Weak Lensing Flexion Study

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Er, Xinzhong

aus

Tianjin, China

Bonn 2010

Contents

0.1 Introduction

As early as thousands of years ago, people already started studying the world of heaven and the universe. It is a long history involving science, philosophy, and religion, like that the ancient Chinese philosophers term the universe by 'Taiji', which means 'supreme ultimate'. Only in recent decades, people gained solid knowledge of physics, which allowed to really explore the universe in a scientific way. In the 1920s, Hubble found that galaxies are receding, which is an evidence for an expansion of the universe, and is considered as the first cosmological observation. Since then, the understanding of the universe has changed and grown rapidly with the progress of both physical theory and telescope technology. Our understanding of the universe is contained in some model of cosmology. The one currently popular is the ΛCDM cosmology, which is favored by most cosmological observations, such as large volume galaxy survey (SDSS)(Tegmark et al. 2004), cosmic microwave background (WMAP) (Komatsu et al. 2009) and cosmic shear (Fu et al. 2008). These measurements are accurate enough to constrain the parameterizations of the cosmological model. With these parameters, some properties of the universe on large-scales can be established, e.g. how the structure in the universe formed or whether the universe is flat.

Despite these successes, we have little knowledge about the fundamental physics of the universe. In the ΛCDM model, about 70% of the matter-energy in the universe consists of Dark Energy and 25% of Cold Dark Matter, whose nature are still unknown. Furthermore, the initial conditions for structure formation are given by some inflation scenarios, whose predictions still need strict confirmation. Finally structure formation in the universe is not fully understood on small scales (smaller than the Mpc scale), where non-linear processes play an important role. Precise measurements of the matter distribution in the universe will help to understand some of these issues.

Weak gravitational lensing provides an excellent tool to test cosmological models. The light of distant galaxies is deflected by the tidal gravitational field of the intervening matter along the line of sight to the observer. The distorted images reveal information about the properties of the mass distribution between the observer and the source galaxies. Since gravitational light deflection is independent of the nature or state of matter, it provides an important tool to study the dark matter in the universe, and the relation of galaxies with the underlying dark matter. Undoubtedly, the gravitational lensing studies will play an important role in future cosmology.

The first order image distortion caused by lensing is called shear and is proportional to the induced image ellipticity. If the image size is not small, compared to the scale over which the shear varies, higher-order distortions, known as flexion, occur. Mathematically, flexion corresponds to the derivative of the shear. In combination with a strong shear, the effect is to deform circular images into arclets. Flexion is sensitive to small scale variations in the potential. It thus probes the inner part of the mass profiles of galaxy clusters, and the substructure within clusters. Furthermore, it can be used to measure the ellipticity of a galaxy dark matter halo, by using galaxy-galaxy lensing techniques.

Gravitational flexion is a new topic. It needs to be developed both observationally and theoretically. First of all, flexion is very difficult to quantify due to its mathematical complexity. Moreover, it cannot be measured directly due to the mass-sheet degeneracy in lensing systems, i.e. the observable quantity is the reduced flexion. In practical observations, there is the additional complication of the point spread function, which can also produce a shear and flexion signal that must be corrected.

Although there are some measurements on flexion the results are not very accurate at the moment. One can only make a rough estimate of how large the flexion contribution can be to the result of a lensing analysis. Current conservative results indicate that flexion measurements can significantly improve cluster mass reconstructions.

In this thesis, I will address the topic of flexion, from theory to measurements. The thesis is organized as follows:

- Chapter 1 sets the foundation needed for the understanding of the standard model of cosmology, introducing the homogeneous, isotropic, universe with structure formation.

- Chapter 2 describes the principle of weak gravitational lensing

- Chapter 3 introduces lensing flexion and reduced flexion, and derives a method to measure reduced flexion through brightness moments.

- Chapter 4 deal with two-dimensional mass reconstruction of galaxy clusters using flexion, and shows how to combine strong lensing with weak lensing shear and flexion, for the mass reconstruction.

- Chapter 5 shows preliminary constraints on the ellipticity of galaxy dark matter halos using galaxy-galaxy flexion, and introduces flexion in aperture statistics.

- Appendix A outlines the algorithm of higher-order KSB, used to measure reduced flexion.

Chapter 1

Cosmological Standard Model

Cosmology is the study of the whole Universe. A cosmological model explains the overall components, structure of the universe, and its time evolution. The standard model of cosmology, which is outlined here, is based on the assumption of large-scale homogeneity and isotropy. The linear perturbation theory on the homogeneous background explains the growth of matter inhomogeneities from small initial fluctuations to the large-scale structures. Nonlinear approaches based on N-body simulations try to model collapsed and virilized objects like galaxy clusters.

In the standard model, the universe begins with the Big Bang, and expands at different rates at different epoches (inflation, radiation-dominated epoch, matter-dominated epoch and present accelerating expansion epoch). The duration of the different epochs depends on the relative density of the universe's constituent components, as can be seen from the names of the epochs. Current observations find that most of the contents of the universe are not normal matter within our knowledge of physics. And these mysterious contents are named dark energy and dark matter, which comprise about 70% and 25% of the universe. The standard cosmological model is thus based on cold dark matter (CDM) with a cosmological constant Λ (a scenario of dark energy) and is abbreviated by the term ΛCDM model.

1.1   The expansion of the Universe

The most important cornerstone of modern cosmology is the assumption that the place we occupy in the universe is not special. This is known as the **cosmological principle**, which is favored by modern observations, such as the precise isotropy in the cosmic microwave background (CMB hereafter). Based on the isotropic cosmic background, linear cosmological perturbation theory explains the growth of matter inhomogeneities and the formation of large scale structures. The standard cosmology is also a great success for General Relativity (Einstein 1916), because gravity is the dominant force in the universe. Based on the cosmological principle, the spacetime geometry can be described by the Robertson-Walker metric, and that leads to the isotropic and homogeneous cosmic solution of the Einstein field equation − the Friedmann equations.

1.1.1 Spacetime geometry and the Einstein equation

The geometry of spacetime is described by a line element, which is

$$\mathrm{d}s^2 \equiv g_{\nu\mu}\mathrm{d}x^\nu \mathrm{d}x^\mu, \tag{1.1}$$

where the indices ν and μ range from 0 to 3, with the first one reserved for the time coordinate and the last three for spatial coordinates. We use the Einstein convention where the repeated indices are summed over. The metric $g_{\nu\mu}$ is usually assumed to be symmetric for simplicity, with four diagonal and six off-diagonal components. To describe the trajectory of particles in the spacetime, we define the Lagrangian as $\mathcal{L} = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$, where dot means derivative with respect to the affine parameter λ, which parameterizes the trajectory. Inserting it into the Euler-Lagrangian equation

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\kappa} - \frac{\partial L}{\partial x^\kappa} = 0, \tag{1.2}$$

we obtain the equation of motion as the geodesic equation $2g_{\nu\sigma}\ddot{x}^\nu + \dot{x}^\mu \dot{x}^\kappa(2g_{\sigma\kappa,\mu} - g_{\mu\kappa,\sigma}) = 0$. Since the indices μ and κ are interchangeable, i.e. $\dot{x}^\mu \dot{x}^\kappa g_{\sigma\kappa,\mu} = \dot{x}^\kappa \dot{x}^\mu g_{\sigma\mu,\kappa}$, the terms in brackets can be written as $(g_{\sigma\kappa,\mu} + g_{\mu\sigma,\kappa} - g_{\mu\kappa,\sigma})$. Defining the Christoffel symbol

$$\Gamma^\nu_{\mu\kappa} = \frac{1}{2}g^{\nu\lambda}\left[\frac{\partial g_{\lambda\mu}}{\partial x^\kappa} + \frac{\partial g_{\lambda\kappa}}{\partial x^\mu} - \frac{\partial g_{\mu\kappa}}{\partial x^\lambda}\right], \tag{1.3}$$

the geodesic equation can be written as

$$\ddot{x}^\nu + \Gamma^\nu_{\mu\kappa}\dot{x}^\mu \dot{x}^\kappa = 0. \tag{1.4}$$

A force-free particle falls along a geodesic, while the evolution of the separation between two nearby geodesics δx^ν reflects the curvature of the local space,

$$\frac{\mathrm{D}^2 \delta x^\nu}{\mathrm{D}\lambda^2} = R^\nu_{\mu\kappa\rho}\delta x^\kappa \dot{x}^\mu \dot{x}^\rho, \tag{1.5}$$

where $R^\nu_{\mu\kappa\rho}$ is the Riemann-Christoffel curvature tensor. In general relativity, gravity can be described as a geometric property of space and time. We thus write the action of $g_{\mu\nu}$ as

$$S_G = \frac{-1}{16\pi G}\int d^4x\sqrt{g}R, \tag{1.6}$$

where G is Newton's gravitational constant, $g = \det g_{\mu\nu}$, and the Ricci scalar $R = g^{\mu\nu}R_{\mu\nu}$ is defined from the Ricci tensor, which is a contraction of the curvature tensor $R_{\mu\nu} = R^\alpha_{\mu\alpha\nu}$.

The variation of the action is

$$\delta S_G = \frac{1}{16\pi G}\int \mathrm{d}^4x\sqrt{g}\left[R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right]\delta g_{\mu\nu}, \tag{1.7}$$

and the term in brackets is the Einstein tensor $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$. The action principle requires that S_G be stationary, in other words, that the variation δS_G is equal to the source, i.e. the energy-momentum tensor. This leads to the field equation of gravity, the Einstein equation

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu}. \tag{1.8}$$

The field equation is very difficult to solve. In fact, analytic solutions can only be found in the case of a highly symmetric metric.

1.1.2 The Robertson-Walker metric and the Friedmann equations

Based on the 'cosmological principle', which states that the universe is spatially isotropic and homogeneous, a solution of the field equations for an expanding universe is found. To describe the expanding universe, the scale factor $a(t)$ is introduced. A comoving distance x_0 is defined such that the physical distance at cosmic time t is $a(t)x_0$. The metric describing the expanding universe is called the **Robertson-Walker** metric, and in spherical polar coordinates it is written as

$$ds^2 = -c^2 dt^2 + a^2(t)\left(\frac{dr^2}{1 - Kr^2} + r^2 d\theta^2 + \sin^2\theta d\phi^2\right), \tag{1.9}$$

where c is the speed of light, and K stands for the spatial curvature,

$$K = \begin{cases} > 0 & \text{spherical;} \\ < 0 & \text{hyperbolic;} \\ 0 & \text{Euclidean, flat.} \end{cases} \tag{1.10}$$

Using this diagonal and spherically symmetric metric on the left side of the field equation (1.8) and introducing a perfect fluid energy-momentum tensor $T_{\mu\nu} = (P/c^2 + \rho)U_\mu U_\nu - P/c^2 g_{\mu\nu}$ on the right side, the field equations reduce to a temporal and a spatial component. These are known as the **Friedmann equations**

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{Kc^2}{a^2} = \frac{8\pi G\rho}{3}, \tag{1.11}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\frac{3P}{c^2} + \rho\right). \tag{1.12}$$

These are the fundamental equations which govern the evolution of the scale factor in an isotropic and homogeneous universe; here ρ is the energy density and P is the pressure. From the two Friedmann equations, it is easy to get an energy conservation equation (here only shown for a flat universe, $K = 0$)

$$\dot{\rho} + 3\left(\frac{P}{c^2} + \rho\right)\frac{\dot{a}}{a} = 0. \tag{1.13}$$

Equation (1.13) coupled with an equation of state $P = \hat{P}(\rho)$, which relates the pressure to the energy density, determines the evolution of energy density $\rho = \hat{\rho}(a)$ in the universe. In particular if $P = w\rho c^2$ with constant w, which is called the equation-of-state parameter, the equation has solution

$$\rho \propto a^{-3(1+w)}, \tag{1.14}$$

and the scale factor varies with the time as $a \propto t^{2/[3(1+w)]}$. Typical values for w are $w = 0, 1/3, -1$, for the cases of non-relativistic matter, relativistic matter and vacuum energy, respectively.

From Eq.(1.11), it is convenient to define a present critical density

$$\rho_c = \frac{3H_0^2}{8\pi G} = 1.878 \times 10^{-29} h^2 \text{g/cm}^3, \tag{1.15}$$

where $H_0 \equiv \dot{a}(t_0)/a(t_0)$ is the **Hubble constant**, and h is the Hubble constant in units of $100\,\text{kms}^{-1}\text{Mpc}^{-1}$. According to Eq.(1.11), the curvature constant K will be positive, 0

or negative if the present density ρ_0 is larger, equal or smaller than ρ_c. The parameters $\Omega_i \equiv \rho_i/\rho_c$ give the present fractional contribution of different components of the universe (i can denote baryons, dark matter, radiation, etc) to the critical density.

The total energy density evolution can be derived from Eqs.(1.14) and (1.11), and is written as

$$\rho = \frac{3H_0^2}{8\pi G} \left[\Omega_\Lambda + \Omega_M \left(\frac{a_0}{a}\right)^3 + \Omega_r \left(\frac{a_0}{a}\right)^4 \right], \tag{1.16}$$

with Ω_Λ, Ω_M and Ω_r being the fraction of vacuum, non-relativistic and relativistic matter energy density to the critical energy density, and the first Friedmann equation evaluated today ($a = a_0$), reduces to

$$\Omega_\Lambda + \Omega_M + \Omega_r + \Omega_K = 1, \quad \text{with} \quad \Omega_K \equiv -\frac{Kc^2}{a_0^2 H_0^2}. \tag{1.17}$$

Inserting this into Eq.(1.11), it gives

$$dt = \frac{dx}{H_0 x \sqrt{\Omega_\Lambda + \Omega_K x^{-2} + \Omega_M x^{-3} + \Omega_r x^{-4}}}, \tag{1.18}$$

where $x = a/a_0$. This equation allows us to write the present age of the universe as an integral from $x = 0$, the beginning of the expansion of the universe to $x = 1$, ($a = a_0$):

$$t_0 = \frac{1}{H_0} \int_0^1 \frac{dx}{x \sqrt{\Omega_\Lambda + \Omega_K x^{-2} + \Omega_M x^{-3} + \Omega_r x^{-4}}}. \tag{1.19}$$

1.1.3   The Cosmological redshift

If the universe is expanding, we should see galaxies receding from us. Therefore the observed wavelength is larger than the emitted one. To calculate these wavelength shifts, let us consider a light ray coming to us along the radial direction. A ray of light obeys the equation $ds^2 = 0$, so for such a trajectory, Eq.(1.9) gives the relation

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = \int_0^{r_1} \frac{dr}{\sqrt{1 - Kr^2}}. \tag{1.20}$$

Here we define that the observer sits at the center of the coordinates, light is emitted from a source at comoving coordinate r_1 and time t_1, and it arrives at the origin $r = 0$ at time t_0. Taking the differential of this relation, we see that the interval δt_1 between the departure of subsequent light signals is related to the interval δt_0 between the arrival of these light signals by

$$\frac{\delta t_1}{a(t_1)} = \frac{\delta t_0}{a(t_0)}. \tag{1.21}$$

The emitted frequency is $\nu_1 \propto 1/\delta t_1$, and the observed frequency is $\nu_0 \propto 1/\delta t_0$, so $\nu_0/\nu_1 = a(t_1)/a(t_0)$. If $a(t)$ is increasing, then this is a redshift. A decrease in frequency by a factor $a(t_1)/a(t_0)$ is equivalent to an increase in wavelength by a factor

$$1 + z = a(t_0)/a(t_1). \tag{1.22}$$

This result shows that galaxies or other sources are receding from us produce a redshift. Therefore, galaxies with redshift z are often said to have a cosmological radial velocity cz (only valid if $z \ll 1$).

1.1.4 Luminosity distance and angular diameter distance

It is difficult to define 'distance' in a curved space-time, since different methods might give different results in a curved and non-static space-time. It is these measurements that can tell us whether the expansion of the universe is accelerating or decelerating, and how fast. The most common method of determining distances in cosmology is based on the measurement of the apparent luminosity of objects of known absolute luminosity. The absolute luminosity L is the energy emitted per second, and the apparent luminosity, which is called flux l, is the energy received per second per square centimeter of receiving area. If the energy is emitted isotropically, we have the relation in Euclidean geometry, $l = L/(4\pi d^2)$. In the expanding universe, we replace $1/d^2$ with $1/r^2a^2(t_0)$ at the time t_0 that the light reaches the observer. Furthermore, the rate of the arrival of photons is lower than the rate at which they are emitted, by the redshift factor $1/(1+z)$, and also the energy of the photons received on Earth is less than the energy with which they were emitted, by the same redshift factor $1/(1+z)$. Putting this together, we can define a **luminosity distance** d_L by

$$l = \frac{L}{4\pi d_L^2} = \frac{L}{4\pi r^2 a^2(t_0)(1+z)^2}. \tag{1.23}$$

From the denominator, we obtain $d_L = a(t_0)r(1+z)$.

