree, $\xi_{\rm T}(\phi) = \xi'_{\rm T}(\phi)$.

Likewise, one can Fourier back-transform eq. (9.16) to obtain a relation between the lensed and the un-lensed CMB power spectra. To evaluate the resulting integral, it is convenient to assume $\sigma(\phi) = \epsilon\phi$, with ϵ being either a constant or a slowly varying function of ϕ. This assumption will be justified below. One then finds

$$P'_{\rm T}(l') = \int_0^\infty \frac{\mathrm{d}l}{\epsilon^2 l} P_{\rm T}(l)\, \exp\left(-\frac{l^2+l'^2}{2\epsilon^2 l^2}\right) \mathrm{I}_0\left(\frac{l'}{\epsilon^2 l}\right)\,. \tag{9.22}$$

For $\epsilon \ll 1$, this expression can be simplified to

$$P'_{\rm T}(l') = \int_0^\infty \frac{\mathrm{d}l}{\sqrt{2\pi}\epsilon l} P_{\rm T}(l)\, \exp\left[-\frac{(l-l')^2}{2\epsilon^2 l^2}\right]\,. \tag{9.23}$$

9.5. Deflection-Angle Variance

9.5.1. Auto-Correlation Function of Deflection Angles

We proceed by evaluating the dispersion $\sigma^2(\phi)$ of the deflection angles. This is conveniently derived from the deflection-angle auto-correlation function,

$$\xi_{\vec{\alpha}}(\phi) \equiv \left\langle \vec{\alpha}\vec{\alpha}' \right\rangle\,. \tag{9.24}$$

Note that the correlation function of $\vec{\alpha}$ is the sum of the correlation functions of the components of $\vec{\alpha}$,

$$\xi_{\vec{\alpha}} = \langle \vec{\alpha}\vec{\alpha}' \rangle = \langle \alpha_1\alpha'_1 \rangle + \langle \alpha_2\alpha'_2 \rangle = \xi_{\alpha_1} + \xi_{\alpha_2}\,. \tag{9.25}$$

In terms of the autocorrelation function, the dispersion $\sigma^2(\phi)$ can be written

$$\sigma^2(\phi) = \frac{1}{2}\left\langle \left[\vec{\alpha}-\vec{\alpha}'\right]^2 \right\rangle = \xi_{\vec{\alpha}}(0) - \xi_{\vec{\alpha}}(\phi)\,. \tag{9.26}$$

The deflection angle is given by eq. (6.12) on page 46 in terms of the Newtonian potential Φ of the density fluctuations δ along the line-of-sight. For lensing of the CMB, the line-of-sight integration extends along the (unperturbed) light ray from the observer at $w = 0$ to the last-scattering surface at $w(z \approx 1000)$; see the derivation in Sect. 6.2. leading to eq. (6.12, page 46).

We introduced the effective convergence in (6.15, page 46) as half the divergence of the deflection angle. In Fourier space, this equation can be inverted to yield the Fourier transform of the deflection angle,

$$\hat{\vec{\alpha}}(\vec{l}) = -\frac{2\mathrm{i}\,\hat{\kappa}_{\rm eff}(\vec{l})}{|\vec{l}|^2}\,\vec{l}\,. \tag{9.27}$$

The deflection-angle power spectrum can therefore be written as

$$P_{\vec{\alpha}}(\phi) = \frac{4}{l^2} P_\kappa(l)\,. \tag{9.28}$$

The deflection-angle autocorrelation function is obtained from eq. (9.28) via Fourier transformation. The result is

$$\xi_{\vec{\alpha}}(\phi) = \int \frac{\mathrm{d}^2\vec{l}}{(2\pi)^2} P_{\vec{\alpha}}(l)\, \exp(-\mathrm{i}\vec{l}\vec{\phi}) = 2\pi \int_0^\infty l\mathrm{d}l\, P_\kappa(l)\, \frac{\mathrm{J}_0(l\phi)}{(\pi l)^2}\,, \tag{9.29}$$

similar to the form (6.60, page 55), but here the filter function is no longer a function of the *product* $l\phi$ only, but of l and ϕ separately,

$$F(l,\phi) = \frac{\mathrm{J}_0(l\phi)}{(\pi l)^2} = \phi^2\,\frac{\mathrm{J}_0(l\phi)}{(\pi l\phi)^2}\,. \tag{9.30}$$

We plot $\phi^{-2}F(l,\phi)$ in Fig. 33. For fixed ϕ, the filter function suppresses small-scale fluctuations, and it tends towards $F(l,\phi) \to (\pi l)^{-2}$ for $l \to 0$.

Inserting $P_\kappa(l)$ into (9.29), we find the explicit expression for the deflection-angle auto-correlation function,

$$\begin{aligned} \xi_{\vec{\alpha}}(\phi) &= \frac{9H_0^4\Omega_0^2}{c^4} \int_0^w \mathrm{d}w'\, W^2(w,w')\, a^{-2}(w') \\ &\times \int_0^\infty \frac{\mathrm{d}k}{2\pi k} P_\delta(k,w')\, \mathrm{J}_0[f_K(w')k\phi]\,. \end{aligned} \tag{9.31}$$

Despite the obvious similarity between this result and the magnification auto-correlation function (6.34 on page 49), it is worth noting two important differences. First, the weighting of the integrand along the line-of-sight differs by a factor of $f_K^2(w')$ because

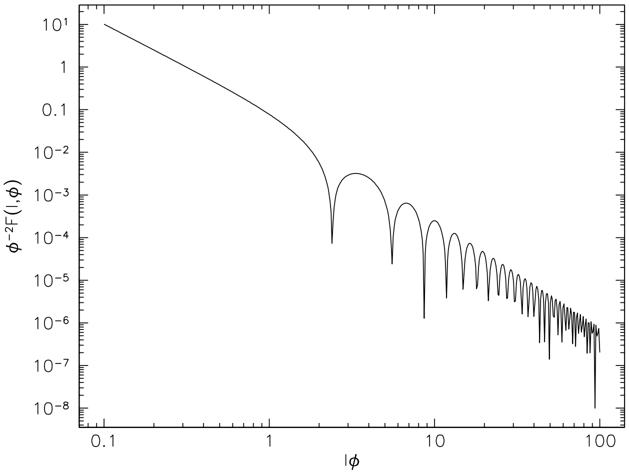

FIG. 33.—The filter function $F(l,\phi)$ as defined in eq. (9.30), divided by ϕ^2, is shown as a function of $l\phi$. Compare Fig. 22 on page 54. For fixed ϕ, the filter function emphasises large-scale projected density perturbations (i.e. structures with small l).

we integrate deflection-angle components rather than the convergence, i.e. first rather than second-order derivatives of the potential Φ. Consequently, structures near the observer are weighted more strongly than for magnification or shear effects. Secondly, the wave-number integral is weighted by k^{-1} rather than k, giving most weight to the largest-scale structures. Since their evolution remains linear up to the present, it is expected that non-linear density evolution is much less important for lensing of the CMB than it is for cosmic magnification or shear.

9.5.2. Typical Angular Scale

A typical angular scale ϕ_g for the coherence of gravitational light deflection can be obtained as

$$\phi_g \equiv \left[\frac{1}{\xi_{\vec\alpha}(0)}\left(\left|\frac{\partial^2\xi_{\vec\alpha}(\phi)}{\partial\phi^2}\right|_{\phi=0}\right)\right]^{-1/2} . \tag{9.32}$$

As eq. (9.31) shows, the deflection-angle auto-correlation function depends on ϕ only through the argument of the Bessel function $\mathrm{J}_0(x)$. For small arguments x, the second-order derivative of the $\mathrm{J}_0(x)$ is approximately $\mathrm{J}_0{}''(x) \approx -\mathrm{J}_0(x)/2$. Differentiating $\xi_{\vec\alpha}(\phi)$ twice with respect to ϕ, and comparing the result to the expression for the magnification auto-correlation function $\xi_\mu(\phi)$ in eq. (6.34, page 49), we find

$$\frac{\partial^2\xi_{\vec\alpha}(\phi)}{\partial\phi^2} \approx -\frac{1}{2}\,\xi_\mu(\phi) , \tag{9.33}$$

and thus

$$\phi_g^2 \approx 2\,\frac{\xi_{\vec\alpha}(0)}{\xi_\mu(0)} . \tag{9.34}$$

We shall estimate ϕ_g later after giving a simple expression for $\xi_{\vec\alpha}(\phi)$. The angle ϕ_g gives an estimate of the scale over which gravitational light deflection is coherent.

9.5.3. Special Cases and Qualitative Expectations

We mentioned before that it is less critical here to assume linear density evolution because large-scale density perturbations dominate in the expression for $\xi_{\vec\alpha}(\phi)$. Specialising further to an Einstein-de Sitter universe so that $w \approx 2c/H_0$, eq. (9.31) simplifies to

$$\xi_{\vec\alpha}(\phi) = \frac{9H_0^4}{c^4}\,w\int_0^1 \mathrm{d}y\,(1-y)^2\int_0^\infty \frac{\mathrm{d}k}{2\pi k}\,P_\delta^{(0)}(k)\,\mathrm{J}_0(wyk\phi) , \tag{9.35}$$

with $wy \equiv w'$.

Adopting the model spectra for HDM and CDM specified in eq. (6.37, page 50) and expanding $\xi_{\vec\alpha}(\phi)$ in a power series in ϕ, we find, to second order in ϕ,

$$\xi_{\vec\alpha}(\phi) = A'\,wk_0\left\{\begin{array}{ll}\frac{3}{2\pi}\left[1-\frac{\phi^2}{20}(wk_0)^2\right] & \text{for HDM}\\ \frac{3\sqrt{3}}{8}\left[1-\frac{3\,\phi^2}{40}(wk_0)^2\right] & \text{for CDM}\end{array}\right. . \tag{9.36}$$

Combining these expressions with eqs. (9.34) and (6.38, page 50), we find for the deflection-angle coherence scale ϕ_g

$$\phi_g \approx 3\,(wk_0)^{-1} . \tag{9.37}$$

It is intuitively clear that ϕ_g should be determined by $(wk_0)^{-1}$. Since k_0^{-1} is the typical length scale of light-deflecting density perturbations, it subtends an angle $(wk_0)^{-1}$ at distance w. Thus the coherence angle of light deflection is given by the angle under which the deflecting density perturbation typically appears. The source distance w in the case of the CMB is the comoving distance to $z = 1{,}000$. In the Einstein-de Sitter case, $w = 2$ in units of the Hubble length. Hence, with $k_0^{-1} \approx 12\,(\Omega_0 h^2)\,\mathrm{Mpc}$ [cf. eq. (2.49), page 8], we have $wk_0 \approx 500$. Therefore, the angular scale of the deflection-angle auto-correlation is of order

$$\phi_g \approx 6\times10^{-3} \approx 20' . \tag{9.38}$$

To lowest order in ϕ, the deflection-angle dispersion (9.26) reads

$$\sigma^2(\phi) \propto (wk_0)^3\,\phi^2 . \tag{9.39}$$

The dispersion $\sigma(\phi)$ is plotted in Fig. 34 for the four cosmological models specified in Tab. 1 on page 45 for linear and non-linear evolution of the density fluctuations.