There is another way of defining distance, by comparing angular sizes with physical dimensions. A source of physical size S at comoving radial coordinate r_1 and time t_1 is observed to subtend a small angle θ. The **angular diameter distance** d_A is defined so that θ is given by the usual relation of Euclidean geometry $\theta = S/d_A$. In the Robertson-Walker metric, we see that

$$d_A = a(t_1)r. \tag{1.24}$$

A comparison of this result with Eq.(1.23) shows $d_L/d_A = (1+z)^2$. The angular distance is the most important distance used in gravitational lensing and in the fluctuations of the CMB.

1.2 Early Universe

The story of the early universe takes place at very high temperatures ($> 10^4 K$), and what happened in the extremely early and high temperature era is limited by our current observation and knowledge of the physics. For example, the inflation model at the moment is a paradigm not a theory. There exist a large number of models within the inflationary framework, none of which have yet convincingly emerged as a standard model of inflation.

1.2.1 Inflation

The standard model assumes that before the period of radiation dominance, during which the Robertson-Walker scale factor $a(t)$ was growing as $\sqrt{t}$, there was an earlier period called 'inflation', during which $a(t)$ grew more or less exponentially. It was Guth (1981) who first incited interest in the possibility of inflation by noting that the existence of an era of inflation would solve some of the outstanding problems of cosmology, the Three Puzzles. But his version of inflation had a fatal problem, and was soon replaced with another inflation model due to Linde (1982) and Albrecht & Steinhardt (1982).
The Flatness Puzzle

Current observations such as CMB fluctuations, are consistent with a very small spatial curvature parameter $\Omega_K \ll 1$. But Ω_K is just the present value of the dimensionless time dependent curvature parameter $-K/a(t)^2 H^2 = -K/\dot{a}(t)^2$, where $H = \dot{a}(t)/a(t)$ (in this section, the speed of light $c = 1$). In the matter-dominated era, $a(t)$ has been increasing as $t^{2/3}$ and $|K|/\dot{a}^2$ has also been increasing as $t^{2/3}$, and varies as T^{-1}, where T is temperature. Thus, if $|\Omega_K| < 1$, then at a temperature of $10^4 K$, $|K|/\dot{a}^2$ could not have been greater than about 10^{-4}. Earlier, in the radiation era, $a(t)$ was increasing as $t^{1/2}$, and $|K|/\dot{a}^2$ was increasing as T^{-2}, so it is necessary that $|K|/\dot{a}^2$ was as most about 10^{-16} at the temperature $T \simeq 10^{10} K$, and even smaller at earlier times. But there is no reason to expect the curvature to have been so very small.

Suppose the universe began with a period of inflation during which $a(t)$ increased by N e-foldings. Then $\dot{a}/a$ would have been roughly constant, $|K|/\dot{a}^2$ would have been decreasing more or less like a^{-2}. If $|K|/a^2 H^2$ had a value of order unity at the beginning of inflation, then at the time t_I (end of the inflation), it would have a value $|K|/a_I^2 H_I^2$ of order e^{-2N}, and today we will have

$$|\Omega_K| = \frac{|K|}{a_0^2 H_0^2} \sim e^{-2N} \left(\frac{a_I H_I}{a_0 H_0}\right)^2. \tag{1.25}$$

Thus the flatness problem is avoided if the expansion during inflation has the lower bound $e^N > a_I H_I/(a_0 H_0)$.

The Horizon Puzzle

The observation of the high degree of isotropy of the CMB posed a problem. The distance that light could have traveled during the lifetime of the universe is finite, and the largest distance is the horizon. The horizon size in a matter- or radiation-dominated universe is of order t, which, because $a(t)$ has increased as $t^{2/3}$ since the time of last scattering, was of order $d_H \simeq H_0^{-1}(1 + z_L)^{-3/2}$, where z_L is the redshift for the time of last scattering. The angular diameter distance d_A to that time is of order $H_0^{-1}(1 + z_L)^{-1}$, so the horizon at the time of last scattering now subtends an angle of order $d_H/d_A \simeq (1 + z_L)^{-1/2}$ radians, which for $z_L \simeq 1090$ is about 1.6 degrees. Therefore it is difficult to understand why regions that are separated by more than a few degrees on the sky are so similar in the CMB.

The Monopole Puzzle

Another problem arises from combining cosmology with particle physics. In Grand Unified Theories, magnetic monopoles are predicted to have been produced with a high abundance at a very early stage in the universe. They are predicted to be extraordinarily massive, around 10^{16}GeV (electron volt). Such particles would be non-relativistic for most of the universe's history. If monopoles did not find each other to annihilate, the number density of monopoles could reach at least one monopole per nucleon, which is in disagreement with what is observed.

Scalar fields and Slow roll

Inflation is the name given to any epoch during which the scale factor of the universe is accelerating exponentially, $\ddot{a} > 0$. An alternative expression of the condition for inflation is given in terms of H^{-1}/a, the comoving Hubble length. The condition for inflation is that the comoving Hubble length is decreasing with time. In other words, the observable universe becomes smaller during inflation (Liddle & Lyth 2000). Assuming that we work within general relativity, the condition for inflation can be written as a requirement on the material driving the expansion: $\rho + 3P < 0$. For inflation, we need a material with the unusual property of a negative pressure. Such a material is a scalar field which was introduced by particle

physicists. The energy density and pressure of a homogeneous scalar field $\phi(t)$ are

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi); \;\; P = \frac{1}{2}\dot{\phi}^2 - V(\phi), \tag{1.26}$$

where $V(\phi)$ is the potential of the scalar field, which we hope to derive from some particle physics motivation. So the energy conservation equation $\dot{\rho} = -3H(\rho + P)$ takes the form

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \tag{1.27}$$

where prime means derivative with respect to ϕ, and H is the time-dependent expansion rate, which during the period of scalar field energy domination is

$$H^2 = \frac{8\pi G \rho}{3} = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right). \tag{1.28}$$

From Eqs.(1.27) and (1.28) we can get the formula for $\dot{H}$,

$$\dot{H} = -4\pi G\dot{\phi}^2. \tag{1.29}$$

In order to have a nearly exponential expansion, the fractional change $\dot{H}/H^2$ must be much less than unity. This requires that $\dot{\phi}^2 \ll |V(\phi)|$. Thus, Eqs.(1.27) and (1.28) become

$$H^2 \simeq \frac{8\pi G}{3}V(\phi), \;\; 3H\dot{\phi} \simeq -V'(\phi), \tag{1.30}$$

which is called the condition of slow-roll approximation. We can further define two slow-roll parameters

$$\epsilon(\phi) = \frac{1}{16\pi G}\left(\frac{V'}{V}\right)^2, \;\; \eta(\phi) = \frac{1}{8\pi G}\frac{V''}{V}. \tag{1.31}$$

with the conditions $\epsilon(\phi) \ll 1$, $|\eta(\phi)| \ll 1$ given by the slow-roll approximation. Note that they are not sufficient conditions because they only restrict the form of the potential, whereas the value of $\dot{\phi}$ can be chosen freely. Under these conditions the expansion is generally not strictly exponential, but it can easily be exponentially large. Suppose that during some time interval the field $\phi(t)$ shifts from an initial value ϕ_1 to a final value ϕ_2, with $0 < V(\phi_2) < V(\phi_1)$, and the slow-roll conditions are valid. If the Robertson-Walker scale factor grows like e^{Ht}, it will increase by a factor

$$\frac{a(t_2)}{a(t_1)} = \exp\left[\int_{t_1}^{t_2} H \mathrm{d}t\right] = \exp\left[\int_{\phi_1}^{\phi_2} \frac{H}{\dot{\phi}}\mathrm{d}\phi\right] \simeq \exp\left[-\int_{\phi_1}^{\phi_2} \frac{8\pi G V(\phi)}{V'(\phi)}\mathrm{d}\phi\right]. \tag{1.32}$$

In the last step the slow-roll approximation is used. If $|V'(\phi)/V(\phi)| \ll \sqrt{16\pi G}$, then this factor is much greater than $\sqrt{4\pi G}|\phi_1 - \phi_2|$, so this guarantees that we get a large number of e-foldings. As an example of the exact solution for inflation, we consider the potential to take the form

$$V(\phi) = V_0 \exp(-\lambda\phi), \tag{1.33}$$

with V_0 and λ arbitrary constants. The equation of motion (Eq.1.27) then has the particular solution

$$\phi(t) = \frac{1}{\lambda}\ln\left(\frac{8\pi G V_0 \epsilon^2 t^2}{3 - \epsilon}\right), \;\; \text{and} \;\; a \propto t^{1/\epsilon}. \tag{1.34}$$

The increase in the Robertson-Walker scale factor during a time interval from t_1 to t_2 in which the field drops from ϕ_1 to ϕ_2 is thus

$$\frac{a(t_2)}{a(t_1)} = \left(\frac{t_2}{t_1}\right)^{1/\epsilon} = \exp[(\phi_2 - \phi_1)\lambda/2\epsilon]. \tag{1.35}$$

In order to end inflation, there must be a coupling of the inflation scalar field to other fields, including the fields of ordinary matter and radiation, so that the energy density of the inflation field is able to decrease as

$$\rho_\phi(t) = \rho_\phi(t_I)\left(\frac{a(t_I)}{a(t)}\right)^3 e^{-\Gamma(t-t_I)}, \tag{1.36}$$

where Γ is the rate of decay of the ϕ quanta into other particles, and t_I is taken at the beginning of the inflation oscillation and decay. Taking account of the flow of energy from the inflation, the conservation equation of the particles yields

$$\dot{\rho}_M + 3H(\rho_M + P_M) = \Gamma\rho_\phi. \tag{1.37}$$

we assume that the decay products of the inflation field are highly relativistic, so that $P_M = \rho_M/3$, then

$$\rho_M(t) = \frac{\rho_\phi(t_I)\Gamma a^3(t_I)}{a^4(t)}\int_{t_I}^t a(t')e^{-\Gamma(t'-t_I)}\mathrm{d}t' \tag{1.38}$$

In contemporary models of inflation this is the source of all the matter and radiation in the present universe. The matter energy density starts equal to zero at t_I, then rises at first, and finally falls as the density is attenuated by the expansion of the universe. So far, the details of inflation are unknown, and the whole idea of inflation remains a speculation, though one that is increasingly plausible. In addition to the classic problems that inflation solved at the beginning, it has had one significant experimental success: a prediction of some of the properties of the CMB and Large-Scale Structure.

1.2.2   Brief thermal history

In this section, the key point is temperature. At sufficiently early times and extremely high temperatures $> 10^{10}K$ (and high densities), the collision rate of all kinds of particles (maybe even dark matter particles at higher temperature) can keep them in thermal equilibrium. With the expansion of the universe and with a decreasing temperature, the particle species start to decouple. Some of the decoupling processes leave fossil imprints from which we can learn the story of the early universe.

At the epoch when temperatures were between 10^4K and $10^{11}K$, the collision rate of photons with electrons and other particles was greater than the expansion rate of the universe. Thus the photons and particles can be assumed to have been in thermal equilibrium. The condition of thermal equilibrium tells us that the entropy in a co-moving volume is fixed

$$s(T)a^3 = \text{constant}, \tag{1.39}$$

where s is the entropy density. The variation of the temperature with time is governed by this equation and the first Friedmann equation (1.11). The energy density of massless particles

$$\rho(T) = g\int_0^\infty \frac{4\pi p^3 \mathrm{d}p}{(2\pi\hbar)^3}\left(\frac{1}{\exp(p/k_BT)\pm 1}\right) = \begin{cases} ga_BT^4/2, & \text{for Bosons} \\ 7ga_BT^4/16, & \text{for Fermions} \end{cases} \tag{1.40}$$

where g is the number of spin states of the particles, $\hbar$ is the reduced Planck constant, k_B is the Stefan-Boltzmann constant, and a_B is a factor calculated from the integral. Thus, during the epoch that the dominant constituent of the universe is a highly relativistic ideal gas, the entropy and energy densities are given by

$$s(T) = \frac{2\mathcal{N}a_B T^3}{3}, \;\; \rho(T) = \frac{\mathcal{N}a_B T^4}{2},$$ (1.41)

where $\mathcal{N}$ is the number of particle types, counting particles and antiparticles and each spin state separately, and the extra factor of 7/8 for fermions.

At the time when the temperature was around $10^{11}K$, the neutrinos were kept in thermal equilibrium by neutral current reactions, like $e^+ + e^- \rightleftharpoons \mu + \mu^-$. Hence the constituents of the universe at this time were photons with two spin states, plus three generations of neutrinos and their anti-partners, each with one spin state, plus electrons and positrons, each with two spin states, all in equilibrium and highly relativistic, giving

$$\mathcal{N} = 2 + \frac{7}{8}(6 + 4) = \frac{43}{4}.$$ (1.42)

The interesting point here is that the energy density depends on $\mathcal{N}$, and can effect the expansion rate, and thus the temperature at which the particles decouple. On the other hand, in the weak interaction between a neutron and a proton, e.g. $n + \nu \leftrightarrow p + e^-$, the ratio of neutron conversion to proton conversion also depends on the temperature

$$\frac{\lambda(p \to n)}{\lambda(n \to p)} = \exp\left(-Q/k_B T\right), \;\; \text{with} \;\; Q = m_n - m_p = 1.293 \text{ MeV},$$ (1.43)

which means that the different number of particle types $\mathcal{N}$ would lead to a different temperature for when the neutrinos decouple from the others, and a different ratio of neutrons to protons.

At a temperature of about $10^{10}K$ neutrinos were just going out of equilibrium. The collision rate of neutrinos with electrons or positrons was much lower than the expansion rate. The neutrinos then began to move with the speed of light, and did not interact with the other kinds of particles. This makes the number density distribution n_ν retain the form of the Fermi-Dirac distribution without a chemical potential, and with a temperature $T_\nu \propto 1/a$. At this temperature we need to take into account the finite mass of the electron, so the temperature T of the electrons, positrons and photons (still in coupled) will no longer be the same as the neutrino temperature T_ν. When the temperature dropped to $10^9 K$, meaning $k_B T < m_e c^2$, the electrons and positrons annihilated, $\mathcal{N}$ changed from $2 + \frac{7}{8}4$ to 2. The entropy conservation law (Eqs.1.39 and 1.41) gives $a^3 T^3 \mathcal{N} \sim$ constant, so the temperature of photons rose by a factor of $(11/4)^{1/3}$. In the other words, the ratio T/T_ν rose from about unity for $T > 10^{11}K$ to about 1.401 at $T < 10^9 K$. In particular, at the present time, the $T_{\text{CMB}} = 2.725K$ gives the neutrino temperature $1.945K$, so there is a cosmic neutrino background which we are not yet able to detect.

After the era of electron-positron annihilation, the energy density of the universe was dominated for a long time by photons, neutrinos and antineutrinos, all of them highly relativistic, so we still have $s(T) \propto T^3$. Not until the lower temperature $\sim 4000K$, must we take into account the energy density of non-relativistic matter, and it took about hundreds of thousands of years for the universe to cool down to that point.

1.2.3   Recombination and Reionization

The universe is expanding, so we expect that in the past matter was hotter and denser than at present, as discussed in previous sections. If we look far enough backward in time we come to an era when it was too hot for electrons to be bound into atoms, and the rapid collisions of photons with electrons would have kept radiation in thermal equilibrium with matter. Then, the energy density of photons is given by the black-body spectrum

$$\epsilon(\nu, T)d\nu = \frac{8\pi h}{c^3} \frac{\nu^3 d\nu}{\exp(h\nu/k_B T) - 1},\tag{1.44}$$

where ν is the frequency of a photon, and h is Planck's constant. As time passed, the matter became cooler and less dense, and when the temperature dropped to about $10^5 K$, the photons no longer exchanged energy efficiently with electrons. Photons continued to be scattered by free electrons, but without gain or loss of energy. This terminated when the free electrons become bound into hydrogen and helium atoms, ending the scattering of photons. At that time, the cosmic temperature T was of order $3000 K$, and the redshift was about 1090. This is called **recombination**. The photons then began a free expansion, with their spectrum keeping the same form. This is the **cosmic microwave radiation background** observed today, which keeps the emitted form but with a much lower temperature, $2.725 K$ (Fig. 1.1).

The neutral phase of the baryonic pre-galactic medium lasts from $z = 1100$ to $z \sim 14$. During this period inhomogeneities in the matter field that were initially from quantum fluctuations and enhanced by inflation, grow by gravitational instability. The primordial density perturbations and formation of the structure will be described in Sec. 1.3.

From $z \sim 14$ neutral hydrogen is reionized by UV radiation from the first generation of luminous objects, either QSOs or the first stars. Unlike recombination, cosmic reionization occurs over a large interval of redshift, from $z \sim 14$ to 6. The epoch of a neutral medium is called 'Dark Ages', and the process of reionization of this medium is still poorly studied (see Fan et al. 2006; Ciardi & Ferrara 2005, for reviews). The future low-frequency radio telescopes, like LOFAR and SKA, offer a probe of the neutral media by measuring the 21-cm line of neutral hydrogen (Zaroubi & Silk 2005). Indeed, the 21-cm line, which is associated with the hyperfine transition of HI, is an ideal probe of neutral hydrogen since it is not absorbed by dust. The spin temperature quantifies how much HI atoms are in the higher energy level of the hyperfine transition. It has information on the formation of the first generation objects as well as large-scale HI distribution.

Observational evidence suggests that the first sources of light in the universe are stellar-type objects, although the existence of primordial mini-quasars is not completely ruled out. The objects which formed first are predicted to have mass corresponding to virial temperatures $T_{\mathrm{vir}} < 10^4 K$. Once the gas has virialized in the potential wells of preexisting dark matter halos, and is cooling to further collapse, star formation is ignited. Then energy injection and radiation can affect the subsequent galaxy formation process and intergalactic medium evolution.

1.3   Large-Scale Structure

The Universe is inhomogeneous on scales below $\sim 200\, h^{-1} \mathrm{Mpc}$, showing structures such as filaments, clusters and galaxies. The established mechanism for structure formation is that
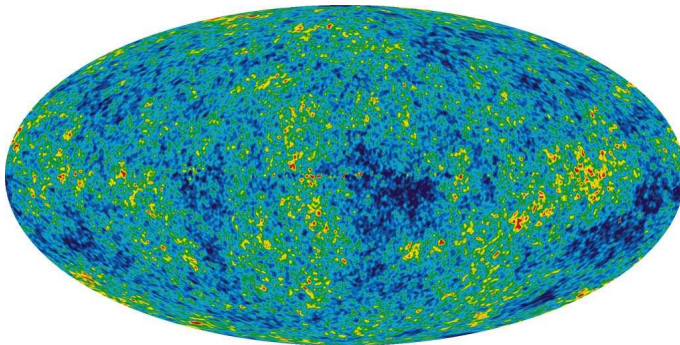
Figure 1.1: The cosmic microwave temperature fluctuations from the 5-year WMAP data seen over the full sky. The average temperature is 2.725K, and the colors represent the tiny temperature fluctuations, as in a weather map. Red regions are warmer and blue regions are colder by about 0.0002 degrees. Credit: NASA/WMAP Science Team.

primordial density perturbations grow through gravitational instability to form what we observe today (here structure mainly means structure of matter and photons). The equation governing the overdensities δ schematically reads (Dodelson 2003)

$$\ddot{\delta} + [\text{Pressure} - \text{Gravity}]\,\delta = 0. \tag{1.45}$$

These forces act in opposite ways, Gravity acts to increase overdensities. While if the 'Pressure' term in Eq.(1.45) is strong, inhomogeneities do not grow.

1.3.1 Boltzmann equation

The Boltzmann equation is one of the most important equations in statistical mechanics that deals with system not in thermodynamic equilibrium. We are interested in the anisotropy and inhomogeneities in the cosmic distribution of photons and matter. Their evolution can be obtained from the Boltzmann equation for all species in an expanding universe

$$\frac{\mathrm{d}f(t,x,p)}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} = C[f], \tag{1.46}$$

where $f(t,x,p)$ is the distribution function in 'particle' position and momentum phase space, at time t, position x and momentum p. The $C[f]$ term contains all possible collision terms, and is related to the 'Pressure' term in Eq.(1.45), and

In order to get some useful physics from the Boltzmann equation, we need to determine the coefficients $\mathrm{d}x/\mathrm{d}t$, $\mathrm{d}p/\mathrm{d}t$, hence the equations of motion, and specify the form of the metric. Assuming small perturbations around the flat FRW universe model, the metric can be written as

$$g_{00} = -1 - 2\Psi(x,t),\ g_{0i} = 0,\ g_{ij} = a^2\delta_{ij}(1 + 2\Phi(x,t)). \tag{1.47}$$

The coefficients $\mathrm{d}x/\mathrm{d}t$, $\mathrm{d}p/\mathrm{d}t$ can thus be obtained by applying this metric to the geodesic equation (1.4). Here the perturbation Ψ corresponds to the Newtonian potential, and Φ is the perturbation to the spatial curvature. It is important to mention that Eq.(1.47) contains only scalar perturbations. It is possible that the metric also has vector and tensor perturbations. For example, inflation predicts that there will be tensor perturbations which will create cosmic relic gravitational waves (Rubakov et al. 1982). Since we focus here on the matter perturbations, which are mainly coupled to the scalar perturbations, we only need Ψ and Φ, and treat them as small quantities. In current popular cosmological models, dark matter plays an important role in structure formation and in determining the gravitational field in

the universe. We will thus consider the derivation of Eq.(1.46) for cold dark matter (CDM) only. First, there is no collision term in Eq.(1.46) since dark matter by definition does not interact with any of the other constituents. It is convenient to define the number density for dark matter particles

$$n \equiv \int \frac{\mathrm{d}^3 p}{(2\pi)^3} f, \tag{1.48}$$

and the energy-momentum tensor

$$P^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = \left(\frac{\mathrm{d}t}{\mathrm{d}\lambda}, \ \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\lambda} \right) = (E, \ \boldsymbol{p}), \tag{1.49}$$

where E is the energy and $\boldsymbol{p}$ is the momentum vector. The velocity can be written in terms of E and $\boldsymbol{p}$,

$$\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \frac{\boldsymbol{p}}{E}. \tag{1.50}$$

Thus, the velocity for dark matter particles is

$$\boldsymbol{v} = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} f \frac{p\hat{\boldsymbol{p}}}{E}. \tag{1.51}$$

Since we are interested in the evolution of the dark matter distribution, we can use the integral of Eq.(1.46) over the momentum p phase space, which is called nth moment $\int \mathrm{d}^3 p/(2\pi)^3 p^n f$. Then the zeroth moment of the Eq.(1.46) for CDM leads to the cosmological generalization of the continuity equation:

$$\frac{\partial n}{\partial t} + \frac{1}{a}\frac{\partial n v^i}{\partial x^i} + 3\left[\frac{\dot{a}}{a} + \frac{\partial \Phi}{\partial t} \right] n = 0. \tag{1.52}$$

The first two terms are the standard continuity equation from fluid mechanics, the last term arises due to the FRW metric and its perturbations. We neglect the metric perturbations Φ, and replace the number density as mass density ρ by $\rho = n m_p$ in Eq.(1.52), where m_p is the mass of single particle, then the equation looks more familiar

$$\frac{\partial \rho}{\partial t} + 3H\rho + \frac{1}{a}\nabla(\rho \boldsymbol{v}) = 0, \tag{1.53}$$

here ∇ is the gradient with respect to comoving spatial coordinates. In a similar way, the first moment of the Boltzmann equation (1.46) for CDM becomes the momentum conservation equation, or Euler equation

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\nabla)\boldsymbol{v} = -\frac{1}{a}\nabla\Psi. \tag{1.54}$$

These equations together with the Poisson equation

$$\nabla^2 \Psi = 4\pi G \rho a^2, \tag{1.55}$$

govern the evolution of the density ρ (of dark matter here), the potential Ψ and the peculiar velocity field $\boldsymbol{v}$ in an expanding universe. The peculiar velocity can also be written as the proper velocity $\boldsymbol{u}$ minus the Hubble flow $\boldsymbol{v} = \boldsymbol{u} - H\boldsymbol{x}$.

It is useful to define the density contrast

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}}, \tag{1.56}$$

where $\bar{\rho}$ is the background density of a homogeneous and isotropic universe. This δ which departs from the homogeneity is the beginning of the large scale structure. Therefore in terms of δ, Eq.(1.53)-(1.55) become

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot [(1+\delta)\boldsymbol{v}] = 0 \quad \text{Continuity equation}, \tag{1.57}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + H\boldsymbol{v} + \frac{1}{a}(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\frac{1}{a}\nabla\Psi \quad \text{Euler equation}, \tag{1.58}$$

$$\nabla^2\Psi = \frac{3H_0^2\Omega_{\mathrm{M}}}{2a}\delta \quad \text{Poisson equation}. \tag{1.59}$$

1.3.2 Linear Evolution

Considering only terms linear in δ and v, Eq.(1.57) and Eq.(1.58) become $\dot{\boldsymbol{v}} + H\boldsymbol{v} = -a^{-1}\nabla\Psi$ and $\dot{\delta} + a^{-1}\nabla\boldsymbol{v} = 0$, respectively. Together with the Poisson equation, they can be combined into a second-order linear differential equation for δ

$$\ddot{\delta} + 2H\dot{\delta} - \frac{3H_0^2\Omega_{\mathrm{M}}}{2a^3}\delta = 0. \tag{1.60}$$

The general solution has the form

$$\delta(x, a) = D_+(a)\delta_+(x) + D_-(a)\delta_-(x), \tag{1.61}$$

where $D_\pm(t)$ are two linearly independent solutions of

$$\ddot{D} + 2H\dot{D} = \frac{3H_0^2\Omega_{\mathrm{M}}}{2a^3}D. \tag{1.62}$$

$D_-(a)$ is called the decaying mode and can be neglected for later times since density perturbations in this mode quickly decay and play no role in structure formation. The other one D_+ is called the growing mode. In general it can be written as

$$D_+(a) = \frac{5\Omega_{\mathrm{M}}}{2}\frac{H(a)}{H_0}\int_0^a \mathrm{d}a' \left[1 + \Omega_{\mathrm{M}}\left(\frac{1}{a'} - 1\right) + \Omega_\Lambda\left(a'^2 - 1\right)\right]^{-3/2}, \tag{1.63}$$

where we assume Ω_{r} and Ω_K are 0. If $D_+(a)$ is normalized to unity today $(D_+(a_0) = 1)$, it is called the **growth factor**. Different solutions are obtained if different values of the density parameters are considered. The explicitly solution for the growing mode of the EdS universe $(\Omega_m = 1, \Omega_\Lambda = 0)$ provides a rough approximation for others.

The perturbation evolution can be written for each Fourier mode of δ with wavelength k, and one has to account for the exact expansion history during each different epoch of the background evolution. The result depends on the wavelength mode being smaller or larger than the evolving horizon d_{H}. An important feature for the growth of a given mode, is if it enters the horizon in the epoch of radiation domination $(a \ll a_{\mathrm{eq}})$ or matter domination $(a \gg a_{\mathrm{eq}})$. The evolution of δ is summarized in Table 1.1. From the table is clear that short

Table 1.1: The evolution of a linear perturbation during different epochs for different wavelength modes.

Long-wavelength ($\lambda \gg d_{\mathrm{H}}$)	short-wavelength ($\lambda \ll d_{\mathrm{H}}$)	
$\delta \propto a^2$	no growth	if $a \ll a_{\mathrm{eq}}$,
$\delta \propto a$	$\delta \propto a$	if $a \gg a_{\mathrm{eq}}$.

wavelength modes that enter the horizon during radiation domination are suppressed by a factor $(a_{\mathrm{enter}}/a_{\mathrm{eq}})^2$ with respect to long wavelength modes.

To account for this scale-dependent linear growth, one can define a **transfer function**

$$T(k) = \frac{\delta(k, a=1)}{\delta(k=0, a=1)} \frac{\delta(k=0, a)}{\delta(k, a)}, \tag{1.64}$$

where $k = 0$ indicates some large scale (formally infinite). The initial perturbations at scale factor a_i are outside the horizon. Then $T(k)$ is independent of a_i for all relevant scales. For CDM perturbations, the behavior of the transfer function is asymptotic for both large and small k

$$T(k) \approx \begin{cases} 1 & \text{for} \quad \lambda \gg \lambda_{\mathrm{eq}}; \\ (\lambda/\lambda_{\mathrm{eq}})^2 & \text{for} \quad \lambda \ll \lambda_{\mathrm{eq}}, \end{cases} \tag{1.65}$$

where $\lambda = 2\pi/k$ is the comoving wavelength of a perturbation, and

$$\lambda_{\mathrm{eq}} = \frac{2\pi}{k_{\mathrm{eq}}} = d_{\mathrm{H}}(a_{\mathrm{enter}}). \tag{1.66}$$

The behavior for intermediate value of k can be calculated numerically through some public software, like CMBfast[1]. And a fitting formula has been given by Bardeen et al. (1986),

$$T(k) = \frac{\ln(1 + 2.34q)}{2.34q}[1 + 3.89q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4]^{-1/4}. \tag{1.67}$$

with $q = k/(\Gamma h \mathrm{Mpc}^{-1})$, and the **shape parameter** defined as

$$\Gamma = \Omega_{\mathrm{M}} h. \tag{1.68}$$

A generalization taking baryons into account was given by Sugiyama (1995),

$$\Gamma = \Omega_{\mathrm{M}} h \exp[-\Omega_b(1 + \sqrt{2h}/\Omega_{\mathrm{M}})], \tag{1.69}$$

where Ω_b is the fraction of baryon density to the critical energy density. This approach is only an approximation. Eisenstein & Hu (1998) found a new transfer function with a more complete treatment of baryons.

It is interesting to note that from the observations of CMB fluctuations, one knows that at recombination ($z \approx 1090$) the baryon density perturbations were of the order of 10^{-5}. Since then, with a growth of $\delta \propto a$, they could only grow by a factor of $\sim 10^3$ reaching today a density contrast of about 10^{-2}. This is a very strong indication that the major

---

[1]www.cmbfast.org

fraction of the matter is not baryonic, but only interacts gravitationally – dark matter – since non-linear perturbations much larger than this exist. In order to describe the non-linear evolution, one has to find exact solutions of the non-linear dynamical equations, which is impossible in an analytic way. Therefore a semi-analytic method, like using a perturbative approach, or numerical simulations are needed. Section 1.3.5 will mention aspects of non-linear perturbation theory.

1.3.3 Correlation function and power spectrum

The matter density contrast field δ in the universe has to be described statistically. We consider our universe as a random field, determined by its moments (mean, n-point correlation functions). To completely describe its properties an infinite number of moments is needed in general. For the special case of a Gaussian random field, the mean and two-point correlation function are sufficient to completely describe the field. In the early universe, the initial inhomogeneities are Gaussian (a prediction of Inflation theory). Later, by gravitational instability, non-linear and non-Gaussian features arise in the matter field and higher-order moments of the density field need to be considered. Since the mean of the density contrast vanishes by construction ($\langle \delta(\boldsymbol{x}) \rangle = 0$), the second and third moments are studied. In Fourier space, these are the power spectrum and bispectrum, respectively.

Generally, the two-point correlation function (2PCF) is defined as

$$\langle \delta(\boldsymbol{x})\delta(\boldsymbol{y}) \rangle = C(|\boldsymbol{x} - \boldsymbol{y}|), \tag{1.70}$$

where angular brackets denote an ensemble average. The 2PCF depends only on the separation $|\boldsymbol{x} - \boldsymbol{y}|$ of the two points if the δ field is statistically homogeneous. The power spectrum of the density fluctuations P_δ is defined by

$$\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}') \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}') P_\delta(|\boldsymbol{k}|), \tag{1.71}$$

where $\delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}')$ is the Dirac delta function.

From the results of last section, the power spectrum for a given time can be calculated from the initial one, $P_\delta(k, a_i)$, as

$$P_\delta(k, a) = T^2(k) \frac{D_+^2(a)}{D_+^2(a_i)} P_\delta(k, a_i). \tag{1.72}$$

Usually one assumes a simple and scale-free power law $P_\delta(k, a_i) \propto k^{n_s}$, with spectral index $n_s \lesssim 1$. The power spectrum normalization is determined from observations, usually considering σ_8, the variance of the density fluctuations in spheres of radius $8\,h^{-1}\mathrm{Mpc}$. This particular scale was chosen because the dispersion of the number of galaxies measured in this volume is of order unity, $\sigma_{8,\mathrm{gal}} \sim 1$. Assuming that galaxies are tracers of underlying dark matter, one can relate the galaxies and dark matter fluctuation amplitudes. A simple linear relation is $\sigma_{8,\mathrm{gal}}^2 = b^2 \sigma_8^2$, where b is the linear bias factor.