The behaviour of $\sigma(\phi)$ expressed in eq. (9.39) can qualitatively be understood describing the change in the transverse separation between light paths as a random walk. Consider two light paths separated by an angle ϕ such that their comoving transverse separation at distance w is $w\phi$. Let k^{-1} be the typical scale of a potential fluctuation Φ. We can then distinguish two different cases depending on whether $w\phi$ is larger or smaller than k^{-1}. If $w\phi > k^{-1}$, the transverse separation between the light paths is much larger than the typical potential fluctuations, and their deflection will be incoherent. It will be coherent in the opposite case, i.e. if $w\phi < k^{-1}$.

When the light paths are coherently scattered passing a potential fluctuation, their angular separation changes by $\delta\phi_1 \approx w\phi\nabla_\perp(2k^{-1}\nabla_\perp\Phi/c^2)$, which is the change in the deflection angle across $w\phi$. If we replace the gradients by the inverse of the typical scale, k, we have $\delta\phi_1 \approx 2\,w\phi k\Phi/c^2$. Along a distance w, there are $N \approx kw$ such potential fluctuations, so that the total change in angular separation is expected to be $\delta\phi \approx N^{1/2}\delta\phi_1$.

In case of incoherent scattering, the total deflection of each light path is expected to be $\delta\phi \approx N^{1/2}(2k^{-1}\nabla_\perp\Phi/c^2) \approx N^{1/2}\,2\Phi/c^2$, independent of ϕ. Therefore,

$$\sigma^2(\phi) \approx \left\{\begin{array}{lll} N\delta\phi_1^2 \approx (2\Phi/c^2)^2\,(wk)^3\,\phi^2 & \text{for} & \phi < (wk)^{-1}\\ N(2\Phi/c^2)^2 \approx (2\Phi/c^2)^2\,(wk) & \text{for} & \phi > (wk)^{-1}\end{array}\right. . \tag{9.40}$$

This illustrates that the dependence of $\sigma^2(\phi)$ on $(wk)^3\,\phi^2$ for small ϕ is merely a consequence of the random coherent scattering of neighbouring light rays at potential fluctuations. For large ϕ, $\sigma(\phi)$ becomes constant, and so $\sigma(\phi)\phi^{-1} \to 0$. As Fig. 34 shows, the dispersion $\sigma(\phi)$ increases linearly with ϕ for small ϕ and flattens gradually for $\phi > \phi_g \approx (10-20)'$ as expected, because ϕ_g divides coherent from incoherent scattering.

9.5.4. Numerical Results

The previous results were obtained by specialising to linear evolution of the density contrast in an Einstein-de Sitter universe. For arbitrary cosmological parameters, the deflection-angle dispersion has to be computed numerically. We show in Fig. 34 examples for $\sigma(\phi)$ numerically calculated for the four cosmological models detailed in Tab. 1 on page 45. Two curves are plotted for each model. The somewhat steeper curves were obtained for linear, the others for non-linear density evolution.

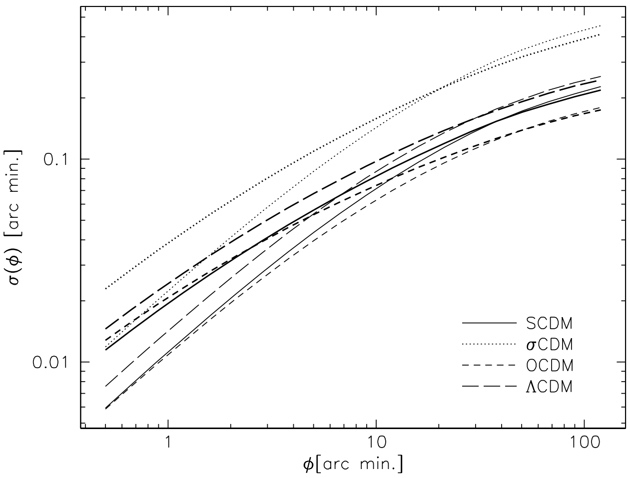

FIG. 34.—The deflection-angle variance $\sigma(\phi)$ is shown for the four cosmological models specified in Tab. 1 on page 45. Two curves are shown for each model, one for linear and one for non-linear evolution of the density fluctuations. Solid curves: SCDM; dotted curve: σCDM; short-dashed curves: OCDM; and long-dashed curves: ΛCDM. The somewhat steeper curves are for linear density evolution. Generally, the deflection-angle variance increases linearly with ϕ for small ϕ, and flattens gradually for $\phi \gtrsim 20'$. At $\phi \approx 10'$, $\sigma(\phi)$ reaches $\approx 0.1'$, or $\approx 0.01\phi$, for the cluster-normalised model universes (all except σCDM; dotted curves). As expected, the effect of non-linear density evolution is fairly moderate, and most pronounced on small angular scales, $\phi \lesssim 10'$.

Figure 34 shows that typical values for the deflection-angle variance in cluster-normalised model universes are of order $\sigma(\phi) \approx (0.03-0.1)'$ on angular scales between $\phi \approx (1-10)'$. While the results for different cosmological parameters are fairly close for cluster-normalised CDM, $\sigma(\phi)$ is larger by about a factor of two for CDM in an Einstein-de Sitter model normalised to $\sigma_8 = 1$. For the other cosmological models, the differences between different choices for the normalisation are less pronounced. The curves shown in Fig. 34 confirm the qualitative behaviour estimated in the previous section: The variance $\sigma(\phi)$ increases approximately linearly with ϕ as long as ϕ is small, and it gradually flattens off at angular scales $\phi \gtrsim \phi_g \approx 20'$.

In earlier chapters, we saw that non-linear density evolution has a large impact on weak gravitational lensing effects, e.g. on the magnification auto-correlation function $\xi_\mu(\phi)$. As mentioned before, this is not the case for the deflection-angle auto-correlation function $\xi_{\vec\alpha}(\phi)$ and the variance $\sigma(\phi)$ derived from it, because the filter function $F(l,\phi)$ relevant here suppresses small-scale density fluctuations for which the effect of non-linear evolution are strongest. Therefore, non-linear evolution is expected to have less impact here. Only on small angular scales ϕ, the filter function extends into the sufficiently non-linear regime. The curves in Fig. 34 confirm and quantify this expectation. Only on scales of $\phi \lesssim 10'$, the non-linear evolution does have some effect. Obviously, non-linear evolution increases the deflection-angle variance in a manner quite independent of cosmology. At angular scales $\phi \approx 1'$, the increase amounts to roughly a factor of two above the linear results.

9.6. Change of CMB Temperature Fluctuations

9.6.1. Summary of Previous Results

We are now ready to justify assumptions and approximations made earlier, and to quantify the impact of weak gravitational lensing on the Cosmic Microwave Background. The main assumptions were that (i) the deflection-angle variance $\sigma(\phi)$ is small, and (ii) $\sigma(\phi) \approx \varepsilon\phi$, with ε a (small) constant or a function slowly varying with ϕ. The results obtained in the previous section show that $\sigma(\phi)$ is typically about two orders of magnitude smaller than ϕ, confirming $\varepsilon \ll 1$. Likewise, Fig. 34 shows that the assumption $\sigma(\phi) \propto \phi$ is valid on angular scales smaller than the coherence scale for the deflection, $\phi \lesssim \phi_g \approx 20'$. As we have seen, this proportionality is a mere consequence of random coherent scattering of neighbouring light rays in the fluctuating potential field. For angles larger than ϕ_g, $\sigma(\phi)$ gradually levels off to become constant, so that the ratio between $\sigma(\phi)$ and ϕ tends to zero while ϕ increases further beyond ϕ_g. We can thus broadly summarise the numerical results on the deflection-angle variance by

$$\sigma(\phi) \approx \begin{cases} 0.01\,\phi & \text{for} \quad \phi \lesssim 20' \\ 0.7' & \text{for} \quad \phi \gg 20' \end{cases} , \tag{9.41}$$

which is valid for cluster-normalised CDM quite independent of the cosmological model; in particular, $\sigma(\phi) < 1' \approx 3 \times 10^{-4}$ radians for all ϕ.

9.6.2. Simplifications

Accordingly, the argument of the exponential in eq. (9.16) is a truly small number. Even for large $l \approx 10^3$, $l^2\sigma^2(\phi) \ll 1$. We can thus safely expand the exponential into a power series, keeping only the lowest-order terms. Then, eq. (9.16) simplifies to

$$\xi'_{\rm T}(\phi) = \xi_{\rm T}(\phi) - \sigma^2(\phi)\int_0^\infty \frac{l^3 \mathrm{d}l}{4\pi}\, P_{\rm T}(l)\, \mathrm{J}_0(l\phi)\,, \tag{9.42}$$

where we have used that the auto-correlation function $\xi_{\rm T}(\phi)$ is the Fourier transform of the power spectrum $P_{\rm T}(l)$. Employing again the approximate relation $\mathrm{J}_0{}''(x) \approx -\mathrm{J}_0(x)/2$ which holds for small x, we notice that

$$\int_0^\infty \frac{l^3 \mathrm{d}l}{4\pi}\, P_{\rm T}(l)\, \mathrm{J}_0(l\phi) \approx -\frac{\partial^2 \xi_{\rm T}(\phi)}{\partial\phi^2}\,. \tag{9.43}$$

We can introduce a typical angular scale ϕ_c for the CMB temperature fluctuations in the same manner as for light deflection in eq. (9.32). We define ϕ_c by

$$\phi_c^{-2} \equiv -\frac{1}{\xi_{\rm T}(0)}\, \frac{\partial^2 \xi_{\rm T}(\phi)}{\partial\phi^2}\bigg|_{\phi=0}\,, \tag{9.44}$$

so that, up to second order in ϕ, eq. (9.42) can be approximated as

$$\xi'_T(\phi) \approx \xi_T(\phi) - \frac{\sigma^2(\phi)}{\phi_c^2}\,\xi_T(0)\,. \tag{9.45}$$

We saw earlier that $\sigma(\phi) \approx \varepsilon\phi$ for $\phi \lesssim \phi_g$. Equation (9.45) can then further be simplified to read

$$\xi'_T(\phi) \approx \xi_T(\phi) - \varepsilon^2\,\xi_T(0)\,\frac{\phi^2}{\phi_c^2}\,. \tag{9.46}$$

In analogy to eq. (9.26), we can write the mean-square temperature fluctuations of the CMB between two beams separated by an angle ϕ as

$$\sigma_T^2(\phi) = \left\langle [\tau(\vec\theta) - \tau(\vec\theta + \vec\phi)]^2 \right\rangle = 2\,[\xi_T(0) - \xi_T(\phi)]\,. \tag{9.47}$$

Weak gravitational lensing changes this relative variance to

$$\sigma_T'^2 = 2\,\left[\xi'_T(0) - \xi'_T(\phi)\right]\,. \tag{9.48}$$

Using eq. (9.46), we see that the relative variance is *increased* by the amount

$$\Delta\sigma_T^2(\phi) = \sigma_T'^2(\phi) - \sigma_T^2(\phi) \approx \varepsilon^2\,\xi_T(0)\,\frac{\phi^2}{\phi_c^2}\,. \tag{9.49}$$

Now, the auto-correlation function at zero lag, $\xi_T(0)$, is the temperature-fluctuation variance, σ_T^2. Hence, we have for the *rms* change in the temperature variation

$$\left[\Delta\sigma_T^2(\phi)\right]^{1/2} = \varepsilon\,\sigma_T\,\frac{\phi}{\phi_c}\,. \tag{9.50}$$

Weak gravitational lensing thus changes the CMB temperature fluctuations only by a very small amount, of order $\varepsilon \approx 10^{-2}$ for $\phi \approx \phi_c$.