1.3.4 The bispectrum

In regions that are highly non-linear due to gravitational collapse, non-Gaussian features arise. A precise treatment requires higher order moments, nPCFs as a natural extension of

Figure 1.2: The power spectrum of density fluctuations in the Universe, as determined by different methods. Here, $\Delta^2(k) \propto k^3 P(k)$ vs scale (wavelength) is plotted, where $k = 2\pi/\lambda$ is the wavenumber. Going from large to small scales, the results presented here are obtained from CMB temperature fluctuations, from the abundance of galaxy clusters, from the large-scale distribution of galaxies, from cosmic shear, and from the statistical properties of the Lyα forest. One can see that the power spectrum of a ΛCDM model is able to describe all these data over many orders of magnitude in scale (Source: Max Tegmark)
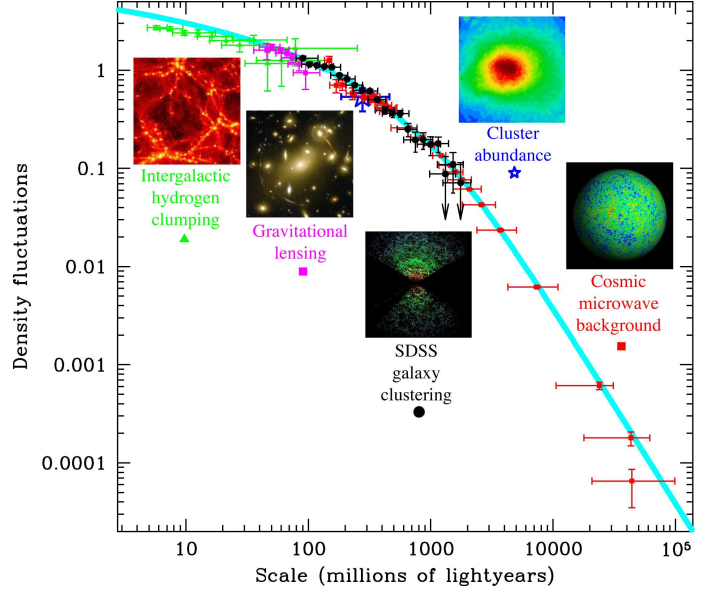
2PCFs. The Fourier transform of the 3PCF is the bispectrum $B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, a)$

$$\langle \delta(\boldsymbol{k}_1, a)\delta(\boldsymbol{k}_2, a)\delta(\boldsymbol{k}_3, a) \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3)[B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, a) + B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_3, a) + B_\delta(\boldsymbol{k}_2, \boldsymbol{k}_3, a)]. \tag{1.73}$$

Assuming an initial Gaussian density field, the bispectrum can be approximated through second-order perturbation theory (Bernardeau et al. 2002),

$$B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, a) = F_2(k_1, k_2, \cos\phi)P_\delta(k_1, a)P_\delta(k_2, a), \tag{1.74}$$

where P_δ is the linear power spectrum and $\cos\phi = (\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)/(k_1 k_2)$. Unlike the power spectrum, the bispectrum is a function of three variables, i.e. the unique condition for a triangle. The kernel function in an EdS-universe is

$$F_2(k_1, k_2, \cos\phi) = \frac{10}{7} + \cos\phi\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}\cos^2\phi. \tag{1.75}$$

It is convenient to define the reduced bispectrum Q as

$$Q(k_1, k_2, \cos\phi, a) = \frac{B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, a) + B_\delta(\boldsymbol{k}_2, \boldsymbol{k}_3, a) + B_\delta(\boldsymbol{k}_3, \boldsymbol{k}_1, a)}{P_\delta(k_1, a)P_\delta(k_2, a) + P_\delta(k_2, a)P_\delta(k_3, a) + P_\delta(k_3, a)P_\delta(k_1, a)}. \tag{1.76}$$

Since the bispectrum is proportional to P_δ^2, Q is mostly independent of time and cosmology, but dependent on the triangle configuration of the three Fourier vectors $\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3$.

1.3.5   Non-linear evolution

In the low-redshift universe the density contrast reaches values much higher than unity, up to $\delta \sim 1000$ for galaxy clusters. These objects can no longer be described with linear theory. One possible description is provided by the halo model, an empirical model which combines results from simulations and theoretical models. The main assumption is that all dark matter is distributed in spherically symmetric dark matter halos. The halo model was developed by

Seljak (2000); Ma & Fry (2000); Scoccimarro et al. (2001). Galaxies are then introduced according to the halo occupation distribution $P(N|m)$, which is a conditional probability of a halo of mass m to contain N galaxies. Then nPCF or power spectrum for galaxies can be calculated approximately.

On small scale, the nPCF depend on the halo density profile. The halo radial mass profile can be identified from numerical simulations. A universal functional form of the halo density profile was found in simulations by Navarro et al. (1996), which is called the **NFW**-profile

$$\rho(r) = \frac{\rho_s}{(r/r_{\rm s})(1 + r/r_{\rm s})^2}, \tag{1.77}$$

where ρ_s is the amplitude of the density profile, and r_s specifies a characteristic radius. For $r \ll r_s$, $\rho \propto r^{-1}$, and for $r \gg r_s$, the profile follows $\rho \propto r^{-3}$, which means r_s is the radius where the slope of the density profile changes.

The virial mass, which is approximately the mass contained in a region within which the average density is ~ 200 times the critical density at redshift z. $M_{\rm vir} = 4\pi/3 r_{200}^3 \, 200\rho_{\rm cr}(z)$, can be expressed as

$$M_{\rm vir} = \int_0^{r_{200}} \frac{\rho_s}{r/r_s(1 + r/r_s)^2} 4\pi r^2 {\rm d}r. \tag{1.78}$$

The characteristic density can be expressed as

$$\rho_s = \frac{200}{3}\rho_{\rm cr}(z)\frac{c^3}{\ln(1 + c) - c/(1 + c)}, \tag{1.79}$$

where the concentration parameter c is defined as

$$c \equiv r_{200}/r_s. \tag{1.80}$$

The larger the value of c, the more strongly the mass is concentrated towards the inner region. The NFW profile has two parameters, either ρ_s and r_s, or r_{200} and c, that describe the shape of the mass distribution. These parameters can be estimated from observation, such as weak lensing and X-ray of clusters.

There is another density profile, mathematically simple, which is not a bad approximation, given by

$$\rho(r) = \frac{\sigma_v^2}{2\pi G r^2}, \tag{1.81}$$

where σ_v is the velocity dispersion of the 'particles'. This simpler analytic model is called Singular Isothermal Sphere, and describes a spherically symmetric, self-gravitating system. The lensing properties of these dark matter halo models are discussed in Chapter 5.

1.3.6 The substructure problem

The ΛCDM cosmology is up to now the most successful model in describing cosmological observations. However there are some results in apparent conflict with observations, and it is particularly interesting to look into these problems, since we can gain more knowledge about the universe, and eventually could find out some new physics.

ΛCDM cosmological simulations of structure formation show that a halo of mass M contains numerous halos of much lower mass, sub-halos. But they are not observed: For example, only around 20 satellite galaxies are found around the Milky Way, whereas the simulations

predict several hundreds of subhalos (Kauffmann et al. 1993; Klypin et al. 1999; Diemand et al. 2007). The presence of substructure over a wide range of scales is a direct consequence of hierarchical structure formation. The missing sub-halos in observation is thus a problem for CDM models. On the other hand, simulations predict mass concentrations whereas observations look for light. Moreover the simulations only take CDM particles into account, the baryonic process being still too complex for current computers.

This is a problem under debate (Kravtsov et al. 2004; Gao et al. 2004; Zentner et al. 2005; Sawala et al. 2010). One powerful tool for probing sub-halos within galaxies and clusters is gravitational lensing, since it measures the total amount of mass, dark and luminous. This problem will be addressed in the following chapters of this thesis.

Chapter 2

Gravitational Lensing

In this chapter, we come to another application of the Einstein equation besides the Friedmann cosmology−`gravitational lensing` (GL hereafter), which is the name of the phenomenon of deflection of light in gravitational fields. The behavior of light rays in a gravitational field was studied before General Relativity, 18th century by Newton and Laplace among others. However in 1919, Eddington confirmed the value of the deflection angle produced by the Sun during a Solar eclipse, which strongly supported the newborn gravity theory−GR. And this may be considered the first observation of GL.

In 1924, Chwolson (1924) published a short note, in which he considered a star in the foreground of a distant star. In case of perfect alignment of lens and source, a ring-shaped image of the background star should form around the lens star. Such kind of image is now known as the Einstein ring (should be called 'Chwolson ring' fairly). Not until the 1970's, did the observations make some real progress, thanks to the discovery of quasars, which are ideal distant sources for GL.

In my thesis, I only study GL with light deflection in the weak-field, small deflection limit; strong-field cases are not considered here (like near a black hole or neutron star). For details about GL, the reader is referred to Schneider et al. (1992, 2006).

2.1 Basic lensing

2.1.1 Deflection angle

To calculate the image distortion of a background source by GL, we need to know first what is the deflection angle for a light ray. This problem can be solved by applying the Schwarzschild metric (Schwarzschild 1916) for a point mass case

$$\mathrm{d}s^2 = \left[1 - \frac{2MG}{c^2 r}\right] c^2 \mathrm{d}t^2 - \left[1 - \frac{2MG}{c^2 r}\right]^{-1} \mathrm{d}r^2 - r^2 \mathrm{d}\theta^2 - r^2 \sin^2\theta\ \mathrm{d}\phi^2, \qquad (2.1)$$

where M is the mass of the point and r is the distance to the point. Recalling the geodesic equation (Eq.1.4), the equation of motion yields

$$\frac{1}{r^2}\left(1 - \frac{2MG}{rc^2}\right)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 - \frac{r^2}{r_0^2}\frac{1 - \frac{2MG}{r_0 c^2}}{1 - \frac{2MG}{rc^2}} + 1 = 0, \qquad (2.2)$$

where r_0 is the impact parameter, and is much larger than the Schwarzschild radius of the lens ($R_s = 2GM/c^2$). The absolute value of the deflection angle can be obtained in the weak field approximation

$$\hat{\alpha} = 2|\phi(r_0) - \phi(\infty)| - \pi \approx \frac{4GM}{c^2 r_0}. \tag{2.3}$$

In order to determine the deflection angle of an extended mass lens with density $\rho(\boldsymbol{r})$, we can simply add up the deflection caused by each mass element $dm(\boldsymbol{r}) = \rho(\boldsymbol{r})dV$ of the mass distribution. Then the deflection angle provided by an individual lens is

$$\hat{\boldsymbol{\alpha}}(r) = \frac{4G}{c^2} \int d^2 r' \int dx_3' \, \rho(\boldsymbol{r}', x_3') \frac{\boldsymbol{r}' - \boldsymbol{r}_0}{|\boldsymbol{r}' - \boldsymbol{r}_0|^2} = \frac{4G}{c^2} \int d^2 r' \Sigma(\boldsymbol{r}') \frac{\boldsymbol{r}' - \boldsymbol{r}_0}{|\boldsymbol{r}' - \boldsymbol{r}_0|^2}, \tag{2.4}$$

where we have defined the surface mass density

$$\Sigma(\boldsymbol{r}) = \int dx_3 \, \rho(\boldsymbol{r}, x_3). \tag{2.5}$$

The expression is valid as long as the extent of the mass distribution along the line of sight is much smaller than the distances between source, lens and observer. This thin-lens approximation is well satisfied when considering lensing by galaxies or clusters, but fails for cosmic shear where the mass distribution causes multiple deflections all along the line of sight.

2.1.2   Lensing equation

Once given the deflection angle, it is easy to set up an equation known as the lens equation which relates the position of the images and the source by the geometrical configuration of the lens system (Fig.2.1). The source is located in the background at an angular position β. The light emitted from the source is deflected in the gravitational potential of the foreground mass distribution on its way to the observer. Thus, the image is observed at an angular position θ. With the scaled deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta}) = (D_{ds}/D_s)\hat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta})$, we can get a concise form for the lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{2.6}$$

The solution of the lens equation for a given source position $\boldsymbol{\beta}$ is the observed position $\boldsymbol{\theta}$ of the image. If more than one solution(image) can be found, this corresponds to multiple images.

It is convenient to define the critical surface density as

$$\Sigma_{cr} = \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}}, \tag{2.7}$$

then the dimensionless surface mass density $\kappa(\boldsymbol{\theta}) = \Sigma(D_d\boldsymbol{\theta})/\Sigma_{cr}$ is also the convergence. A mass distribution with $\kappa > 1$ at some position $\boldsymbol{\theta}$ produces multiple images when there is a source at an appropriate position. The deflection angle can be rewritten using κ

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2}. \tag{2.8}$$

We therefore can define the deflection potential by

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla\psi(\boldsymbol{\theta}), \quad \psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\theta' \, \kappa(\boldsymbol{\theta}') \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'|, \tag{2.9}$$
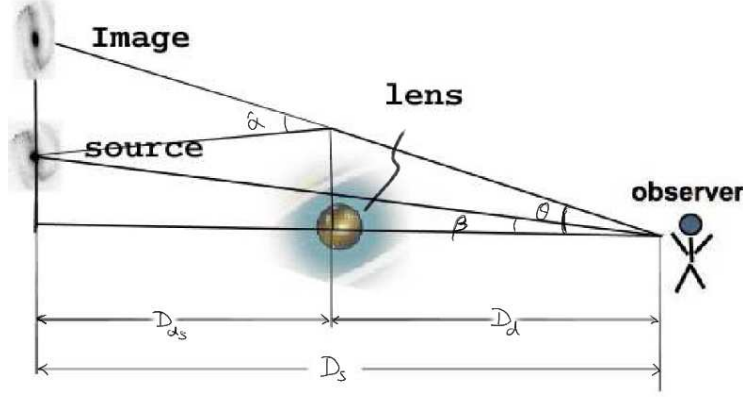
Figure 2.1: Geometry of a thin lens system: D_d, D_s and D_ds are the angular diameter distances from the observer to the lens, from the observer to the source and from the lens to the source. Note that this figure is a 2-D projection of lensing system. The lens and source plane are defined as being perpendicular to the optical axis which is the line connecting observer and lens.

which is the 2-D projection of the gravitational potential. It satisfies the 2-D Poisson equation

$$\nabla^2 \psi = \psi_{,11} + \psi_{,22} = 2\kappa, \tag{2.10}$$

where indices separated by a comma denote partial derivatives with respect to θ_i. The **shear** is introduced

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}), \quad \gamma_2 = \psi_{,12}. \tag{2.11}$$

The two-component shear $\gamma = (\gamma_1, \gamma_2)$ is a spin-2 quantity, where a spin-N quantity means that it has the same value after rotation by $2\pi/N$. It is usual to write the shear as a complex number $\gamma = \gamma_1 + \mathrm{i}\gamma_2 = |\gamma|\mathrm{e}^{2\mathrm{i}\varphi}$, where the polar transformation property of the shear has been accounted for by the factor 2 in front of the polar angle φ. The linear lensing equation for an extended source in complex notation can thus be written as

$$\boldsymbol{\beta} = (1 - \kappa)\boldsymbol{\theta} - \gamma\boldsymbol{\theta}^*, \tag{2.12}$$

where $\boldsymbol{\theta}^*$ is the complex conjugate of $\boldsymbol{\theta}$. It is also common to express the shear with respect to a specified direction, like a foreground galaxy (Fig. 2.2). This defines the tangential and cross shear

$$\gamma_\mathrm{t} = -\gamma_1 \cos 2\varphi - \gamma_2 \sin 2\varphi, \quad \gamma_\times = \gamma_1 \sin 2\varphi - \gamma_2 \cos 2\varphi, \tag{2.13}$$

where φ is the polar angle to the specified direction.

2.2 Image distortion

Besides point sources (like QSO or stars), there are also extended sources in the universe, like galaxies. Since they are usually small sources in angular size, and considering only small deflection angles, the lens equation can be Taylor expanded. To linear order we get,

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \boldsymbol{\beta}(\boldsymbol{\theta}_0) + \mathcal{A}(\boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0), \tag{2.14}$$

Figure 2.2: Illustration of the tangential and cross component of shear (Eq.2.13), measured with respect to the center of the circles. Left panel: $\gamma_t > 0$, $\gamma_\times = 0$; right panel: $\gamma_t > 0$, $\gamma_\times = 1$.
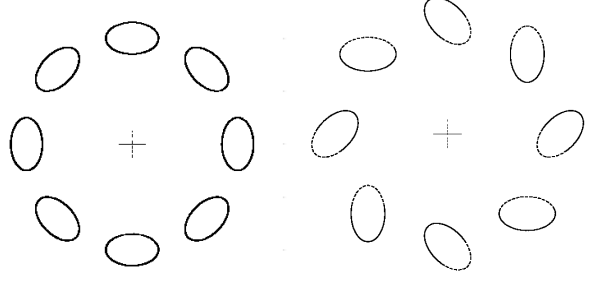
where $\boldsymbol{\beta}(\boldsymbol{\theta}_0)$ and $\boldsymbol{\theta}_0$ are the central positions of the source and image, and $\mathcal{A}$ is the Jacobian matrix

$$\mathcal{A}_{ij}(\boldsymbol{\theta}) = \frac{\partial \beta_i}{\partial \theta_j} = \left(\delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j} \right). \tag{2.15}$$

Generally, the matrix can be decomposed into 3 parts: the trace, a symmetrical traceless matrix and an antisymmetrical matrix,

$$\mathcal{A} = \begin{pmatrix} 1 - \kappa & 0 \\ 0 & 1 - \kappa \end{pmatrix} + \begin{pmatrix} -\gamma_1 & -\gamma_2 \\ -\gamma_2 & \gamma_1 \end{pmatrix} + \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}, \tag{2.16}$$

where the quantities κ, $\gamma_1 + i\gamma_2$, and ω are known as the convergence, (complex) shear and rotation, respectively. They all depend on the redshifts of source and lens. The rotation component is a non-gravitational effect, and is not considered here. The matrix is symmetrical, if the rotation is zero $\omega = 0$. The Jacobian can be simplified in terms of potential $\mathcal{A}_{ij} = \beta_{i,j} = \delta_{ij} - \psi_{,ij}$.

In general, a circular source is mapped to an elliptical image by shear, and a source is enlarged by convergence. Liouvilles theorem guarantees surface brightness conservation by gravitational lensing, thus the observed intensity I at a position $\boldsymbol{\theta}$ is related to the intensity in the source plane I^s via $I(\boldsymbol{\theta}) = I^s(\boldsymbol{\beta}(\boldsymbol{\theta}))$. An image is magnified by a factor equal to the ratio of the integrated intensity distributions $I(\boldsymbol{\theta})$ and $I^s(\boldsymbol{\beta})$, which is the inverse of the determinant of $\mathcal{A}$,

$$\mu = \frac{1}{\det \mathcal{A}} = \frac{1}{(1 - \kappa)^2 - (\gamma_1^2 + \gamma_2^2)}. \tag{2.17}$$

Here μ can be either sign, depending on the parity of the lensed image with respect to the unperturbed image. Regions with different parity are separated by **critical curves** in the lens plane, which are defined by $\det \mathcal{A} = 0$. Critical curves are mapped back onto **caustics** in the source plane. If $\det \mathcal{A} = 0$, the magnification would become infinite, which is impossible in reality. In this case, the linear lensing equation breaks down, thus the magnification by Eq.2.17 is no longer valid here. If an image lies near or crosses the critical curve, it will be highly magnified and distorted, looking like an arc shape (Fig.2.3).

In the regime where $\kappa > 1$, which called strong lensing regime, multiple images may form. Mathematically this means that multiple solutions of the lens equation (2.6) exist. The positions and magnifications of the multiple images depend on the mass distribution of the lens. In other words, using these strong lensing information one can constrain the lens property.
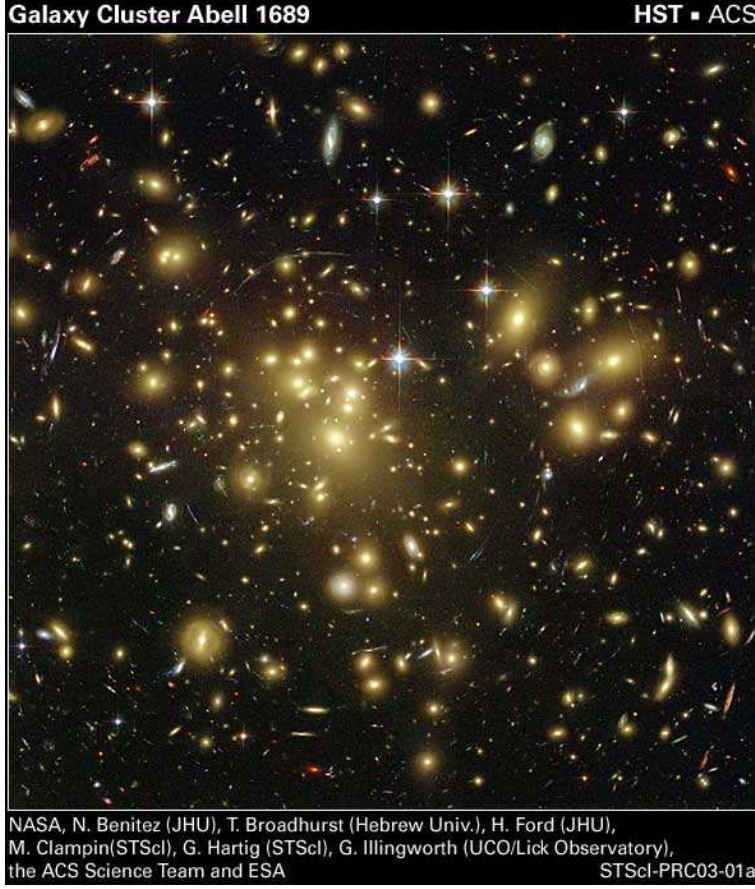
Figure 2.3: One of the most massive galaxy clusters, Abell 1689, made from the Advanced Camera of Hubble Space Telescope. The gravity of the cluster acts as a 2-million-light-year-wide lens in space. Some of the faintest objects in the picture are probably over redshift 6. There are several arcs, which are lensed images of the distance background galaxies.

2.2.1 The relation between convergence and shear

The image distortion in weak lensing yields an estimate of the local shear. If we are capable of measuring the shear field, we will be able to make a degenerate distribution of the projected mass (by 'degenerate', I will explain in Sect. 2.2.3). In order to get a relation between the shear and convergence, we can rewrite them in Fourier space (Kaiser & Squires 1993),

$$\hat{\kappa}(\boldsymbol{l}) = -\frac{1}{2}(l_1^2 + l_2^2)\hat{\psi}(\boldsymbol{l}), \; \hat{\gamma}_1(\boldsymbol{l}) = -\frac{1}{2}(l_1^2 - l_2^2)\hat{\psi}(\boldsymbol{l}), \; \hat{\gamma}_2(\boldsymbol{l}) = -l_1 l_2 \hat{\psi}(\boldsymbol{l}), \qquad (2.18)$$

and obtain

$$\hat{\gamma}(\boldsymbol{l}) = \hat{\gamma}_1(\boldsymbol{l}) + \mathrm{i}\hat{\gamma}_2(\boldsymbol{l}) = \frac{(l_1 + l_2\mathrm{i})^2}{l^2}\hat{\kappa}(\boldsymbol{l}) = \mathrm{e}^{2\mathrm{i}\beta}\hat{\kappa}(\boldsymbol{l}), \qquad (2.19)$$

where β is the polar angle of the Fourier vector $\boldsymbol{l}$. These relations are not defined for $l = 0$, which corresponds to a uniform surface mass density and no shear signal. This relation may be inverted

$$\hat{\kappa}(\boldsymbol{l}) = \frac{(l_1 - \mathrm{i}l_2)^2}{l^2}\hat{\gamma}(\boldsymbol{l}) = \mathrm{e}^{-2\mathrm{i}\beta}\hat{\gamma}(\boldsymbol{l}). \qquad (2.20)$$

The inverse Fourier transformation of Eq.(2.20) follows immediately

$$\kappa(\boldsymbol{\theta}) - \kappa_0 = \frac{1}{\pi}\int_{\mathbb{R}^2} \mathrm{d}^2\theta' \gamma(\boldsymbol{\theta}')D^*(\boldsymbol{\theta} - \boldsymbol{\theta}'), \quad \text{with kernel} \; D(\boldsymbol{\theta}) = \frac{\theta_2^2 - \theta_1^2 - 2\mathrm{i}\theta_1\theta_2}{|\boldsymbol{\theta}|^4}, \qquad (2.21)$$

which means that κ can be determined from the shear only up to an additive constant. Moreover, in principle $\kappa(\boldsymbol{\theta})$ should be real, (an imaginary component of mass does not make real sense). However, noise in real data can produce an imaginary component. In analogy to Eq.(1.71), the power spectra of the convergence and shear are defined and found to be equal, $P_\kappa = P_\gamma$ with

$$\langle \hat{\kappa}(\boldsymbol{l})\hat{\kappa}^*(\boldsymbol{l})\rangle = \langle \hat{\gamma}(\boldsymbol{l})\hat{\gamma}^*(\boldsymbol{l})\rangle = (2\pi)^2 \delta_{\mathrm{D}}(\boldsymbol{l} - \boldsymbol{l}')P_\kappa(l). \tag{2.22}$$

2.2.2 Shear measurements

The local gravitational shear transforms round sources into ellipses. Thus, the shear could be estimated from the shape of the observed galaxies if all galaxies were intrinsically circular. However, they are not, and the shear cannot be measured from a single galaxy (in the weak lensing case). Yet, it is reasonable to assume that the intrinsic orientation of galaxies is random. The local shear therefore can be estimated by averaging over an ensemble of background galaxies.

The shapes of galaxies are irregular, and poorly represented by ellipses. A number of approaches have been put forward to improve the shear measurement (Bridle et al. 2009), mainly brightness moments and shapelets. The shapelets consist of the two-dimensional Cartesian Gauss-Hermite functions, famous as the energy eigenstate of the 2-D quantum harmonic oscillator. A 2-D function $f(\boldsymbol{x})$, such as the image of a galaxy, can thus be decomposed as

$$f(\boldsymbol{\theta}) = \sum_{m,n=0}^{\infty} f_{mn}\, B_{mn}(\boldsymbol{\theta}, \beta), \text{ with coefficients } f_{mn} = \int \mathrm{d}^2\theta\, f(\boldsymbol{\theta})\, B_{mn}(\boldsymbol{\theta}, \beta), \tag{2.23}$$

where

$$B_{mn}(\boldsymbol{\theta}, \beta) = \beta^{-1}\psi_n(\beta^{-1}\theta_1)\psi_m(\beta^{-1}\theta_2). \tag{2.24}$$

Here β is a scale factor chosen for the galaxy, and ψ_n are reduced Hermite polynomials. The shape parameters, like ellipticity can be characterized in terms of the shapelet coefficients (Refregier & Bacon 2003; Massey & Refregier 2005; Kuijken 2006).

Since surface brightness is conserved by lensing, it is convenient to define the brightness moment operation

$$Mom[F(\beta)] = \int \mathrm{d}^2\beta\, F(\beta)\, I^s(\beta) = \int \mathrm{d}^2\theta\, \det\!\mathcal{A}\, F(\beta(\theta))\, I(\theta), \tag{2.25}$$

where $F(\beta)$ is a function of the source coordinate, with the origin of the image plane as the center of light of the image

$$\int d^2\theta\, \theta\, I(\theta) = 0. \tag{2.26}$$

In particular, if setting $F = 1$, one finds that the zeroth-order moment

$$Mom[1] = S_0 = \int \mathrm{d}^2\beta\, I^s(\beta) = \det\!\mathcal{A}_0\, S, \tag{2.27}$$

here S is the flux of the lensed image, so $S = S_0/\det\!\mathcal{A}_0$.

The origin of the coordinates in the source plane is the image of the origin in the lens plane as mapped with the lens equation. But, this does not coincide with the center of light

of the source, which is given by

$$
\begin{aligned}
\bar{\beta} &= \frac{1}{S_0} \int d^2\beta \, \beta \, I^s(\beta) = \frac{1}{S_0} \int d^2\theta \, \det\!\mathcal{A} \, \beta(\theta) \, I(\theta); & (2.28)\\
&= \frac{1}{S_0} \int d^2\theta \, \left[(1-\kappa)^2 - \gamma\gamma^*\right] \left[(1-\kappa)\theta - \gamma\theta^*\right] I(\theta). & (2.29)
\end{aligned}
$$

The linear term in θ vanishes, due to Eq.(2.26), thus $\bar{\beta} \approx 0$ if κ and γ are nearly constants over the image. Therefore, using the linear lens equation, the centroid shift therefore does not affect much in the 2nd order brightness moment of the source

$$
Q_2^s = \frac{1}{S_0} \int d^2\beta \, (\beta - \bar{\beta})^2 \, I^s(\beta); \quad Q_0^s = \frac{1}{S_0} \int d^2\beta \, (\beta - \bar{\beta})(\beta - \bar{\beta})^* \, I^s(\beta). \tag{2.30}
$$

Similarly, the 2nd order brightness moment can be written as

$$
Q_2 = \frac{1}{S} \int d^2\theta \, \theta^2 \, I(\theta); \quad Q_0 = \frac{1}{S} \int d^2\theta \, \theta\theta^* \, I(\theta). \tag{2.31}
$$

Note that if we use Eq.(2.12) to expand the integrand (Eq.2.30), we will obtain the relation between the source and image brightness moments, which is only sufficient to constrain 2 parameters. There are however, 3 parameters need to be solved, i.e. κ, γ_1 and γ_2.

2.2.3 Mass-sheet degeneracy

For a general lens system the source position β and the source size are unknown. This leads to a non-unique determination of the lens surface mass density κ, namely the observation of a gravitational lens system is unchanged if the surface mass density κ is transformed as $\kappa(\theta) \to \kappa_\lambda(\theta) = \lambda\kappa(\theta) + (1-\lambda)$ (Gorenstein et al. 1988). The family κ_λ satisfies the Poisson equation $\Delta\psi_\lambda = 2\kappa_\lambda$ with $\psi_\lambda(\theta) = (1-\lambda)\theta^2/2 + \lambda\psi(\theta)$.

In the case of weak lensing, the shapes of images are invariant under the transformation (Schneider & Seitz 1995). Therefore the shear is not an observable, but only the **reduced shear** $g = \gamma/(1-\kappa)$ is. We can thus rewrite the lensing equation (2.12)

$$
\hat{\beta} \equiv \frac{\beta}{1-\kappa} = \theta - g\theta^*. \tag{2.32}
$$

With this reduced lens equation, the relation between the source and image brightness moments become

$$
\begin{aligned}
Q_2^s &= Q_2 - 2gQ_0 + g^2 Q_2^*;\\
Q_0^s &= -g^* Q_2 + (1 + gg^*)Q_0 - gQ_2^*.
\end{aligned} \tag{2.33}
$$

If we define the complex ellipticity $\chi = Q_2/Q_0$, then we have the relation (Seitz & Schneider 1995)

$$
\chi^s = \frac{\chi - 2g + g^2\chi^*}{1 + gg^* - g^*\chi - g\chi^*}, \tag{2.34}
$$

which under the assumption $\langle\chi^s\rangle = 0$ gives the approximate estimator for the reduced shear $g \approx 1/2\langle\chi\rangle$. In another definition of ellipticity

$$
\epsilon = \frac{1-r}{1+r} \, e^{2i\phi}, \tag{2.35}
$$

where r is axis ratio of elliptical isophotes, and ϕ yields the direction of distortion, the unbiased estimator for the reduced shear is achieved by $g = \langle \epsilon \rangle$ (Seitz & Schneider 1997).

The ambiguity of the lens surface mass density κ is called mass-sheet degeneracy (as mentioned in Sect. 2.2.1). With additional knowledge use of the redshift distribution of the background galaxies, the mass-sheet degeneracy can be weakly broken (Bradač et al. 2004).