9.6.3. The Lensed CMB Power Spectrum

However, we saw in eq. (9.23) that the gravitationally lensed CMB power spectrum is smoothed compared to the intrinsic power spectrum. Modes on an angular scale ϕ are mixed with modes on angular scales $\phi \pm \sigma(\phi)$, i.e. the relative broadening $\delta\phi/\phi$ is of order $2\sigma(\phi)/\phi$. For $\phi \lesssim \phi_g \approx 20'$, this relative broadening is of order $2\varepsilon \approx 2\times10^{-2}$, while it becomes negligible for substantially larger scales because $\sigma(\phi)$ becomes constant. This effect is illustrated in Fig. 35, where we show the unlensed and lensed CMB power spectra for CDM in an Einstein-de Sitter universe.

The figure clearly shows that lensing smoothes the CMB power spectrum on small angular scales (large l), while it leaves large angular scales unaffected. Lensing effects become visible at $l \gtrsim 500$, corresponding to an angular scale of $\phi \lesssim (\pi/500)\,\mathrm{rad} \approx 20'$, corresponding to the scale where coherent gravitational light deflection sets in. An important effect of lensing is seen at the high-l tail of the power spectra, where the lensed power spectrum falls systematically above the unlensed one (Metcalf & Silk 1997). This happens because the Gaussian convolution kernel in eq. (9.23) becomes very broad for very large l, so that the lensed power spectrum at l' can be substantially increased by intrinsic power from significantly smaller l. In other words, lensing mixes power from larger angular scales into the otherwise feature-less damping tail of $P_T(l)$.

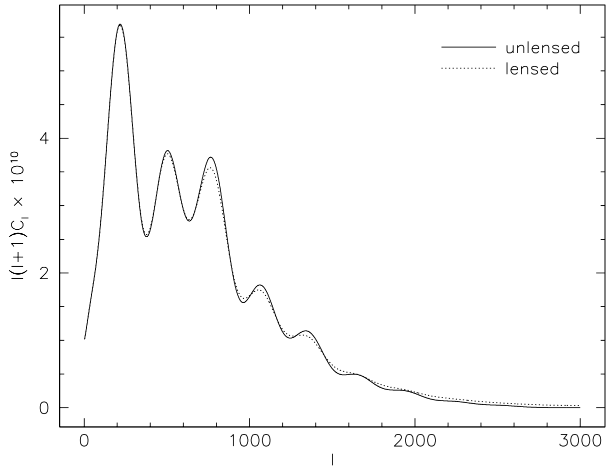

FIG. 35.—The CMB power spectrum coefficients $l(l+1)C_l$ are shown as a function of l. The solid line displays the intrinsic power spectrum, the dotted line the lensed power spectrum for an Einstein-de Sitter universe filled with cold dark matter. Evidently, lensing smoothes the spectrum at small angular scales (large l), while it has no visible effect on larger scales. The curves were produced with the `CMBfast` code, see Zaldarriaga & Seljak (1998).

9.7. Discussion

Several different approximations entered the preceding derivations. Firstly, the deflection-angle variance $\sigma(\phi)$ was generally assumed to be small, and for some expressions to be proportional to ϕ with a small constant of proportionality ε. The numerical results showed that the first assumption is very well satisfied, and the second assumption is valid for $\phi \lesssim \phi_g$, the latter being the coherence scale of gravitational light deflection.

We further assumed the deflection-angle field to be a Gaussian random field, the justification being that the deflecting matter distribution is also a Gaussian random field. While this fails to be exactly true at late stages of the cosmic evolution, we have seen that the resulting expression can also be obtained when $\sigma(\phi)$ is small and $\vec\alpha$ is not a Gaussian random field; hence, in practice this assumption is not a limitation of validity.

A final approximation consists in the Born approximation. This should also be a reasonable assumption at least in the case considered here, where we focus on *statistical* properties of light propagation. Even if the light rays would be bent considerably, the statistical properties of the potential gradient along their true trajectories are the same as along the approximated unperturbed rays.

Having found all the assumptions made well justifiable, we can conclude that the random walk of light rays towards the surface of recombination leads to smoothing of small-scale features in the CMB, while large-scale features remain unaffected. The border line between small and large angular scales is determined by the angular coherence scale of gravitational light deflection by large-scale matter distributions, which we found to be of order $\phi_g \approx 20'$, corresponding to $l_g = 2\pi\phi_g^{-1} \approx 1{,}000$. For the smallest angular scales, well into the damping tail of the intrinsic CMB power spectrum, this smoothing leads to a substantial redistribution of power, which causes the lensed CMB power spectrum to fall systematically above the unlensed one at $l \gtrsim 2000$, or $\phi \lesssim 2\pi l^{-1} \approx 10'$. Future space-bound CMB observations, e.g. by the Planck Surveyor satellite, will achieve angular resolutions of order $\gtrsim 5'$, so that the lensed regime of the CMB power spectrum will be well accessible. Highly accurate analyses of the data of such missions will therefore need to take lensing effects by large-scale structures into account.

One of the foremost goals of CMB observations is to derive cosmological parameters from the angular CMB power spectrum C_l. Unfortunately, there exists a parameter degeneracy in the sense that for any given set of cosmological parameters fitting a given CMB spectrum, a whole family of cosmological models can be found that will fit the spectrum (almost) equally well (Zaldarriaga et al. 1997). Metcalf & Silk (1998) and Stompor & Efstathiou (1999) showed that the rise in the damping-tail amplitude due to gravitational lensing of the CMB can be used to break this degeneracy once CMB observations with sufficiently high angular resolution become available.

We discussed in Sect. 4.2. how shapes of galaxy images can be quantified with the tensor Q_{ij} of second surface-brightness moments. Techniques for the reconstruction of the intervening projected matter distribution are then based on (complex) ellipticities constructed from Q_{ij}, e.g. the quantity χ defined in (4.4). Similar reconstruction techniques can be developed by constructing quantities comparable to χ from the CMB temperature fluctuations $\tau(\vec{\theta})$. Two such quantities were suggested in the literature, namely

$$\tau_{,1}^2 - \tau_{,2}^2 + 2\mathrm{i}\,\tau_{,1}\tau_{,2} \tag{9.51}$$

(Zaldarriaga & Seljak 1999) and

$$\tau_{,11} - \tau_{,22} + 2\mathrm{i}\,\tau_{,12} \tag{9.52}$$

(Bernardeau 1997). As usual, comma-preceded indices i denote differentiation with respect to θ_i.

The transformation of the tensor $\tau_{,i}\tau_{,j}$ between the lensed and unlensed CMB anisotropy distribution is mediated by the effective surface mass distribution $\kappa_{\rm eff}(\vec{\theta})$, defined as in (6.16) with w set to the comoving distance to the last-scattering surface at $z \sim 1000$. As shown by Zaldarriaga & Seljak (1999) and Seljak & Zaldarriaga (1999), one can reconstruct the power spectrum of the projected surface density from the observed statistical properties of $\tau_{,i}\tau_{,j}$; in fact, this power spectrum can be obtained either from the trace-part of this tensor, corresponding to $\kappa_{\rm eff}$ itself, or from the trace-free part, corresponding to the power spectrum of the shear which, as was shown earlier, is the same as that of $\kappa_{\rm eff}$. In contrast to similar studies based on the distortion of faint galaxies, the power-spectrum estimate from the CMB has the advantage that the redshift of the source is known. Furthermore, the power spectrum of the projected matter distribution can be obtained over a wide range of angular scales, corresponding to a wide range of spatial scales.

Even if the temperature anisotropies are intrinsically Gaussian, lensing will induce non-Gaussian features of the measured temperature map (e.g., Winitzki 1998). Hence, measurements of non-Gaussian temperature fluctuations must be interpreted with care. However, the lensing-induced non-Gaussian features on small angular scales are correlated with large-scale temperature gradients (Zaldarriaga 1999), thus providing a signature of the presence of lensing effects in the maps.

Lensing of the CMB can also be correlated with lensing effects of faint galaxies at lower redshift. The shear acting on these galaxies is part of the shear acting on the CMB, the difference being due to the different redshift of galaxies and the last scattering surface. Hence, one expects a correlation between these two shears (van Waerbeke et al. 1999a), as can be measured by correlating galaxy ellipticities with either of the quantities (9.51) or (9.52).

Finally, it is worth noting that gravitational lensing mixes different types of CMB polarisation (the "electric" and "magnetic", or E and B modes, respectively) and can thus create B-type polarisation even when only E-type polarisation is intrinsically present (Zaldarriaga & Seljak 1998). This effect, however, is fairly small in typical cosmological models and will only marginally affect future CMB polarisation measurements.

10. Summary and Outlook

We have summarised the basic ideas, theoretical developments, and first applications of weak gravitational lensing. In particular, we showed how the projected mass distribution of clusters can be reconstructed from the image distortion of background galaxies, using parameter-free methods, how the statistical mass distribution of galaxies can be obtained from galaxy-galaxy lensing, and how the larger-scale mass distribution in the Universe affects observations of galaxy shapes and fluxes of background sources, as well as the statistical properties of the CMB. Furthermore, weak lensing can be used to construct a mass-selected sample of clusters of galaxies, making use only of their tidal gravitational field which leaves an imprint on the image shapes of background galaxies. We have also discussed how the redshift distribution of these faint and distant galaxies can be derived from lensing itself, well beyond the magnitude limit which is currently available through spectroscopy.