2.3   Cosmic shear

The cosmic shear describes the weak lensing by matter inhomogeneities on large-scale structure. The distortions are small, therefore the cosmic shear has to be detected in a statistical way over a large number of high-quality galaxy images. Here I mainly follow Bartelmann & Schneider (2001)

2.3.1   Light propagation

For cosmic shear, the light of background galaxies is distorted continuously along the light path, the thin lens approximation therefore fails (too many 'lens planes' ends up with no 'lens plane'). The propagation of light bundles is described by the geodesic deviation equation. In a weakly inhomogeneous universe, it is assumed that the Newtonian potential is small $|\Phi| \ll c^2$, and varies on scales much smaller than the horizon. For the comoving separation $\boldsymbol{x}$ between two light rays, the propagation equation is

$$\frac{\mathrm{d}^2 \boldsymbol{x}}{\mathrm{d}\chi^2} + K\boldsymbol{x} = -\frac{2}{c^2} \left[\nabla_\perp \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi), \chi) - \nabla_\perp \Phi^{(0)}(\chi) \right], \tag{2.36}$$

where χ is the comoving distance along the light ray, and K is the spatial curvature. $\nabla_\perp$ denotes the comoving gradient perpendicular to the light path. The solution of Eq.(2.36) can be written by the Green's functions

$$\boldsymbol{x}(\boldsymbol{\theta}, \chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' \, f_K(\chi - \chi') \left[\nabla_\perp \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \chi') - \nabla_\perp \Phi^{(0)}(\chi') \right]. \tag{2.37}$$

Without 'lens' the source would be seen at the angular separation $\boldsymbol{\beta} = \boldsymbol{x}/f_K$ from the fiducial ray. Hence, $\boldsymbol{\beta}$ can be interpreted as the unlensed angular position in a hypothetical source plane at comoving distance χ. In analogy with standard lens theory, we can define the Jacobian matrix

$$\mathcal{A}(\boldsymbol{\theta}, \chi) = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \frac{1}{f_K(\chi)} \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} \tag{2.38}$$

$$= \delta_{ij} - \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' \, \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \chi')}{\partial x_i \partial x_k} \, \mathcal{A}_{kj}(\boldsymbol{\theta}, \chi'), \tag{2.39}$$

to describe the locally linearized lens mapping. Expanding $\mathcal{A}$ in power of Φ and keeping the linear terms

$$\mathcal{A}_{ij} = \delta_{ij} - \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' \, \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi(f_K(\chi')\boldsymbol{\theta}, \chi')}{\partial x_i \partial x_j}. \tag{2.40}$$

To linear order, the distortion can be obtained by an integral along the unperturbed ray $\boldsymbol{x} = f_K(\chi)\boldsymbol{\theta}$, corrections are of order Φ^2. This is called Born approximation. In this case,

lensing by the 3−D matter distribution can be treated formally as in normal lensing, with the deflection potential

$$\psi(\boldsymbol{\theta}, \chi) = \frac{2}{c^2} \int_0^\chi \mathrm{d}\chi' \frac{f_K(\chi - \chi')}{f_K(\chi')f_K(\chi)} \Phi(f_K(\chi')\boldsymbol{\theta}, \chi'), \tag{2.41}$$

then $\mathcal{A}_{ij} = \delta_{ij} - \psi_{,ij}$. Correspondingly, we have the effective surface mass density and shear as

$$\kappa = \frac{1}{2}(\psi_{,11} + \psi_{,22}), \quad \gamma = \frac{1}{2}(\psi_{,11} - \psi_{,22}) + \mathrm{i}\psi_{,12}. \tag{2.42}$$

Making use of the 3−D Poisson equation $\nabla^2 \Psi = 3H_0^2 \Omega_\mathrm{M} \delta/(2a)$, we obtain

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{3H_0^2 \Omega_\mathrm{M}}{2c^2} \int_0^\chi \mathrm{d}\chi' \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\delta(f_k(\chi')\boldsymbol{\theta}, \chi')}{a(\chi')}. \tag{2.43}$$

For a source redshift distribution $p_z(z)\mathrm{d}z = p_\chi(\chi)\mathrm{d}\chi$, the effective surface mass density becomes

$$\kappa(\boldsymbol{\theta}) = \int \mathrm{d}\chi\, p_\chi(\chi)\, \kappa(\boldsymbol{\theta}, \chi) = \frac{3H_0^2 \Omega_\mathrm{M}}{2c^2} \int_0^{\chi_\mathrm{h}} \mathrm{d}\chi\, g(\chi)\, f_K(\chi) \frac{\delta(f_K(\chi)\boldsymbol{\theta}, \chi)}{a(\chi)}, \tag{2.44}$$

with the effective source-redshift weighted lens efficiency factor

$$g(\chi) = \int_\chi^{\chi_\mathrm{h}} \mathrm{d}\chi'\, p_\chi(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')}, \tag{2.45}$$

where χ denotes the distance to the horizon.

2.3.2 The E- and B- mode of the shear field

From the relations between the shear, the convergence and the deflection potential, one can derive the relation (Kaiser 1995)

$$\nabla_\mathrm{c}\kappa = \nabla_\mathrm{c}^*\gamma = \mathcal{F}, \tag{2.46}$$

where $\mathcal{F}$ is flexion, which will be discussed in the next chapter, and the complex gradient is defined as

$$\nabla_\mathrm{c} = \frac{\partial}{\partial\theta_1} + \mathrm{i}\frac{\partial}{\partial\theta_2}; \quad \nabla_\mathrm{c}^* = \frac{\partial}{\partial\theta_1} - \mathrm{i}\frac{\partial}{\partial\theta_2}. \tag{2.47}$$

In reality, when measuring $\mathcal{F}$ from data, it will have a non-gradient component due to noise and systematic measurement errors. Moreover, other sources than measurement effects can cause a non-zero curl component of $\mathcal{F}$, e.g. intrinsic shape of galaxies.

One therefore introduces the new quantities κ^E and κ^B (Crittenden et al. 2002), called E- and B- mode of the convergence (in analogy to the electromagnetic field) by

$$\nabla^2\kappa^\mathrm{E} = \nabla_\mathrm{c}\nabla_\mathrm{c}^*\kappa^\mathrm{E} = \nabla \cdot \boldsymbol{F}; \quad \nabla^2\kappa^\mathrm{B} = \nabla \times \boldsymbol{F}, \tag{2.48}$$

where $\boldsymbol{F} = (\mathcal{F}_1, \mathcal{F}_2)$. One can further define the E- and B-mode potentials via the Poisson equation $\nabla^2\psi^{\mathrm{E},\mathrm{B}} = 2\kappa^{\mathrm{E},\mathrm{B}}$, and combine the E- and B-mode for simplification,

$$\psi = \psi^\mathrm{E} + \mathrm{i}\psi^\mathrm{B}; \quad \kappa = \kappa^\mathrm{E} + \mathrm{i}\kappa^\mathrm{B}. \tag{2.49}$$

The complex shear thus can be written as

$$\gamma = \frac{1}{2}(\psi_{,11}^{\mathrm{E}} - \psi_{,22}^{\mathrm{E}}) - \psi_{,12}^{\mathrm{B}} + \mathrm{i}\left[\psi_{,12}^{\mathrm{E}} + \frac{1}{2}(\psi_{,11}^{\mathrm{B}} - \psi_{,22}^{\mathrm{B}})\right]. \tag{2.50}$$

The shear and convergence power spectrum as defined in Eq.(2.22) can be generalized to a field which has both an E- and B-mode. The E-mode, the B-mode and the mixed power spectra can be defined as

$$
\begin{aligned}
\mathrm{EE} &= \langle\hat{\kappa}^{\mathrm{E}}(l)\hat{\kappa}^{\mathrm{E}}(l')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{l}-\boldsymbol{l}')P_{\kappa}^{\mathrm{E}}(l); \\
\mathrm{BB} &= \langle\hat{\kappa}^{\mathrm{B}}(l)\hat{\kappa}^{\mathrm{B}}(l')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{l}-\boldsymbol{l}')P_{\kappa}^{\mathrm{B}}(l); \\
\mathrm{EB} &= \langle\hat{\kappa}^{\mathrm{E}}(l)\hat{\kappa}^{\mathrm{B}}(l')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{l}-\boldsymbol{l}')P_{\kappa}^{\mathrm{EB}}(l),
\end{aligned} \tag{2.51}
$$

and can be derived from the shear correlation function

$$
\begin{aligned}
\langle\hat{\gamma}(l)\hat{\gamma}(l)^*\rangle &= (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{l}-\boldsymbol{l}')[P_{\kappa}^{\mathrm{E}}(l)+P_{\kappa}^{\mathrm{B}}(l)]; \\
\langle\hat{\gamma}(l)\hat{\gamma}(l)\rangle &= (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{l}-\boldsymbol{l}')[P_{\kappa}^{\mathrm{E}}(l)-P_{\kappa}^{\mathrm{B}}(l)+2\mathrm{i}P_{\kappa}^{\mathrm{EB}}(l)].
\end{aligned} \tag{2.52}
$$

There are different sources of the B-mode in the shear field. It can arise from the measurement errors. Also the lensing itself: source redshift clustering and lens-lens coupling which is not accounted for in the Born approximation (Schneider et al. 2002b, 1998). However, these effect are assumed to be small. The intrinsic alignment, which is the intrinsic correlation of galaxy orientations, can show up in both the E- and B-mode shear signal.

2.3.3   Cosmic shear statistics

Since the projected density κ is a projection of δ, it also has to be described as a random field. The convergence power spectrum, which is defined in Eq.(2.22), is related to that of δ through Limber's equation (Limber 1953)

$$P_{\kappa}(l) = \frac{9H_0^4\Omega_{\mathrm{M}}^2}{4c^4}\int_0^{\chi_{\mathrm{h}}}\mathrm{d}\chi\frac{g^2(\chi)}{a^2(\chi)}P_{\delta}\left(\frac{l}{f_K(\chi)},\chi\right). \tag{2.53}$$

The power spectrum P_{κ} yields information about P_{δ}, and is related to actual observable quantities, the shear two-point correlation functions (2PCFs), which have three components

$$\xi_{\pm}(\theta) = \langle\gamma_{\mathrm{t}}(\theta')\gamma_{\mathrm{t}}(\theta'+\theta)\rangle \pm \langle\gamma_{\times}(\theta')\gamma_{\times}(\theta'+\theta)\rangle; \ \ \xi_{\times} = \langle\gamma_{\times}\gamma_{\mathrm{t}}\rangle. \tag{2.54}$$

Because the large-scale structure is statistically homogeneous and isotropic, the 2PCFs only depend on the modulus θ of the vector connecting the two points.

The Fourier transform of the 2PCF of the shear is the shear power spectrum. Since $P_{\gamma} = P_{\kappa}$ (Eq.2.22), the 2PCF can be written in terms of the convergence power spectrum, decomposed into E- and B-mode,

$$
\begin{aligned}
\xi_{\pm}(\theta) &= \frac{1}{2\pi}\int\mathrm{d}l\,l\,J_{0,4}(l\theta)[P_{\kappa}^{\mathrm{E}}(l)\pm P_{\kappa}^{\mathrm{B}}(l)]; \\
\xi_{\times}(\theta) &= \frac{1}{2\pi}\int\mathrm{d}l\,l\,J_4(l\theta)P_{\kappa}^{\mathrm{EB}}(l),
\end{aligned} \tag{2.55}
$$

where J_{ν} is the first kind Bessel function of order ν.

It is important to possess a measurement which separates the E- and the B-mode power spectra, since the first contains cosmological information, while the latter is usually due to systematic effects. E- and B-mode correlation functions, $\xi_{\pm}^{\mathrm{E,B}}$, could be calculated from the above defined 2PCF $\xi_{\pm}$ (Crittenden et al. 2002). However, such calculation require the knowledge of either arbitrarily large or small angular separations, which is not accessible (Schneider et al. 2002a). Other statistics were defined concerning this problem.

The aperture mass M_{ap} was introduced by Kaiser (1994) and Schneider (1996) as a quantity which measures a weighted integral of the local surface mass density κ in an aperture of radius θ, centered at $\boldsymbol{\theta}_0$. Alternatively it can be expressed in terms of the tangential shear γ_{t}, where tangential is with respect to the aperture center $\boldsymbol{\theta}_0$

$$M_{\mathrm{ap}}(\theta, \boldsymbol{\theta}_0) = \int \mathrm{d}^2\theta' U(|\boldsymbol{\theta} - \boldsymbol{\theta}'|)\kappa(\boldsymbol{\theta}') = \int \mathrm{d}^2\theta' Q_\theta(|\boldsymbol{\theta} - \boldsymbol{\theta}'|)\gamma_{\mathrm{t}}(\boldsymbol{\theta}'), \qquad (2.56)$$

where $U(|\boldsymbol{\theta}|)$ is a compensated filter function, i.e. $\int \mathrm{d}\theta\theta U(\theta) = 0$. The fact of using a compensated filter leaves M_{ap} invariant to a constant additive surface mass density. The filter Q is obtained from U by

$$Q(\theta) = \frac{2}{\theta^2} \int_0^\theta \mathrm{d}\theta'\theta' U(\theta') - U(\theta). \qquad (2.57)$$

If we use a function of the form

$$Q(\theta) = \frac{6}{\pi r^2} \frac{\theta^2}{r^2} \left(1 - \left(\frac{\theta^2}{r}\right)^2\right), \qquad (2.58)$$

where r is the radius of an aperture, the second moment or dispersion of $M_{\mathrm{ap}}(\theta)$ is related to the power spectrum through (Schneider et al. 1998)

$$\langle M_{\mathrm{ap}}^2 \rangle(\theta) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, P_\kappa(l) W_{\mathrm{ap}}(\theta l), \quad \text{with}\ \ W_{\mathrm{ap}}(\eta) = \frac{576 J_4^2(\eta)}{\eta^4}. \qquad (2.59)$$

$\langle M_{\mathrm{ap}}^2 \rangle$ has the property of separating the E- and B-mode power spectrum. To see this, we can define a complex aperture mass

$$M(\theta, \boldsymbol{\theta}_0) = M_{\mathrm{ap}}(\theta, \boldsymbol{\theta}_0) + \mathrm{i}M_\perp(\theta, \boldsymbol{\theta}_0), \qquad (2.60)$$

where

$$M_\perp(\theta, \boldsymbol{\theta}_0) = \int_0^\infty \mathrm{d}^2\theta' Q_\theta(|\boldsymbol{\theta} - \boldsymbol{\theta}'|)\gamma_\times(\boldsymbol{\theta}'). \qquad (2.61)$$

With the complex aperture mass, one can define two second-order quantities $\langle MM \rangle(\theta)$ and $\langle MM^* \rangle(\theta)$, where the ensemble average is replaced by a spatial average over aperture centers $\boldsymbol{\theta}_0$. Using Eq.(2.56), they write as

$$\begin{aligned} \langle MM^* \rangle(\theta) &= \langle M_{\mathrm{ap}}^2 \rangle(\theta) + \langle M_\perp^2 \rangle(\theta) = \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, [P_\kappa^{\mathrm{E}}(l) + P_\kappa^{\mathrm{B}}(l)]\hat{U}^2(\theta l); \\ \langle MM \rangle(\theta) &= \langle M_{\mathrm{ap}}^2 \rangle(\theta) - \langle M_\perp^2 \rangle(\theta) + 2\mathrm{i}\langle M_{\mathrm{ap}} M_\perp \rangle(\theta) \\ &= \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, [P_\kappa^{\mathrm{E}}(l) - P_\kappa^{\mathrm{B}}(l) + 2\mathrm{i}P_\kappa^{\mathrm{EB}}(l)]\hat{U}^2(\theta l). \end{aligned} \qquad (2.62)$$

Combining these expressions, one gets

$$
\begin{aligned}
\langle M_{\mathrm{ap}}^2 \rangle(\theta) &= \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, P_\kappa^{\mathrm{E}}(l) \hat{U}^2(\theta l); \\
\langle M_\perp^2 \rangle(\theta) &= \frac{1}{2\pi} \int_0^\infty \mathrm{d}l\, l\, P_\kappa^{\mathrm{B}}(l) \hat{U}^2(\theta l).
\end{aligned}
\tag{2.63}
$$

The equations show that $\langle M_{\mathrm{ap}}^2 \rangle$ depends only on the E-mode and $\langle M_\perp^2 \rangle$ on the B-mode power spectrum. Here $\hat{U}(\theta l)$ is the Fourier transform of the filter function U_θ. The aperture mass statistic is a localized measure of the power spectrum since the filter function $\hat{U}^2(\theta l)$ is narrow. For the case of E-mode only, the aperture measures can be expressed as finite integrals over the correlation functions,

$$
\langle M_{\mathrm{ap}}^2 \rangle(\theta) = \frac{1}{2} \int_0^{2\theta} \frac{\mathrm{d}\theta'\, \theta'}{\theta^2} \left[\xi_+(\theta') T_+\left(\frac{\theta'}{\theta}\right) + \xi_-(\theta') T_-\left(\frac{\theta'}{\theta}\right) \right];
\tag{2.64}
$$
$$
\langle M_\perp^2 \rangle(\theta) = \frac{1}{2} \int_0^{2\theta} \frac{\mathrm{d}\theta'\, \theta'}{\theta^2} \left[\xi_+(\theta') T_+\left(\frac{\theta'}{\theta}\right) - \xi_-(\theta') T_-\left(\frac{\theta'}{\theta}\right) \right],
$$

where the two function $T_\pm$ are given in Schneider et al. (2002b). The aperture mass method for E- and B-mode decomposition still require the 2PCF to be known down to arbitrary small angular separations (Eq.2.64). In addition, measures in apertures are not very effective, since large areas of the sky containing bright stars or foreground galaxies have to be omitted in order not to bias the result.

Schneider & Kilbinger (2007); Eifler et al. (2010); Fu & Kilbinger (2010) introduced a new method to perform an E- and B-mode decomposition using a shear 2PCF measured on finite interval, which is called **Ring statistics**.

The ring statistics only involve correlations between points in an annulus $\zeta_1 \leq \theta \leq \zeta_2$ and points in a concentric annulus $\zeta_3 \leq \theta \leq \zeta_4$ (see Fig. 2.4). The rings are non-overlapping, i.e. $\zeta_i < \zeta_j$ if $i < j$. The ring statistics correlator are thus calculated from shear 2PCF covering the range $[\eta\Psi = \zeta_3 - \zeta_2,\ \Psi = \zeta_2 + \zeta_4]$, and are defined as

$$
\langle RR_{\mathrm{E}} \rangle(\Psi, \eta) = \int_{\eta\Psi}^\Psi \frac{\mathrm{d}\theta}{2\theta} [\xi_+(\theta) Z_+(\theta, \eta) + \xi_-(\theta) Z_-(\theta, \eta)];
$$
$$
\langle RR_{\mathrm{B}} \rangle(\Psi, \eta) = \int_{\eta\Psi}^\Psi \frac{\mathrm{d}\theta}{2\theta} [\xi_+(\theta) Z_+(\theta, \eta) - \xi_-(\theta) Z_-(\theta, \eta)],
\tag{2.65}
$$

where the functions $Z_\pm$ are defined in Schneider & Kilbinger (2007) and η (with $\eta < 1$) quantifies the separation between the outer and inner rings. $\langle RR_{\mathrm{E}} \rangle$ is related to the E-mode power spectrum through

$$
\langle RR_{\mathrm{E}} \rangle(\Psi, \eta) = \int_0^\infty \frac{\mathrm{d}l\, l}{2\pi} P_{\mathrm{E}}(l) W_{\mathrm{E}}(l\Psi, \eta)
\tag{2.66}
$$

with

$$
W_{\mathrm{E}}(l\Psi, \eta) = \int_{\eta\Psi}^\Psi \frac{\mathrm{d}\theta}{2\theta} [J_0(l\theta) Z_+(\theta, \eta) + J_4(l\theta) Z_-(\theta, \eta)].
\tag{2.67}
$$

The covering range varies with (Ψ, η). An optimal choice of Ψ and η significantly improves the ring statistics's S/N (Eifler et al. 2010).
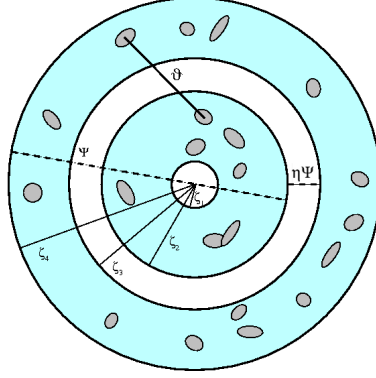
Figure 2.4: Illustration of the Ring statistics: The 2PCF is calculated for each galaxy in the inner ring with all galaxies in the outer ring. Figure from Eifler et al. (2010).

2.4 Galaxy-Galaxy lensing

The phenomenon of image distortion of background galaxies by a foreground galaxy is called galaxy-galaxy lensing (GGL). The first detection of GGL was reported by Brainerd et al. (1996). Its main application is the study of dark matter haloes of galaxies. The halo properties of galaxies can provide a hint on the formation and evolution of clusters and galaxies.

Individual galaxies are not sufficiently massive to produce a significant distortion of the images of background galaxies. Therefore, a population of galaxies is needed, and their statistical properties are studied. The sample of galaxies is split into fore- and background subsamples. The ellipticities of the background galaxies are expected to be tangentially aligned with respect to each of the lensing foreground galaxies. The tangential shear, as function of angular separation from the lens, yields an estimator, that can be used to measure some galaxy halos properties.

If we define the fractional galaxy number density $\kappa_{\rm g} = n_{\rm g}/\bar{n} - 1$, where $n_{\rm g}$ is the galaxy number density and $\bar{n}$ the mean number density on the sky, the mean tangential/cross shear at separation θ can be written as (Hoekstra et al. 2002a)

$$\langle \gamma_{\rm t,\times}(\theta) \rangle = \langle \gamma_{\rm t,\times}(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \kappa_{\rm g}(\boldsymbol{\theta}_0) \rangle, \tag{2.68}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$ are the source and lens galaxy positions respectively, $\theta = |\boldsymbol{\theta} - \boldsymbol{\theta}_0|$, and the ensemble average is taken over the fore- and background pairs. Since $\gamma_{\rm t} + {\rm i}\gamma_\times = -\gamma {\rm e}^{-2{\rm i}\phi}$, where ϕ is the polar angle of γ, the GGL relates to the statistics of both the underlying galaxy and dark matter density fields

$$
\begin{aligned}
\langle \gamma_{\rm t}(\theta) \rangle + {\rm i}\langle \gamma_\times(\theta) \rangle &= -{\rm e}^{-2{\rm i}\phi} \langle \gamma(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \kappa_{\rm g}(\boldsymbol{\theta}_0) \rangle \\
&= -{\rm e}^{-2{\rm i}\phi} \int \frac{{\rm d}^2 l}{(2\pi)^2} \int \frac{{\rm d}^2 l'}{(2\pi)^2} \, {\rm e}^{-{\rm i}\boldsymbol{\theta}\cdot\boldsymbol{l}} \langle \hat{\gamma}(\boldsymbol{l})\hat{\kappa}_{\rm g}(\boldsymbol{l}') \rangle \\
&= -{\rm e}^{2{\rm i}(\beta-\phi)} \int \frac{{\rm d}^2 l}{(2\pi)^2} \int \frac{{\rm d}^2 l'}{(2\pi)^2} \, {\rm e}^{-{\rm i}\boldsymbol{\theta}\cdot\boldsymbol{l}} \langle \hat{\kappa}(\boldsymbol{l})\hat{\kappa}_{\rm g}(\boldsymbol{l}') \rangle, \tag{2.69}
\end{aligned}
$$

where in the last step, the Kaiser-Squires relation (Eq.2.20) is used. If we define the cross power spectrum of the galaxy and matter density field $P_{\kappa {\rm g}}(l) = (2\pi)^2 \delta_{\rm D}(\boldsymbol{l} - \boldsymbol{l}')\langle \hat{\kappa}(\boldsymbol{l})\hat{\kappa}_{\rm g}(\boldsymbol{l}') \rangle$, this becomes

$$\langle \gamma_{\rm t}(\theta) \rangle + {\rm i}\langle \gamma_\times(\theta) \rangle = -\int_0^\infty \frac{{\rm d}l\, l}{(2\pi)^2} \int_0^{2\pi} {\rm d}\alpha \, {\rm e}^{{\rm i}l\theta\cos\alpha} \left[\cos 2\alpha + {\rm i}\sin 2\alpha\right] P_{\kappa {\rm g}}(l), \tag{2.70}$$

The imaginary part, which corresponds to $\langle \gamma_\times(\theta) \rangle$, vanishes in the integral since it is an odd function. The remaining terms yield an expression relating $P_{\kappa g}$ with the tangential shear only

$$\langle \gamma_t(\theta) \rangle = \frac{1}{2\pi} \int dl\, l J_2(\theta l) P_{\kappa g}(l), \tag{2.71}$$

where $J_2(\theta l)$ is a Bessel function. The GGL is a measure for the cross-correlation of galaxies and dark matter distribution.

The GGL signal is obtained by averaging over many galaxies. For the halos of similar galaxy type, one expects to have similar properties. One can try to obtain a model for the surface mass density $\Sigma(\theta)$ associated with a lens galaxy. The mean tangential shear measured on the circle of radius θ can be related to average properties of the foreground halo through (Schneider et al. 2006)

$$\langle \gamma_t(\theta) \rangle = \bar{\kappa}(\theta) - \kappa(\theta), \tag{2.72}$$

where $\bar{\kappa}(\theta)$ is the averaged convergence inside the circle and $\kappa(\theta)$ is the mean convergence on the circle. The convergence is a dimensionless surface mass density, and depends on the redshifts of source and lens. Thus when measuring GGL signal using the ellipticities of galaxies, without redshift information, it is only possible to obtain a surface mass density averaged over the fore- and background galaxies,

$$\langle \gamma_t(\theta) \rangle = \int dz_d\, p_d(z_d) \int dz_s\, p_s(z_s) \frac{\bar{\Sigma}(z_d, \theta) - \Sigma(z_d, \theta)}{\Sigma_{crit}(z_s, z_d)}, \tag{2.73}$$

where $p_d(z_d)$ and $p_s(z_s)$ are the fore- and background galaxy redshift distributions respectively.

Chapter 3

Weak Lensing Flexion

3.1 Lensing Flexion

3.1.1 Higher-order image distortions

As early in 1988 (Fort et al. 1988; Fort & Mellier 1994) the strongly distorted images of background sources, namely the arclets, have already been observed. In these cases, the locally linearized lens equation (Eq.2.12) is not sufficient to describe this phenomena, since the images are not small compared to the length-scale over which the shear varies. On the other hand, Goldberg & Natarajan (2002) have shown that valuable further information is available from the skewedness and arciness of the source images. In Goldberg & Bacon (2005) (GB05 hereafter), this effect has been introduced as **flexion**, and describes the second-order gravitational lensing mapping distortion. However, in GB05 only one flexion component is considered, which provides an incomplete description. In Bacon et al. (2006) (BGRT hereafter), the second flexion component is studied. Moreover, the Twist and Turn effects were recognized by Bacon & Schäfer (2009), and complete the second-order gravitational lensing image distortion set. In addition, a related approach using 'sextupole lensing' has been treated by Irwin & Shmakova (2006, 2005).

Arclets are the direct application for flexion. In general, flexion responds to small-scale variations in the gravitational potential, and therefore can be used for studying small-scale mass distribution, e.g. galaxy-galaxy lensing and substructure within the cluster.

In this chapter, I will introduce the basic theory of weak lensing flexion, and some measurement issues related to flexion. Most results in this chapter, except the subsection of shapelet can be found from Schneider & Er (2008).

3.1.2 Complex flexion notation

For a two dimensional vector $\boldsymbol{x} = (x_1, x_2)$, we define the complex number $\boldsymbol{x} = x_1 + \mathrm{i}x_2$. Under rotations of the coordinate system by an angle φ, $\boldsymbol{x}$ gets multiplied by the phase factor $\mathrm{e}^{-\mathrm{i}\varphi}$. A quantity is said to be spin-N if it has the same value after rotating $2\pi/N$. The product of spin-A and spin-B quantity is spin-(A+B), and the product of spin-A and spin-B^* is spin-(A-B).

The lens equation $\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\psi(\boldsymbol{\theta})$ in locally linear expansion can be written as $\beta_i = \theta_i - \psi_{,ij}\theta_j$, here summation over repeated indices is implied. Now we generalize this equation to a second-order local expansion, which in Cartesian coordinates reads $\beta_i = \theta_i - \psi_{,ij}\theta_j -$
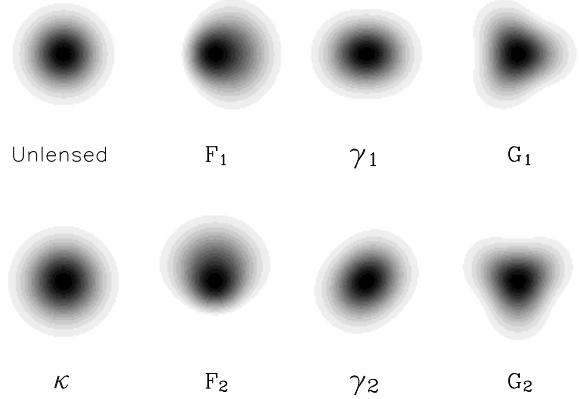
Figure 3.1: Weak lensing image distortions. Here an unlensed Gaussian galaxy has been distorted with convergence, shear and two flexions. Convergence is a spin-0 quantity, first flexion is spin-1, shear is spin-2 and second flexion is spin-3. Figure is from BGRT.

$\psi_{,ijk}\theta_j\theta_k/2$, the third-order derivatives of ψ are related to the gradient of κ and γ. To write these derivatives in complex form, we use the complex gradient (Eq.2.47),

$$\nabla_{\mathrm{c}}\kappa = \frac{1}{2}\left[\psi_{,111} + \psi_{,122} + \mathrm{i}\left(\psi_{,112} + \psi_{,222}\right)\right] \; ; \; \nabla_{\mathrm{c}}\gamma = \frac{1}{2}\left[\psi_{,111} - 3\psi_{,122} + \mathrm{i}\left(3\psi_{,112} - \psi_{,222}\right)\right] \; . \tag{3.1}$$

The complex formalism provides a neat way to generalize the third-order derivatives of the deflection potential. The differential operator ∇_{c} turns a spin-n field into a spin-$(n+1)$ field, whereas ∇_{c}^* reduces the spin by one unit. The derivatives can be expressed as a spin-3 field $\mathcal{G} \equiv \nabla_{\mathrm{c}}\gamma$ and a spin-1 field $\mathcal{F} \equiv \nabla_{\mathrm{c}}^*\gamma$, where we introduced the usual notation for the two flexion quantities. The second-order lens equation in our complex notation then reads

$$\beta = (1-\kappa)\theta - \gamma\theta^* - \frac{1}{4}\mathcal{F}^*\,\theta^2 - \frac{1}{2}\mathcal{F}\,\theta\theta^* - \frac{1}{4}\mathcal{G}\,(\theta^*)^2 \; . \tag{3.2}$$

Since this is no longer a linear equation, a source at β may have more than one image. In fact, up to four images of a source can be obtained, as can be seen for the special case of $\gamma = \mathcal{F} = 0$ and by placing the source at $\beta = 0$. If we set $\mathcal{G} = |\mathcal{G}|\mathrm{e}^{3\mathrm{i}\zeta}$, besides one solution $\theta = 0$, there are three more at $\theta = 4(1-\kappa)/|\mathcal{G}|\,\mathrm{e}^{\mathrm{i}\varphi}$, with $\varphi = \zeta$, $\varphi = \zeta + 2\pi/3$ and $\varphi = \zeta + 4\pi/3$. The occurrence of these solutions lies in the fact that $\mathcal{G}$ is a spin-3 quantity. We shall later need the Jacobian determinant $\det\mathcal{A}$ of the lensing equation, which is

$$\begin{aligned}
\det\mathcal{A} &= (1-\kappa)^2 - \gamma\gamma^* + \boldsymbol{\theta}\cdot\nabla\left[(1-\kappa)^2 - \gamma\gamma^*\right] + \mathcal{O}(\theta^2) \\
&= (1-\kappa)^2 - \gamma\gamma^* \\
&\quad -\theta\left[(1-\kappa)\mathcal{F}^* + \frac{\gamma^*\mathcal{F} + \gamma\,\mathcal{G}^*}{2}\right] - \theta^*\left[(1-\kappa)\mathcal{F} + \frac{\gamma^*\mathcal{G} + \gamma\,\mathcal{F}^*}{2}\right] + \mathcal{O}(\theta^2) \; ,
\end{aligned} \tag{3.3}$$

where the first expression is just the zero- and first-order Taylor expansion of the Jacobian around the origin, and in the second step we made use of the relation $\boldsymbol{\theta}\cdot\nabla = (\theta\nabla_{\mathrm{c}}^* + \theta^*\nabla_{\mathrm{c}})/2$. We point out that Eq.(3.3) is only the first-order expansion; the full Jacobian of Eq.(**??**) contains quadratic terms in θ. In Sect. 3.3, we consider this issue and the critical curves and caustics of the lens equation.

However, the study of the Jacobian $\mathcal{A}$ is incomplete, since there is an asymmetric component of linear Jacobian (Eq.2.16), i.e. rotation w. The derivative of w which is the second order rotation effect is define as turn

$$C \equiv 2\nabla_{\mathrm{c}}w. \tag{3.4}$$

This new distortion mode C describes how the amount of image rotation in the Jacobian varies across the object. In addition, the second order distortion could have 8 components (derivatives of 4 components with respect to θ_1 and θ_2), whereas $\mathcal{F}$, $\mathcal{G}$ and C provide 6. The rest components, which are named twist, do not occur independently due to the symmetry. Nevertheless, rotation, turn and twist effects are non-gravitational distortion mode and not measurable. Thus they are not discussed in the following. For more details, the reader is referred to Bacon & Schäfer (2009).