Given that the first coherent image alignment of faint galaxies around foreground clusters was discovered only a decade ago (Fort et al. 1988; Tyson et al. 1990), the field of weak lensing has undergone a rapid evolution in the last few years, for three main reasons: (i) Theoreticians have recognised the potential power of this new tool for observational cosmology, and have developed specific statistical methods for extracting astrophysically and cosmologically relevant information from astronomical images. (ii) Parallel to that effort, observers have developed new observing strategies and image analysis software in order to minimise the influence of instrumental artefacts on the measured properties of faint images, and to control as much as possible the point-spread function of the resulting image. It is interesting to note that several image analysis methods, particularly aimed at shape measurements of very faint galaxies for weak gravitational lensing, have been developed by a coherent effort of theoreticians and observers (Bonnet & Mellier 1995; Kaiser et al. 1995; Luppino & Kaiser 1997; van Waerbeke et al. 1997; Kaiser 1999; Rhodes et al. 1999; Kuijken 1999), indicating the need for a close interaction between these two groups which is imposed by the research subject.

(iii) The third and perhaps major reason for the rapid evolution is the instrumental development that we are witnessing. Most spectacular was the refurbishment of the Hubble Space Telescope (HST) in Dec. 1993, after which this telescope produced astronomical images of angular resolution unprecedented in optical astronomy. These images have not only been of extreme importance for studying multiple images of galaxy-scale lens systems (where the angular separation is of order one arc second) and for detailed investigations of giant arcs and multiple galaxy images in clusters of galaxies, but also for several of the most interesting results of weak lensing. Owing to the lack of atmospheric smearing and the reduced sky background from space, the shape of fainter and smaller galaxy images can be measured on HST images, increasing the useful number density of background galaxies, and thus reducing the noise due to the intrinsic ellipticity distribution. Two of the most detailed mass maps of clusters have been derived from HST data (Seitz et al. 1996; Hoekstra et al. 1998), and all but one published results on galaxy-galaxy lensing are based on data taken with the HST. In parallel to this, the development of astronomical detectors has progressed quickly. The first weak-lensing observations were carried out with CCD detectors of $\sim 1{,}000^2$ pixels, covering a fairly small field-of-view. A few years ago, the first $(8\,\mathrm{K})^2$ camera was used for astronomical imaging. Its $30' \times 30'$ field can be used to map the mass distribution of clusters at large clustercentric radii, to investigate the potential presence of filaments between neighbouring clusters (Kaiser et al. 1998), or simply to obtain high-quality data on a large area. Such data will be useful for galaxy-galaxy lensing, the search for halos using their lensing properties only, for the investigation of cosmic shear, and for homogeneous galaxy number counts on large fields, needed to obtain a better quantification of the statistical association of AGNs with foreground galaxies.

It is easy to foresee that the instrumental developments will remain the driving force for this research field. By now, several large-format CCD cameras are either being built or already installed, including three cameras with a one square degree field-of-view and adequate sampling of the PSF (MEGAPRIME at CFHT, MEGACAM at the refurbished MMT, and OMEGA-CAM at the newly built VLT Support Telescope at Paranal; see the recent account of wide-field imaging instruments in Arnaboldi et al. 1998). Within a few years, more than a dozen 8- to 10-meter telescopes will be operating, and many of them will be extremely useful for obtaining high-quality astronomical images, due to their sensitivity, their imaging properties and the high quality of the astronomical site. In fact, at least one of them (SUBARU on Mauna Kea) will be equipped with a large-format CCD camera. One might hypothesise that weak gravitational lensing is one of the main science drivers to shift the emphasis of optical astronomers more towards imaging, in contrast to spectroscopy. For example, the VLT Support Telescope will be fully dedicated to imaging, and the fraction of time for wide-field imaging on several other major telescopes will be substantial. The Advanced Camera for Surveys (ACS) is planned to be installed on the HST in 2001. Its larger field-of-view, better sampling, and higher quantum efficiency – compared to the current imaging camera WFPC2 – promises to be particularly useful for weak lensing observations.

Even more ambitious ground-based imaging projects are currently under discussion. Funding has been secured for the VISTA project[17] of a 4 m telescope in Chile with a field-of-view of at least one square degree. Another 4 m *Dark Matter Telescope* with a substantially larger field-of-view (nine square degrees) is being discussed specifically for weak lensing. Kaiser et al. (1999a) proposed a new strategy for deep, wide-field optical imaging at high angular resolution, based on an array of relatively small ($D \sim 1.5$ m) telescopes with fast guiding capacity and a "rubber" focal plane.

Associated with this instrumental progress is the evolution of data-analysis capabilities. Whereas a small-format CCD image can be reduced and analysed 'by hand', this is no longer true for the large-format CCD images. Semi-automatic data-reduction pipelines will become necessary to keep up with the data flow. These pipelines, once properly developed and tested, can lead to a more 'objective' data analysis. In addition, specialised software, such as for the measurement of shapes of faint galaxies, can be implemented, together with tools which allow a correction for PSF anisotropies and smearing.

Staying with instrumental developments for one more moment, the two planned CMB satellite missions (MAP and Planck Surveyor) will provide maps of the CMB at an angular resolution and a signal-to-noise ratio which will most likely lead to the detection of lensing by the large-scale structure on the CMB, as described in Sect. 9.. Last but not least, the currently planned Next Generation Space Telescope (NGST, Kaldeich 1999), with a projected launch date of 2008, will provide a giant step in many fields of observational astronomy, not the least for weak lensing. It combines a large aperture (of order eight meters) with a position far from Earth to reduce sky background and with large-format imaging cameras. Even a relatively short exposure with the NGST, which will be optimised for observations in the near-infrared, will return images with a number density of several hundred background galaxies per square arc minute, for which a shape can be reliably measured; more accurate estimates are presently not feasible due to the large extrapolation into unknown territory. Comparing this number with the currently achievable number density in ground-based observations of about 30 per square arc minute, NGST will revolutionise this field.[18] In addition, the corresponding galaxies will be at much higher mean redshift than currently observable galaxy samples. Taken together, these two facts imply that one can detect massive halos at medium redshifts with only half the velocity dispersion currently necessary to detect them with ground-based data, or that the investigations of the mass distribution of halos can be extended to much higher redshifts than currently possible (see Schneider & Kneib 1998). The ACS on board HST will provide an encouraging hint of the increase in capabilities that NGST has to offer.

Progress may also come from somewhat unexpected directions. Whereas the Sloan Digital Sky Survey (SDSS; e.g. Szalay 1998) will be very shallow compared to more standard weak-lensing observations, its huge angular coverage may compensate for it (Stebbins 1996). The VLA-FIRST survey of radio sources (White et al. 1997) suffers from the sparsely populated radio sky, but this is also compensated by the huge sky coverage (Refregier et al. 1998). The use of both surveys for weak lensing will depend critically on the level down to which the systematics of the instrumental image distortion can be understood and compensated for.

Gravitational lensing has developed from a stand-alone research field into a versatile tool for observational cosmology, and this also applies to weak lensing. But, whereas the usefulness of strong lensing is widely accepted by the astronomical community, weak lensing is only beginning to reach that level of wide appreciation. Part of this difference in attitude may be due to the fact that strong-lensing effects, such as multiple images and giant arcs, can easily be seen on CCD images, and their interpretation can readily be explained also to the non-expert. In contrast, weak lensing effects are revealed only through thorough statistical analysis of the data. Furthermore, the number of people working on weak lensing on the level of data analysis is still quite small, and the methods used to extract shear from CCD data are rather intricate. However, the analysis of CMB data is certainly more complicated than weak lensing analyses, but there are more people in the latter field, who checked and cross-checked their results; also, more people implies that much more development has gone into this field. Therefore, what is needed in weak lensing is a detailed comparison of methods, preferably by several independent groups, analysing the same data sets, together with extensive work on simulated data to investigate down to which level a very weak shear can be extracted from them. Up to now, no show-stopper has been identified which prohibits the detection of shear at the sub-percent level.

Weak-lensing results and techniques will increasingly be combined with other methods. A few examples may suffice to illustrate this point. The analysis of galaxy clusters with (weak) lensing will be combined with results from X–ray measurements of the clusters and their Sunyaev-Zel'dovich decrement. Once

[17] see `http://www-star.qmw.ac.uk/~jpe/vista/`

[18] Whereas with the 8- to 10-meter class ground-based telescopes deeper images can be obtained, this does not drastically affect the 'useful' number density of faint galaxy images. Since fainter galaxies also tend to become smaller, and since a reliable shape estimate of a galaxy is feasible only if its size is not much smaller than the size of the seeing disk, very much deeper images from the ground will not yield much larger number densities of galaxy images which can be used for weak lensing.

these methods are better understood, in particular in terms of their systematics, the question will no longer be, "Are the masses derived with these methods in agreement?", but rather, "What can we learn from their comparison?" For instance, while lensing is insensitive to the distribution of matter along the line-of-sight, the X–ray emission is, and thus their combination provides information on the depth of the cluster (see, e.g., Zaroubi et al. 1998). Joint analyses of weak-lensing, X–ray and Sunyaev-Zel'dovich data on galaxy clusters promise to substantially improve determinations of the baryonic-matter fraction in clusters and of the structure and distribution of cluster-galaxy orbits.

One might expect that clusters will continue for some time to be main targets for weak-lensing studies. In addition to clusters selected by their emission, mass concentrations selected only by their weak-lensing properties shall be investigated in great detail, both with deeper images to obtain a more accurate measurement of the shear, and by X-ray, IR, sub-mm, and optical/IR multi-colour techniques. It would be spectacular, and of great cosmological significance, to find mass concentrations of exceedingly high mass-to-light ratio (well in excess of $1,000$ in solar units), and it is important to understand the distribution of M/L for clusters. A first example may have been found by Erben et al. (2000).

As mentioned before, weak lensing is able to constrain the redshift distribution of very faint objects which do not allow spectroscopic investigation. Thus, lensing can constrain extrapolations of the z-distribution, and the models for the redshift estimates obtained from multi-colour photometry ('photometric redshifts'). On the other hand, photometric redshifts will play an increasingly important role for weak lensing, as they will allow to increase the signal-to-noise ratio of local shear measurements. Furthermore, if source galaxies at increasingly higher redshifts are considered (as will be the case with the upcoming giant telescopes, cf. Clowe et al. 1998), the probability increases that more than one deflector lies between us and this distant screen of sources. To disentangle the corresponding projection effects, the dependence of the lensing strength on the lens and source redshift can be employed. Lenses at different redshifts cause different source-redshift dependences of the measured shear. Hence, photometric redshifts will play an increasingly important role for weak lensing. Whereas a fully three-dimensional mass distribution will probably be difficult to obtain using this relatively weak redshift dependence, a separation of the mass distribution into a small number of lens planes appears feasible.

Combining results from cosmic-shear measurements with the power spectrum of the cosmic density fluctuations as measured from the CMB will allow a sensitive test of the gravitational instability picture for structure formation. As was pointed out by Hu & Tegmark (1999), cosmic-shear measurements can substantially improve the accuracy of the determination of cosmological parameters from CMB experiments, in particular by breaking the degeneracies inherent in the latter (see also Metcalf & Silk 1998). The comparison between observed cosmic shear and theory will at least partly involve the increasingly detailed numerical simulations of cosmic structure evolution, from which predictions for lensing observations can directly be obtained. For example, if the dark matter halos in the numerical simulations are populated with galaxies, e.g., by using semi-empirical theories of galaxy evolution (Kauffmann et al. 1997), detailed prediction for galaxy-galaxy lensing can be derived and compared with observations, thus constraining these theories. The same numerical results will predict the relation between the measured shear and the galaxy distribution on larger scales, which can be compared with the observable correlation between these quantities to investigate the scale- and redshift dependence of the bias factor.

The range of applications of weak lensing will grow in parallel to the new instrumental developments. Keeping in mind that many discoveries in gravitational lensing were not really expected (like the existence of Einstein rings, or giant luminous arcs), it seems likely that the introduction and extensive use of wide-field cameras and giant telescopes will give rise to real surprises.

ACKNOWLEDGEMENTS

We are deeply indebted to Lindsay King and Shude Mao for their very careful reading of the manuscript and their numerous constructive remarks. Helpful and constructive comments were provided by Nick Kaiser, Guido Kruse, Yannick Mellier and Martin White.

We wish to thank Tereasa Brainerd, Klaus Dolag, Jean-Paul Kneib, Yannick Mellier and Robert Schmidt for allowing us to use their original illustrations and figures in this review.

Finally, and most importantly, we express our gratitude towards our past and present collaborators in the field of weak gravitational lensing. Whereas they cannot all be mentioned here, we particularly mention Thomas Erben, Bhuvnesh Jain, Yannick Mellier, Ramesh Narayan, Carolin Seitz, Stella Seitz and Ludovic van Waerbeke for long and fruitful collaborations which have made this review possible.

This work was supported in part by TMR Network "Gravitational Lensing: New Constraints on Cosmology and the Distribution of Dark Matter" of the EC under contract No. ERBFMRX-CT97-0172, and by the "Sonderforschungsbereich 375" on Astro-Particle Physics by the Deutsche Forschungsgemeinschaft.

REFERENCES

Abell, G. O., 1958, ApJS, 3, 211
Amendola, L., Frieman, J. A., & Waga, I., 1999, MNRAS, 309, 465
Antonucci, R., 1993, ARA&A, 31, 473
Arnaboldi, M., Capaccioli, M., Mancini, D., Rafanelli, P., Scaramella, R., & Sedmak, G., 1998, in *Wide-Field Surveys in Cosmology, 14th IAP meeting*, Editions Frontieres
Athreya, R., Mellier, Y., van Waerbeke, L., Fort, B., Pelló, R., & Dantel-Fort, M., 1999, preprint astro-ph/9909518
Audit, E. & Simmons, J. F. L., 1999, MNRAS, 303, 87
Babul, A. & Lee, M. H., 1991, MNRAS, 250, 407
Bacon, D., Refregier, A., & Ellis, R., 2000, preprint astro-ph/0003008
Bahcall, N. & Fan, X., 1998, ApJ, 504, 1
Bahcall, N. A., 1977, ARA&A, 15, 505
Bahcall, N. A., 1988, ARA&A, 26, 631
Banday, A. J., Górski, K. M., Bennett, C. L., et al., 1997, ApJ, 475, 393
Bardeen, J. M., 1980, PhRvD, 22, 1882
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S., 1986, ApJ, 304, 15
Bartelmann, M., 1995a, A&A, 303, 643
Bartelmann, M., 1995b, A&A, 298, 661
Bartelmann, M., Ehlers, J., & Schneider, P., 1993, A&A, 280, 351
Bartelmann, M., Huss, A., Colberg, J. M., Jenkins, A., & Pearce, F. R., 1998, A&A, 330, 1
Bartelmann, M. & Narayan, R., 1995, ApJ, 451, 60
Bartelmann, M., Narayan, R., Seitz, S., & Schneider, P., 1996, ApJ, 464, L115
Bartelmann, M. & Schneider, P., 1991, A&A, 248, 349
Bartelmann, M. & Schneider, P., 1992, A&A, 259, 413
Bartelmann, M. & Schneider, P., 1993a, A&A, 268, 1
Bartelmann, M. & Schneider, P., 1993b, A&A, 271, 421
Bartelmann, M. & Schneider, P., 1994, A&A, 284, 1
Bartelmann, M. & Schneider, P., 1999, A&A, 345, 17
Bartelmann, M., Schneider, P., & Hasinger, G., 1994, A&A, 290, 399
Bartsch, A., Schneider, P., & Bartelmann, M., 1997, A&A, 319, 375
Begelman, M. C., Blandford, R. D., & Rees, M. J., 1984, Rev. Mod. Phys., 56, 255
Benítez, N. & Martínez-González, E., 1995, ApJ, 448, L89
Benítez, N. & Sanz, J. L., 1999, ApJ, 525, L1
Bernardeau, F., 1997, A&A, 324, 15
Bernardeau, F., 1998, A&A, 338, 375
Bernardeau, F., van Waerbeke, L., & Mellier, Y., 1997, A&A, 322, 1
Bertin, E. & Arnouts, S., 1996, A&AS, 117, 393
Biggs, A. D., Browne, I. W. A., Helbig, P., Koopmans, L. V. E., Wilkinson, P. N., & Perley, R. A., 1999, MNRAS, 304, 349
Binggeli, B., Sandage, A., & Tammann, G. A., 1988, ARA&A, 26, 509
Birkinshaw, M. & Gull, S. F., 1983, Nat, 302, 315
Blanchard, A. & Schneider, J., 1987, A&A, 184, 1
Blandford, R. & Narayan, R., 1986, ApJ, 310, 568

Blandford, R. D., Cohen, J., Kundić, T., Brainerd, T., & Hogg, D., 1998, in *American Astronomical Society Meeting 192, #07.04*
Blandford, R. D. & Narayan, R., 1992, ARA&A, 30, 311
Blandford, R. D., Netzer, H., & Woltjer, L., 1990, *Active galactic nuclei*, Springer Verlag, Berlin
Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V., 1991, MNRAS, 251, 600
Bonnet, H., Fort, B., Kneib, J.-P., Mellier, Y., & Soucail, G., 1993, A&A, 280, L7
Bonnet, H. & Mellier, Y., 1995, A&A, 303, 331
Bonnet, H., Mellier, Y., & Fort, B., 1994, ApJ, 427, L83
Borgeest, U., v. Linde, J., & Refsdal, S., 1991, A&A, 251, L35
Börner, G., 1988, *The early universe*, Springer, Heidelberg
Börner, G. & Ehlers, J., 1988, A&A, 204, 1
Bouchet, F. R., Juszkiewicz, R., Colombi, S., & Pellat, R., 1992, ApJ, 394, L5
Bower, R. G. & Smail, I., 1997, MNRAS, 290, 292
Boyle, B. J., Shanks, T., & Peterson, B. A., 1988, MNRAS, 235, 935
Brainerd, T. G., Blandford, R. D., & Smail, I., 1996, ApJ, 466, 623
Brainerd, T. G., Smail, I., & Mould, J., 1995, MNRAS, 275, 781
Brainerd, T. G., Wright, C. O., Goldberg, D. M., & Villumsen, J. V., 1999, ApJ, 524, 9
Bridle, S. L., Hobson, M. P., Lasenby, A. N., & Saunders, R., 1998, MNRAS, 299, 895
Broadhurst, T., Ellis, R., Koo, D., & Szalay, A., 1990, Nat, 343, 726
Broadhurst, T., Huang, X., Frye, B., & Ellis, R. S., 1999, preprint astro-ph/9902316
Broadhurst, T. J., Ellis, R. S., & Shanks, T., 1988, MNRAS, 235, 827
Broadhurst, T. J., Taylor, A. N., & Peacock, J. A., 1995, ApJ, 438, 49
Buta, R., Mitra, S., De Vaucouleurs, G., & Corwin Jr., H. G., 1994, AJ, 107, 118
Carlberg, R. G., Yee, H. K. C., & Ellingson, E., 1994, ApJ, 437, 63
Carroll, S. M., Press, W. H., & Turner, E. L., 1992, ARA&A, 30, 499
Cayón, L., Martinez-Gonzáles, E., & Sanz, J. L., 1993a, ApJ, 413, 10
Cayón, L., Martinez-Gonzáles, E., & Sanz, J. L., 1993b, ApJ, 403, 471
Chiba, M. & Futamase, T., 1999, Prog. Theor. Phys. Supp., 133, 115
Clowe, D., Kaiser, N., Luppino, G. A., & Gioia, I., 2000, ApJ, in press
Clowe, D., Luppino, G. A., Kaiser, N., Henry, J. P., & Gioia, I. M., 1998, ApJ, 497, L61
Cohen, J. G., Blandford, R. D., Hogg, D. W., Pahre, M. A., & Shopbell, P., 1999, ApJ, 512, 30
Cole, S. & Efstathiou, G., 1989, MNRAS, 239, 195
Coles, P., Lucchin, F., Matarrese, S., & Moscardini, L., 1998, MNRAS, 300, 183
Colless, M., 1998, in S. Colombi, Y. Mellier, and B. Raban (eds.), *Wide-Field Surveys in Cosmology*, p. 77, Editions Frontieres
Colless, M., Ellis, R. S., Broadhurst, T. J., Taylor, K., & Peterson, B. A., 1993, MNRAS, 261, 19
Colless, M., Ellis, R. S., Shaw, G., & Taylor, K., 1991, MNRAS, 253, 686
Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A., 1995, AJ, 110, 2655
Cooray, A. R., 1999a, A&A, 348, 31
Cooray, A. R., 1999b, A&A, 348, 673
Crampton, D., Le Fèvre, O., Lilly, S. J., & Hammer, F., 1995, ApJ, 455, 96
Dalcanton, J. J., Canizares, C. R., Granados, A., Steidel, C. C., & Stocke, J. T., 1994, ApJ, 424, 550
Dashevskii, V. M., Zel'dovich, & Ya.B., 1965, SvA, 8, 854
Davis, M. & Peebles, P. J. E., 1983, ApJ, 267, 465
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, J. R., Buta, R. J., Paturel, G., & Fouque, P., 1991, *Third reference catalogue of bright galaxies, version 9*, Springer Verlag, New York
Deiser, N., 1995, *Statistische Methoden zur Bestimmung des Linsenparameters*, Diploma-thesis, LMU München
Dell'Antonio, I. P. & Tyson, J. A., 1996, ApJ, 473, L17
Dolag, K. & Bartelmann, M., 1997, MNRAS, 291, 446
Donahue, M., Voit, G. M., Gioia, I., Luppino, G., Hughes, J. P., & Stocke, J. T., 1998, ApJ, 502, 550
Dressler, A. & Gunn, J. E., 1992, ApJS, 78, 1
Dye, S. & Taylor, A., 1998, MNRAS, 300, L23
Dyer, C. C., 1976, MNRAS, 175, 429
Dyer, C. C. & Roder, R. C., 1973, ApJ, 180, L31
Ebbels, T. et al., 1998, MNRAS, 295, 75
Ebbels, T. M. D. et al., 1996, MNRAS, 281, L75
Eddington, A. S., 1920, *Space, time and gravitation*, Cambridge University Press, Cambridge
Efstathiou, G., 1996, in R. Schaeffer, J. Silk, and M. Spiro (eds.), *Cosmology and large-scale structure*, p. 133, Amsterdam, North Holland
Efstathiou, G., Ellis, R. S., & Peterson, B. A., 1988, MNRAS, 232, 431
Einstein, A. & Strauss, E. G., 1945, Rev. Mod. Phys., 17, 20
Eke, V. R., Cole, S., & Frenk, C. S., 1996, MNRAS, 282, 263
Ellis, R. S., 1997, ARA&A, 35, 389
Erben, T., 1997, *Die Bestimmung von Galaxieneigenschaften durch Galaxy-Galaxy-Lensing*, Diploma-thesis, TU München
Erben, T., van Waerbeke, L., Mellier, Y., et al., 2000, A&A, 355, 23
Etherington, I. M. H., 1933, Phil. Mag., 15, 761
Faber, S. M. & Gallagher, J. S., 1979, ARA&A, 17, 135
Faber, S. M. & Jackson, R. E., 1976, ApJ, 204, 668
Fahlman, G., Kaiser, N., Squires, G., & Woods, D., 1994, ApJ, 437, 56
Falco, E. E., Gorenstein, M. V., & Shapiro, I. I., 1985, ApJ, 289, L1
Falco, E. E., Kochanek, C. S., & Munoz, J. A., 1998, ApJ, 494, 47
Fan, X., Strauss, M. A., Schneider, D. P., Gunn, J. E., et al., 1999, AJ, 118, 1
Faraoni, V., 1993, A&A, 272, 385
Fischer, P., 1999, AJ, 117, 2024
Fischer, P., Bernstein, G., Rhee, G., & Tyson, J. A., 1997, AJ, 113, 521
Fischer, P., McKay, T., Sheldon, E., et al., 1999, preprint astro-ph/9912119
Fischer, P. & Tyson, J. A., 1997, AJ, 114, 14
Fischer, P., Tyson, J. A., Bernstein, G. M., & Guhathakurta, P., 1994, ApJ, 431, L71
Fixsen, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., & Wright, E. L., 1996, ApJ, 473, 576
Forman, W. & Jones, C., 1982, ARA&A, 20, 547
Fort, B. & Mellier, Y., 1994, A&AR, 5, 239
Fort, B., Mellier, Y., & Dantel-Fort, M., 1997, A&A, 321, 353
Fort, B., Mellier, Y., Dantel-Fort, M., Bonnet, H., & Kneib, J.-P., 1996, A&A, 310, 705
Fort, B., Prieur, J. L., Mathez, G., Mellier, Y., & Soucail, G., 1988, A&A, 200, L17
Franx, M., Illingworth, G. D., Kelson, D. D., van Dokkum, P. G., & Tran, K.-V., 1997, ApJ, 486, L75
Freedman, W. L., 1996, in N. Turok (ed.), *Critical dialogs in cosmology, proc. Princeton 250th anniversary*, p. 92, Singapore: World Scientific
Friedmann, A., 1922, Z. Phys., 10, 377
Friedmann, A., 1924, Z. Phys., 21, 326
Frieman, J. A., 1996, Comments on Astrophysics
Fry, J. N., 1984, ApJ, 279, 499
Frye, B. & Broadhurst, T., 1998, ApJ, 499, L115
Fugmann, W., 1990, A&A, 240, 11
Fukugita, M., Futamase, T., Kasai, M., & Turner, E. L., 1992, ApJ, 393, 3
Futamase, T., 1989, MNRAS, 237, 187
Futamase, T. & Sasaki, M., 1989, Phys. Rev. D, 40, 2502
Gautret, L., Fort, B., & Mellier, Y., 2000, A&A, 353, 10
Gaztanaga, E. & Bernardeau, F., 1998, A&A, 331, 829
Geiger, B. & Schneider, P., 1998, MNRAS, 295, 497
Geiger, B. & Schneider, P., 1999, MNRAS, 302, 118
Giovanelli, R. & Haynes, M. P., 1991, ARA&A, 29, 499
Goroff, M. H., Grinstein, B., Rey, S.-J., & Wise, M. B., 1986, ApJ, 311, 6
Gould, A., 1995, ApJ, 440, 510
Gradshteyn, I. S. & Ryzhik, I. M., 1994, *Table of integrals, series and products*, Academic Press, 5th edition
Griffiths, R. E., Casertano, S., Im, M., & Ratnatunga, K. U., 1996, MNRAS, 282, 1159
Gunn, J. E., 1967, ApJ, 150, 737
Gunn, J. E. & Knapp, G. R., 1993, in *ASP Conf. Ser. 43: Sky Surveys. Protostars to Protogalaxies*, p. 267
Gurvits, L. I. & I.G., M., 1986, Nat, 324, 349
Gwyn, S. D. J. & Hartwick, F. D. A., 1996, ApJ, 468, L77
Hamilton, A. J. S., Matthews, A., Kumar, P., & Lu, E., 1991, ApJ, 374, L1
Hamuy, M., Phillips, M. M., Suntzeff, N. B., Schommer, R. A., Maza, J., & Aviles, R., 1996, AJ, 112, 2391
Harrison, E. R., 1970, PhRvD, 1, 2726
Hartwick, F. D. A. & Schade, D., 1990, ARA&A, 28, 437
Hattori, M., Kneib, J.-P., & Makino, N., 1999, Prog. Theor. Phys. Supp., 133, 1
Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F., Langston, G. I., & Lawrence, C. R., 1988, Nat, 333, 537
Hoekstra, H., Franx, M., & Kuijken, K., 1999, preprint astro-ph/9911106
Hoekstra, H., Franx, M., & Kuijken, K., 2000, ApJ, 532, 88
Hoekstra, H., Franx, M., Kuijken, K., & Squires, G., 1998, ApJ, 504, 636
Hogg, D. W., Cohen, J. G., Blandford, R., et al., 1998, AJ, 115, 1418
Holz, D. E., 1998, ApJ, 506, L1
Hu, W., 1995, *Ph.D. thesis*, Univ. of California, Berkeley
Hu, W., 1999, ApJ, 522, L21
Hu, W. & Tegmark, M., 1999, ApJ, 514, L65
Huchra, J., Gorenstein, M., Kent, S., Shapiro, I., Smith, G., Horine, E., & Perley, R., 1985, AJ, 90, 691
Hudson, M. J., Gwyn, S. D. J., Dahle, H., & Kaiser, N., 1998, ApJ, 503, 531
Hui, L., 1999a, ApJ, 519, L9
Hui, L., 1999b, ApJ, 519, L9
Jacobs, M. W., Linder, E. V., & Wagoner, R. V., 1993, Phys. Rev. D, 48, 4623
Jain, B., Connolly, A., Szalay, A., et al., 1999, preprint
Jain, B., Mo, H. J., & White, S. D. M., 1995, MNRAS, 276, L25
Jain, B. & Seljak, U., 1997, ApJ, 484, 560
Jain, B., Seljak, U., & White, S., 2000, ApJ, 530, 547
Jain, B. & van Waerbeke, L., 2000, ApJ, 530, L1
Jaroszynski, M., 1991, MNRAS, 249, 430
Jaroszynski, M., Park, C., Paczynski, B., & Gott, J. R., 1990, ApJ, 365, 22
Jarvis, J. F. & Tyson, J. A., 1981, AJ, 86, 476
Joffre, M., Fischer, P., Frieman, J., et al., 1999, preprint astro-ph/9911285
Kaiser, N., 1984, ApJ, 284, L9
Kaiser, N., 1992, ApJ, 388, 272
Kaiser, N., 1995, ApJ, 439, L1
Kaiser, N., 1998, ApJ, 498, 26
Kaiser, N., 1999, ApJ, submitted, preprint astro-ph/9904003
Kaiser, N. & Squires, G., 1993, ApJ, 404, 441
Kaiser, N., Squires, G., & Broadhurst, T., 1995, ApJ, 449, 460
Kaiser, N., Squires, G., Fahlman, G., & Woods, D., 1994, in *Clusters of galaxies, proc. XIVth Moriond astrophysics meeting, Méribel, France, 1994*, p. 269
Kaiser, N. & Stebbins, A., 1984, Nat, 310, 391
Kaiser, N., Tonry, J. L., & Luppino, G. A., 1999a, PASP, submitted, preprint astro-ph/9912181
Kaiser, N., Wilson, G., Luppino, G., & Dahle, H., 1999b, preprint astro-ph/9907229

Kaiser, N., Wilson, G., Luppino, G., Kofman, L., Gioia, I., Metzger, M., & Dahle, H., 1998, preprint astro-ph/9809268
Kaiser, N., Wilson, G., & Luppino, G. A., 2000, preprint astro-ph/0003338
Kaldeich, B. (ed.), 1999, *The Next Generation Space Telescope: Science Drivers and Technological Challenges*, 34th Liège Astrophysics Colloquium, ESA
Kashlinsky, A., 1988, ApJ, 331, L1
Kassiola, A., Kovner, I., & Fort, B., 1992, ApJ, 400, 41
Kauffmann, G., Nusser, A., & Steinmetz, M., 1997, MNRAS, 286, 795
Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M., 1996, ApJ, 471, 643
Kneib, J. P. et al., 1995, A&A, 303, 27
Kneib, J.-P., Mathez, G., Fort, B., Mellier, Y., Soucail, G., & Longaretti, P.-Y., 1994, A&A, 286, 701
Kneib, J.-P., Mellier, Y., Fort, B., & Mathez, G., 1993, A&A, 273, 367
Kochanek, C. S., 1990, MNRAS, 247, 135
Kochanek, C. S., 1993, ApJ, 419, 12
Kochanek, C. S., 1996, ApJ, 466, 638
Koo, D. C. & Kron, R. G., 1992, ARA&A, 30, 613
Kovner, I. & Milgrom, M., 1987, ApJ, 321, L113
Kristian, J. & Sachs, R. K., 1966, ApJ, 143, 379
Krolik, J. H., 1999, *Active galactic nuclei: from the central black hole to the galactic environment*, Princeton University Press, Princeton, NJ
Kruse, G. & Schneider, P., 1999a, preprint, astro-ph/9904192
Kruse, G. & Schneider, P., 1999b, MNRAS, 302, 821
Kühr, H., Witzel, A., Pauliny-Toth, I. I. K., & Nauber, U., 1981, A&AS, 45, 367
Kuijken, K., 1999, A&A, 352, 355
Kundić, T. et al., 1997, ApJ, 482, 75
Lacey, C. & Cole, S., 1993, MNRAS, 262, 627
Lacey, C. & Cole, S., 1994, MNRAS, 271, 676
Lifshitz, E. M., 1946, J. Phys. USSR, 10, 116
Lilly, S. J., 1993, ApJ, 411, 501
Lilly, S. J., Cowie, L. L., & Gardner, J. P., 1991, ApJ, 369, 79
Lilly, S. J., Le Fèvre, O., Crampton, D., Hammer, F., & Tresse, L., 1995, ApJ, 455, 50
Limber, D. N., 1953, ApJ, 117, 134
Linder, E. V., 1988, A&A, 206, 199
Linder, E. V., 1990a, MNRAS, 243, 353
Linder, E. V., 1990b, MNRAS, 243, 362
Lombardi, M. & Bertin, G., 1998a, A&A, 330, 791
Lombardi, M. & Bertin, G., 1998b, A&A, 335, 1
Lombardi, M. & Bertin, G., 1999a, A&A, 348, 38
Lombardi, M. & Bertin, G., 1999b, A&A, 342, 337
Lombardi, M. & Bertin, G., 1999c, A&A, 342, 337
Loveday, J. & Pier, J., 1998, in S. Colombi, Y. Mellier, and B. Raban (eds.), *Wide-Field Surveys in Cosmology*, p. 317, Editions Frontieres
Lucy, L. B., 1994, A&A, 289, 983
Luppino, G. A. & Kaiser, N., 1997, ApJ, 475, 20
Lynds, R. & Petrosian, V., 1986, BAAS, 18, 1014
Lynds, R. & Petrosian, V., 1989, ApJ, 336, 1
Mao, S., 2000, in T. G. Brainerd and C. S. Kochanek (eds.), *Gravitational Lensing: Recent Progress and Future Goals*
Mao, S. & Kochanek, C. S., 1994, MNRAS, 268, 569
Maoz, D. & Rix, H.-W., 1993, ApJ, 416, 425
Martínez-González, E., Sanz, J. L., & Silk, J., 1990, ApJ, 355, L5
Marzke, R. O., Geller, M. J., Huchra, J. P., & Corwin, H. G., J., 1994a, AJ, 108, 437
Marzke, R. O., Huchra, J. P., & Geller, M. J., 1994b, ApJ, 428, 43
Mellier, Y., 1999, ARA&A, 37, 127
Mellier, Y., Dantel-Fort, M., Fort, B., & Bonnet, H., 1994, A&A, 289, L15
Mellier, Y., Fort, B., Soucail, G., Mathez, G., & Cailloux, M., 1991, ApJ, 380, 334
Metcalf, R., 1999, MNRAS, 305, 746
Metcalf, R. B. & Silk, J., 1997, ApJ, 489, 1
Metcalf, R. B. & Silk, J., 1998, ApJ, 492, L1
Metcalf, R. B. & Silk, J., 1999, ApJ, 519, L1
Metzler, C. A., White, M., Norman, M., & Loken, C., 1999, ApJ, 520, L9
Miralda-Escudé, J., 1991, ApJ, 380, 1
Miralda-Escude, J., 1991, ApJ, 370, 1
Misner, C. W., Thorne, K. S., & Wheeler, J. A., 1973, *Gravitation*, Freeman, New York
Moessner, R. & Jain, B., 1998, MNRAS, 294, L18
Moessner, R., Jain, B., & Villumsen, J. V., 1998, MNRAS, 294, 291
Mould, J., Blandford, R., Villumsen, J., Brainerd, T., Smail, I., Small, T., & Kells, W., 1994, MNRAS, 271, 31
Munshi, D. & Coles, P., 2000, MNRAS, 313, 148
Munshi, D. & Jain, B., 1999a, preprint astro-ph/9912330
Munshi, D. & Jain, B., 1999b, preprint
Naim, A., Lahav, O., Buta, R. J., et al., 1995a, MNRAS, 274, 1107
Naim, A., Lahav, O., Sodre Jr., L., & Storrie-Lombardi, M. C., 1995b, MNRAS, 275, 567
Narayan, R. & Bartelmann, M., 1999, in A. Dekel and J. P. Ostriker (eds.), *Formation of structure in the universe*, p. 360, Cambridge University Press
Narayan, R. & Nityananda, R., 1986, ARA&A, 24, 127
Narayan, R. & Wallington, S., 1993, in *Gravitational lenses in the universe, 31st Liège International Astrophysical Colloquium, June 1993*, p. 217
Natarajan, P., J.-P., K., Smail, I., & Ellis, R. S., 1998, ApJ, 499, 600
Natarajan, P. & Kneib, J.-P., 1997, MNRAS, 287, 833
Navarro, J., Frenk, C., & White, S., 1996a, ApJ, 462, 563
Navarro, J., Frenk, C., & White, S., 1997, ApJ, 486, 493
Navarro, J. F., Frenk, C. S., & White, S. D. M., 1996b, ApJ, 462, 563
Norman, D. J. & Impey, C. D., 1999, AJ, 118, 613
Norman, D. J. & Williams, L. L. R., 1999, ApJ, submitted, preprint astro-ph/9908177
Nottale, L., 1984, MNRAS, 206, 713
Ohanian, H. C., 1983, ApJ, 271, 551
Paczyński, B., 1987, Nat, 325, 572
Paczyński, B., 1996, ARA&A, 34, 419
Padmanabhan, T., 1993, *Structure formation in the Universe*, Cambridge University Press
Peacock, J. A., 1997, MNRAS, 284, 885
Peacock, J. A., 1999, *Cosmological physics*, Cambridge University Press
Peacock, J. A. & Dodds, S. J., 1996, MNRAS, 280, L19
Peebles, P. J. E., 1980, *The large-scale structure of the Universe*, Princeton University Press, Princeton, NJ
Peebles, P. J. E., 1993, *Principles of physical cosmology*, Princeton University Press
Peebles, P. J. E. & Yu, J. T., 1970, ApJ, 162, 815
Pei, Y. C., 1995, ApJ, 438, 623
Pelló, R. et al., 1999a, A&A, 346, 359
Pelló, R., Kneib, J. P., Le Borgne, J. F., et al., 1999b, A&A, 346, 359
Pello, R., Le Borgne, J.-F., Soucail, G., Mellier, Y., & Sanahuja, B., 1991, ApJ, 366, 405
Penzias, A. A. & Wilson, R. W., 1965, ApJ, 142, 419
Peterson, B. M., 1997, *An introduction to active galactic nuclei*, Cambridge University Press, Cambridge
Phillips, M. M., 1993, ApJ, 413, L105
Press, W. & Schechter, P., 1974, ApJ, 187, 425
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T., 1986, *Numerical Recipes*, Cambridge University Press, Cambridge
Press, W. H. & Gunn, J. E., 1973, ApJ, 185, 397
Pyne, T. & Birkinshaw, M., 1996, ApJ, 458, 46
Reblinsky, K. & Bartelmann, M., 1999a, A&A, 345, 1
Reblinsky, K. & Bartelmann, M., 1999b, A&A, 345, 1
Reblinsky, K., Kruse, G., Jain, B., & Schneider, P., 1999, A&A, 351, 815
Rees, M. J., 1984, ARA&A, 22, 471
Rees, M. J. & Sciama, D. W., 1968, Nat, 217, 511
Refregier, A., Brown, S. T., Kamionkowski, M., Helfand, D. J., Cress, C. M., Babul, A., Becker, R., & White, R. L., 1998, in *Wide Field Surveys in Cosmology, 14th IAP meeting held May 26-30, 1998, Paris. Publisher: Editions Frontieres*, p. 209
Refsdal, S., 1964, MNRAS, 128, 307
Refsdal, S. & Surdej, J., 1994, Rep. Prog. Phys., 56, 117
Rhodes, J., Refregier, A., & Groth, E., 1999, ApJ, submitted, preprint astro-ph/9905090
Richstone, D., Loeb, A., & Turner, E. L., 1992, ApJ, 393, 477
Riess, A. G., Filippenko, A. V., Challis, P., et al., 1998, AJ, 116, 1009
Riess, A. G., Press, W. H., & Kirshner, R. P., 1995, ApJ, 438, L17
Riess, A. G., Press, W. H., & Kirshner, R. P., 1996, ApJ, 473, 88
Rix, H.-W., Schneider, D. P., & Bahcall, J. N., 1992, AJ, 104, 959
Robertson, H. P., 1935, ApJ, 82, 284
Rodrigues-Williams, L. L. & Hawkins, M. R. S., 1995, in *Dark Matter, AIP conference proceedings 336, College Park, MD, 1994*, p. 331
Rodrigues-Williams, L. L. & Hogan, C. J., 1994, AJ, 107, 451
Rood, H. J., 1981, RPPh, 44, 1077
Roulet, E. & Mollerach, S., 1997, Phys. Rep., 279, 67
Sachs, R. K. & Wolfe, A. M., 1967, ApJ, 147, 73
Sanz, J. L., Martínez-González, E., & Benítez, N., 1997, MNRAS, 291, 418
Sarazin, C. L., 1986, RvMP, 58, 1
Sasaki, M., 1989, MNRAS, 240, 415
Schechter, P., 1976, ApJ, 203, 297
Schechter, P. L. et al., 1997, ApJ, 475, L85
Schindler, S. & Wambsganss, J., 1996, A&A, 313, 113
Schindler, S. & Wambsganss, J., 1997, A&A, 322, 66
Schneider, P., 1985, A&A, 143, 413
Schneider, P., 1993, A&A, 279, 1
Schneider, P., 1995, A&A, 302, 639
Schneider, P., 1996a, IAU Symposia, 168, 209
Schneider, P., 1996b, MNRAS, 283, 837
Schneider, P., 1998, ApJ, 498, 43
Schneider, P. & Bartelmann, M., 1997, MNRAS, 286, 696
Schneider, P., Ehlers, J., & Falco, E. E., 1992, *Gravitational Lenses*, Springer Verlag, Heidelberg
Schneider, P., King, L., & Erben, T., 2000, A&A, 353, 41
Schneider, P. & Kneib, J.-P., 1998, in *The Next Generation Space Telescope: Science Drivers and Technological Challenges, 34th Liège Astrophysics Colloquium, June 1998*, p. 89
Schneider, P. & Morales-Merino, C., 2000, in preparation
Schneider, P. & Rix, H.-W., 1997, ApJ, 474, 25
Schneider, P. & Seitz, C., 1995, A&A, 294, 411
Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G., 1998a, MNRAS, 296, 873
Schneider, P., van Waerbeke, L., Mellier, Y., Jain, B., Seitz, S., & Fort, B., 1998b, A&A, 333, 767
Schneider, P. & Wagoner, R. V., 1987, ApJ, 314, 154
Schramm, T. & Kayser, R., 1995, A&A, 299, 1
Scoccimarro, R. & Frieman, J., 1999, ApJ, 520, 35
Seitz, C., Kneib, J.-P., Schneider, P., & Seitz, S., 1996, A&A, 314, 707
Seitz, C. & Schneider, P., 1995a, A&A, 297, 287
Seitz, C. & Schneider, P., 1997, A&A, 318, 687

Seitz, S., Collodel, L., Pirzkal, N., et al., 1998a, in *Wide field surveys in cosmology*, p. 203, Editions Frontieres
Seitz, S., Saglia, R., Bender, R., Hopp, U., Belloni, P., & Ziegler, B., 1998b, MNRAS, 298, 325
Seitz, S. & Schneider, P., 1992, A&A, 265, 1
Seitz, S. & Schneider, P., 1995b, A&A, 302, 9
Seitz, S. & Schneider, P., 1996, A&A, 305, 383
Seitz, S. & Schneider, P., 1998, astro-ph/9802051
Seitz, S., Schneider, P., & Bartelmann, M., 1998c, A&A, 337, 325
Seitz, S., Schneider, P., & Ehlers, J., 1994, Class. Quantum Grav., 11, 2345
Seldner, M., Siebers, B., Groth, E. J., & Peebles, P. J. E., 1977, AJ, 82, 249
Seljak, U., 1994, ApJ, 436, 509
Seljak, U., 1996, ApJ, 463, 1
Seljak, U., 1998, ApJ, 506, 64
Seljak, U. & Holz, D. E., 1999, A&A, 351, L10
Seljak, U. & Zaldarriaga, M., 1999, Phys. Rev. Lett., 82, 2636
Smail, I., Ellis, R. S., & Fitchett, M. J., 1994, MNRAS, 270, 245
Smail, I., Ellis, R. S., Fitchett, M. J., & Edge, A. C., 1995a, MNRAS, 273, 277
Smail, I., Hogg, D. W., Yan, L., & Cohen, J. G., 1995b, ApJ, 449, L105
Smoot, G. F., 1997, Lectures held at Strasbourg NATO school on the CMB and cosmology, preprint astro-ph/9705135
Smoot, G. F., Bennett, C. L., Kogut, A., et al., 1992, ApJ, 396, L1
Soucail, G., Fort, B., Mellier, Y., & Picat, J. P., 1987a, A&A, 172, L14
Soucail, G., Mellier, Y., Fort, B., Hammer, F., & Mathez, G., 1987b, A&A, 184, L7
Soucail, G., Mellier, Y., Fort, B., Mathez, G., & Cailloux, M., 1988, A&A, 191, L19
Soucail, G., Ota, N., Böhringer, H., et al., 2000, A&A, 355, 433
Squires, G. et al., 1996a, ApJ, 461, 572
Squires, G. et al., 1997, ApJ, 482, 648
Squires, G. & Kaiser, N., 1996, ApJ, 473, 65
Squires, G., Kaiser, N., Fahlman, G., Babul, A., & Woods, D., 1996b, ApJ, 469, 73
Stebbins, A., 1996, American Astronomical Society Meeting, 189, 8207
Stebbins, A., McKay, T., & Frieman, J., 1996, in C. Kochanek and J. Hewitt (eds.), *Astrophysical applications of gravitational lensing*, p. 75, Kluwer Academic Publishers; Dordrecht
Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M., 1998, ApJ, 492, 428
Stickel, M. & Kühr, H., 1993, A&AS, 100, 395
Stickel, M., Kühr, H., & Fried, J. W., 1993, A&AS, 97, 483
Stocke, J. T., Morris, S. L., Gioia, I. M., Maccacaro, T., Schild, R., Wolter, A., Fleming, T. A., & Henry, J. P., 1991, ApJS, 76, 813
Stompor, R. & Efstathiou, G., 1999, MNRAS, 302, 735
Strauss, M. A. & Willick, J. A., 1995, Phys. Rep., 261, 271
Surpi, G. C. & Harari, D. D., 1999, ApJ, 515, 455
Szalay, A., 1998, American Astronomical Society Meeting, 192, 6405
Taylor, A. N., Dye, S., Broadhurst, T. J., Berí tez, N., & van Kampen, E., 1998, ApJ, 501, 539
Tomita, K., 1989, PhRvD, 40, 3821
Trager, S. C., Faber, S. M., Dressler, A., & Oemler, A., 1997, ApJ, 485, 92
Tran, K.-V. H., Kelson, D. D., van Dokkum, P., Franx, M., Illingworth, G. D., & Magee, D., 1999, ApJ, 522, 39
Tully, R. B. & Fisher, J. R., 1977, A&A, 54, 661
Turner, E. L., Ostriker, J. P., & Gott, J. R., 1984, ApJ, 284, 1
Tyson, J. A., 1986, AJ, 92, 691
Tyson, J. A., 1988, AJ, 96, 1
Tyson, J. A. & Seitzer, P., 1988, ApJ, 335, 552
Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills Jr., A. P., 1984, ApJ, 281, L59
Tyson, J. A., Valdes, F., & Wenk, R. A., 1990, ApJ, 349, L1
Umetsu, K., Tada, M., & Futamase, T., 1999, Prog. Theor. Phys. Supp., 133, 53
Valageas, P., 1999, preprint astro-ph/9911336
Valageas, P., 2000, A&A, 354, 767
Valdes, F., Tyson, J. A., & Jarvis, J. F., 1983, ApJ, 271, 431
van Haarlem, M. P. & van der Hulst, J. M. (eds.), 1999, *Scientific Imperatives at centimeter and meter Wavelengths*
van Kampen, E., 1998, MNRAS, 301, 389
van Waerbeke, L., 1998, A&A, 334, 1
van Waerbeke, L., 2000, MNRAS, 313, 524
van Waerbeke, L., Bernardeau, F., & Benabed, K., 1999a, ApJ, submitted; preprint astro-ph/9910366
van Waerbeke, L., Bernardeau, F., & Mellier, Y., 1999b, A&A, 342, 15
van Waerbeke, L., Mellier, Y., Erben, T., et al., 2000, preprint astro-ph/0002500
van Waerbeke, L., Mellier, Y., Schneider, P., Fort, B., & Mathez, G., 1997, A&A, 317, 303
Viana, P. T. P. & Liddle, A. R., 1996, MNRAS, 281, 323
Villumsen, J., Freudling, W., & DaCosta, L., 1997, ApJ, 481, 578
Voges, W., 1992, in *Environment observation and climate modelling through international space projects. Space sciences with particular emphasis on high-energy astrophysics*, p. 9
Walker, A. G., 1935, Quant. Journ. Math. Oxford Sci., 6, 81
Wallington, S., Kochanek, C. S., & Koo, D. C., 1995, ApJ, 441, 58
Walsh, D., Carswell, R. F., & Weymann, R. J., 1979, Nat, 279, 381
Wambsganss, J., Cen, R., & Ostriker, J. P., 1998, ApJ, 494, 29
Wambsganss, J., Cen, R., Xu, G., & Ostriker, J. P., 1997, ApJ, 475, L81
Warren, S. J. & Hewett, P. C., 1990, RPPh, 53, 1095
Watanabe, K. & Tomita, K., 1991, ApJ, 370, 481
Waxman, E. & Miralda-Escudé, J., 1995, ApJ, 451, 451
Weedman, D. W., 1986, *Quasar astronomy*, Cambridge University Press, Cambridge
Weinberg, S., 1972, *Gravitation and cosmology*, Wiley, New York
Weinberg, S., 1976, ApJ, 208, L1
White, M. & Hu, W., 1999, preprint astro-ph/9909165
White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D., 1997, ApJ, 475, 479
White, S. D. M., Davis, M., Efstathiou, G., & Frenk, C. S., 1987, Nat, 330, 451
White, S. D. M., Efstathiou, G., & Frenk, C. S., 1993, MNRAS, 262, 1023
Williams, L. L. R. & Irwin, M., 1998, MNRAS, 298, 378
Williams, R. E., Blacker, B., Dickinson, M., et al., 1996, AJ, 112, 1335
Wilson, G., Cole, S., & Frenk, C. S., 1996, MNRAS, 280, 199
Winitzki, S., 1998, preprint astro-ph/9806105
Wittman, D., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G., 2000, preprint astro-ph/0003014
Wu, X.-P., 1996, Fund. Cosm. Phys. 17, 1
Wu, X.-P. & Fang, L.-Z., 1996, ApJ, 461, L5
Wu, X.-P. & Han, J., 1995, MNRAS, 272, 705
Zaldarriaga, M., 1999, preprint astro-ph/9910498
Zaldarriaga, M. & Seljak, U., 1998, Phys. Rev. D, 58, 023003
Zaldarriaga, M. & Seljak, U., 1999, Phys. Rev. D, 59, 123507
Zaldarriaga, M., Spergel, D. N., & Seljak, U., 1997, ApJ, 488, 1
Zaritsky, D. & White, S. D. M., 1994, ApJ, 435, 599
Zaroubi, S., Squires, G., Hoffman, Y., & Silk, J., 1998, ApJ, 500, L87
Zel'dovich & Ya.B., 1964, SvA, 8, 13
Zel'dovich, Y. B., 1972, MNRAS, 160, p1
Zensus, J. A. & Pearson, T. J., 1987, *Superluminal radio sources*, Cambridge University Press, Cambridge
Zwicky, F., 1933, Helv. Phys. Acta, 6, 110
Zwicky, F., Herzog, E., & Wild, P., 1968, *Catalogue of galaxies and of clusters of galaxies*, California Institute of Technology, Pasadena

CONTENTS