3.1.3 Reduced flexion

Similar to the weak lensing shear measurement, the flexion also suffers from the mass-sheet degeneracy, i.e. the observable of a gravitational lens system are unchanged if the surface mass density κ is transformed as $\kappa(\boldsymbol{\theta}) \rightarrow \kappa'(\boldsymbol{\theta}) = \lambda\kappa(\boldsymbol{\theta}) + (1 - \lambda)$. In fact, since we expect that the most promising applications of flexion will come from situations where κ is not much smaller than unity, the distinction between shear and reduced shear is more important for flexion than for the usual weak lensing applications. Hence, at best we can expect from higher-order shape measurements to obtain an estimate for the reduced shear and its derivatives. For this reason, we define the **reduced flexion** (Schneider & Er 2008),

$$G_1 \equiv \nabla_c^* g = \frac{\mathcal{F} + g\mathcal{F}^*}{(1 - \kappa)} \; ; \quad G_3 \equiv \nabla_c g = \frac{\mathcal{G} + g\mathcal{F}}{(1 - \kappa)} \; . \tag{3.5}$$

The derivatives of the reduced shear $G_{1,3}$ are those quantities we can actually observe. The mass-sheet transformation is equivalent to an isotropic scaling of the source plane coordinates. Hence, we divide Eq.(3.2) by $(1 - \kappa)$ to obtain a 'reduced lens equation'

$$\hat{\beta} \equiv \frac{\beta}{(1 - \kappa)} = \theta - g\theta^* - \Psi_1^* \theta^2 - 2\Psi_1 \theta\theta^* - \Psi_3 (\theta^*)^2 \text{ with } \Psi_1 = \frac{1}{4}\frac{\mathcal{F}}{(1 - \kappa)} \; ; \quad \Psi_3 = \frac{1}{4}\frac{\mathcal{G}}{(1 - \kappa)} \; . \tag{3.6}$$

The expression for $\mathcal{F}/(1-\kappa)$ in terms of the reduced shear and its derivatives has been derived by Kaiser (1995); in our notation it reads

$$\frac{\mathcal{F}}{(1 - \kappa)} \equiv -\nabla_c \ln(1 - \kappa) = \frac{G_1 - gG_1^*}{1 - gg^*} \Rightarrow \Psi_1 = \frac{G_1 - gG_1^*}{4(1 - gg^*)} \; . \tag{3.7}$$

The expression for the derivative of γ in terms of the reduced shear can be easily obtained from differentiating the definition $\gamma = (1 - \kappa)g$,

$$\frac{\nabla_c \gamma}{(1 - \kappa)} = \frac{\mathcal{G}}{(1 - \kappa)} = G_3 - g\frac{\nabla_c \kappa}{(1 - \kappa)} = G_3 - \frac{g(G_1 - gG_1^*)}{1 - gg^*} \Rightarrow \Psi_3 = \frac{G_3}{4} - \frac{g(G_1 - gG_1^*)}{4(1 - gg^*)} \; . \tag{3.8}$$

The Jacobian determinant $\det\hat{\mathcal{A}}$ of the mapping between the image position θ and the rescaled source position $\hat{\beta}$ then becomes

$$\det\hat{\mathcal{A}} = \frac{\det\mathcal{A}}{(1 - \kappa)^2} = 1 - gg^* - \eta^*\theta - \eta\theta^*,$$

$$\text{where} \quad \eta = \nabla_c^* g - \frac{g(\nabla_c^* g)^*}{2} + \frac{g^*\nabla_c g}{2} = G_1 - \frac{gG_1^*}{2} + \frac{g^*G_3}{2} \tag{3.9}$$

Again, Eq.(3.9) is valid only to linear order in θ. A similar equation for the determinant was obtained in Okura et al. (2007), but they only consider the case of $|g| \ll 1$; this has also consequences for the relations between source and image brightness moments, to be derived further below.

3.2   E/B mode flexion

Flexion has a total of four components, namely the real and imaginary parts of $\mathcal{F}$ and $\mathcal{G}$. A measurement of flexion will thus yield four components, and we might ask whether these components are independent. We recall a similar situation in shear measurements. The shear has two components; on the other hand, the shear is defined as second partial derivatives of the deflection potential, which is a single scalar field. Therefore, the two shear components cannot be mutually independent if they are due to a gravitational lensing signal. We have already seen that the measured shear is not guaranteed to satisfy the condition that the two shear components can be derived from a single scalar deflection potential due to noise and systematic measurement errors. Therefore, one has introduced the notion of E- and B-modes in shear measurement. Formally, the E- and B-mode decomposition can be written in terms of a complex deflection potential $\psi(\theta) = \psi^{\mathrm{E}}(\theta) + \mathrm{i}\psi^{\mathrm{B}}(\theta)$, and thus the surface mass density and the shear (Eqs.2.49 and 2.50).

The distinction between E- and B-mode shear can be obtained by considering second partial derivatives of the shear components. Taking the derivative of Eq.(2.50), one obtains

$$
\begin{aligned}
\mathcal{F} = \nabla_{\mathrm{c}}^* \gamma &= (1/2)\left(\psi^{\mathrm{E}}_{,111} + \psi^{\mathrm{E}}_{,122} - \psi^{\mathrm{B}}_{,112} - \psi^{\mathrm{B}}_{,222}\right) + (\mathrm{i}/2)\left(\psi^{\mathrm{E}}_{,112} + \psi^{\mathrm{E}}_{,222} + \psi^{\mathrm{B}}_{,111} + \psi^{\mathrm{B}}_{,122}\right) \\
&= \kappa^{\mathrm{E}}_{,1} - \kappa^{\mathrm{B}}_{,2} + \mathrm{i}\left(\kappa^{\mathrm{E}}_{,2} + \kappa^{\mathrm{B}}_{,1}\right)\ ,
\end{aligned}
\tag{3.10}
$$

which can be expressed in more compact form as

$$
\mathcal{F} = \nabla_{\mathrm{c}}\left(\kappa^{\mathrm{E}} + \mathrm{i}\kappa^{\mathrm{B}}\right) = \nabla_{\mathrm{c}}\kappa\ .
\tag{3.11}
$$

A further derivative yields for the components

$$
\mathcal{F}_{1,1} = \kappa^{\mathrm{E}}_{,11} - \kappa^{\mathrm{B}}_{,12}\ ;\quad \mathcal{F}_{1,2} = \kappa^{\mathrm{E}}_{,12} - \kappa^{\mathrm{B}}_{,22}\ ;\quad \mathcal{F}_{2,1} = \kappa^{\mathrm{E}}_{,12} + \kappa^{\mathrm{B}}_{,11}\ ;\quad \mathcal{F}_{2,2} = \kappa^{\mathrm{E}}_{,22} + \kappa^{\mathrm{B}}_{,12}\ .
\tag{3.12}
$$

However, it is easier to consider directly the complex derivative of $\mathcal{F}$, from which we obtain

$$
\nabla_{\mathrm{c}}^* \mathcal{F} = \nabla_{\mathrm{c}}^* \nabla_{\mathrm{c}}^* \gamma = \nabla^2\left(\kappa^{\mathrm{E}} + \mathrm{i}\kappa^{\mathrm{B}}\right) = \mathcal{F}_{1,1} + \mathcal{F}_{2,2} + \mathrm{i}\left(\mathcal{F}_{2,1} - \mathcal{F}_{1,2}\right)\ .
\tag{3.13}
$$

Thus, if the shear field is a pure E-mode field, $\nabla_{\mathrm{c}}^* \nabla_{\mathrm{c}}^* \gamma$ is real. An imaginary part of $\nabla_{\mathrm{c}}^* \nabla_{\mathrm{c}}^* \gamma$ is due to a B-mode field. This then yields the local distinction between E- and B-mode shear.

Since the flexion has four components, whereas the lens can be described by a single scalar field, we expect that there are three constraint relations a flexion field has to satisfy if it is due to a lensing potential. In fact, even if we leave the shear field arbitrary (that is, even if we allow it to be composed of E- and B-modes), then we expect two constraint equations, since the flexion field has two components more than the shear field. These constraint equations are easy to obtain. First, if the flexion field is due to a shear field, then we have

$$
\nabla_{\mathrm{c}} \nabla_{\mathrm{c}}^* \gamma = \nabla_{\mathrm{c}}^* \nabla_{\mathrm{c}} \gamma \quad \rightarrow \quad \mathcal{H} := \nabla_{\mathrm{c}} \mathcal{F} - \nabla_{\mathrm{c}}^* \mathcal{G} = 0\ ,
\tag{3.14}
$$

where we defined the spin-2 quantity $\mathcal{H}$. It may describe contributions to the flexion which are not caused by a shear field, such as due to noise, intrinsic source alignments or higher-order terms (e.g. lens-lens coupling) in the propagation equation for light bundles. As a spin-2 field, a non-zero $\mathcal{H}$ can be decomposed into its E- and B-modes. If $\mathcal{H} \equiv 0$, then the spin-3 flexion $\mathcal{G}$ is completely determined by the spin-1 flexion $\mathcal{F}$ up to an additive constant, as can be best seen in Fourier space, for which Eq.(3.14) yields $\hat{\mathcal{G}}(\ell) = -\mathrm{i}\hat{\gamma}(\ell)\,\ell = (\ell/\ell^*)\hat{\mathcal{F}}(\ell)$. Second, if the

flexion field is solely caused by a gravitational lens effect, i.e., by a pure E-mode shear field, then $\nabla_c^* \mathcal{F}$ is real, i.e.,

$$\mathcal{F}_i := \nabla_c^* \mathcal{F} - \nabla_c \mathcal{F}^* = 0 . \tag{3.15}$$

Thus, flexion from a pure E-mode shear field is characterized by the three constraint equations $\mathcal{H} \equiv 0$ and $\mathcal{F}_i \equiv 0$, where the former is a two-component equation.

Turning now to the reduced flexion, the compatibility equations can be obtained as follows. First, if the flexion is due to a shear field, we have

$$H := \nabla_c G_1 - \nabla_c^* G_3 = 0 , \tag{3.16}$$

as follows from the definition Eq.(3.5) of the two flexion components in terms of the reduced shear. Again, if this equation is satisfied, G_3 is completely determined by G_1, up to an additive constant (in Fourier space, $\hat{G}_3(l) = (l/l^*)\hat{G}_1(l)$). Second, if the flexion is caused by a pure E-mode shear, i.e., if the shear is due to a real surface mass density, then we employ the quantity $\ln(1 - \kappa)$, which is real and invariant under mass-sheet transformations, up to an additive constant. Therefore, $K_2 \equiv -\nabla_c^* \nabla_c \ln(1 - \kappa)$ must be real

$$
\begin{aligned}
K_2 &= \nabla_c^* \left(\frac{\nabla_c \kappa}{1 - \kappa} \right) = \nabla_c^* \left[\frac{1}{1 - gg^*} (G_1 - G_1^* g) \right] \\
&= \frac{[\nabla_c^* G_1 - g(\nabla_c G_1)^*]}{1 - gg^*} + \frac{(G_1^2 g^* + g G_1 G_3^* - G_1 G_1^* - g^2 G_1^* G_3^*)}{(1 - gg^*)^2} ,
\end{aligned}
\tag{3.17}
$$

so that a flexion coming from an E-mode shear field satisfies $K_2 = K_2^*$.

3.2.1 The axially-symmetric case

To illustrate these compatibility relations, an example of axi-symmetric flexion field is shown here. For that, we introduce polar coordinates (θ, φ); hence, in this subsection only, θ is the radial coordinate, not a complex number. The gradient operators then become

$$\nabla_c = e^{i\varphi} \left(\frac{\partial}{\partial \theta} + \frac{i}{\theta} \frac{\partial}{\partial \varphi} \right) ; \quad \nabla_c^* = e^{-i\varphi} \left(\frac{\partial}{\partial \theta} - \frac{i}{\theta} \frac{\partial}{\partial \varphi} \right) .$$

We first assume that the flexion derives from a shear field, which in the axi-symmetric case takes the form $\gamma(\theta, \varphi) = \Gamma(\theta) e^{2i\varphi}$. In the case of a pure E-mode shear, $\Gamma(\theta)$ is real, whereas for pure B-modes, Γ is imaginary. The two flexions then read

$$\mathcal{F}(\theta, \varphi) = \nabla_c^* \gamma = e^{i\varphi} \left(\frac{d\Gamma}{d\theta} + \frac{2\Gamma}{\theta} \right) ; \quad \mathcal{G}(\theta, \varphi) = \nabla_c \gamma = e^{3i\varphi} \left(\frac{d\Gamma}{d\theta} - \frac{2\Gamma}{\theta} \right) . \tag{3.18}$$

A further differentiation then yields the result that

$$\nabla_c \mathcal{F} = e^{2i\varphi} \left(\frac{d^2\Gamma}{d\theta^2} + \frac{1}{\theta} \frac{d\Gamma}{d\theta} - \frac{4\Gamma}{\theta^2} \right) = \nabla_c^* \mathcal{G} ,$$

so that the function $\mathcal{H}$ defined in Eq.3.14 vanishes, $\mathcal{H} \equiv 0$.

If flexion does not derive from a shear field, then $\mathcal{H} \neq 0$; for example, this is the case if $\nabla_c \mathcal{F} = -\nabla_c^* \mathcal{G} = \mathcal{H}/2$, which we shall take in the following. Owing to their spin properties, we can write

$$\mathcal{F}(\theta, \varphi) = F(\theta) e^{i\varphi} ; \quad \mathcal{H}(\theta, \varphi) = H(\theta) e^{2i\varphi} ; \quad \mathcal{G}(\theta, \varphi) = G(\theta) e^{3i\varphi} ,$$
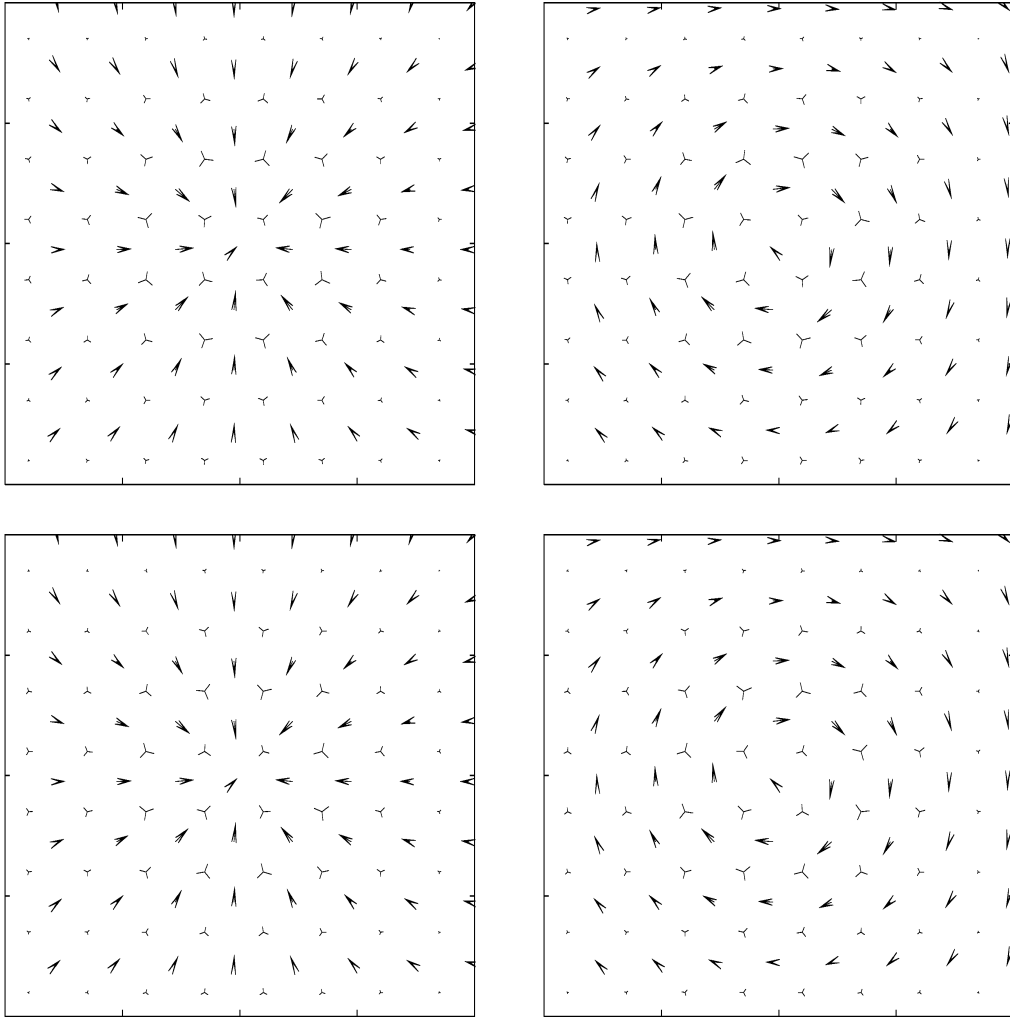
Figure 3.2: The four different flexion fields discussed in the text. The upper left (right) panel shows the flexion corresponding to an axially-symmetric E-mode (B-mode) shear field, where arrows indicate the spin-1 flexion and the skeletons the spin-3 flexion component. In the lower left (right) panel, the flexion fields are displayed which are not due to a shear field, but a non-zero E-mode (B-mode) $\mathcal{F}$ field.

which then leads to the differential equations

$$\frac{\mathrm{d}F}{\mathrm{d}\theta} - \frac{F}{\theta} = \frac{H}{2} \; ; \quad \frac{\mathrm{d}G}{\mathrm{d}\theta} + \frac{3G}{\theta} = -\frac{H}{2}$$

with the solutions

$$F(\theta) = \frac{F_0\theta}{\theta_0} + \frac{\theta}{2}\int_{\theta_0}^{\theta}\mathrm{d}\theta'\,\frac{H(\theta')}{\theta'} \; ; \quad G(\theta) = \frac{G_0\theta_0^3}{\theta^3} - \frac{1}{2\theta^3}\int_{\theta_0}^{\theta}\mathrm{d}\theta'\,\theta'^3\,H(\theta') \;, \tag{3.19}$$

where F_0 and G_0 are constants of integration. In case of a singular isothermal sphere with Einstein radius θ_{E}, one then has

$$\gamma = -\frac{1}{2}\frac{\theta_{\mathrm{E}}}{\theta}\,\mathrm{e}^{2\mathrm{i}\varphi} \; ; \quad \mathcal{F} = -\frac{1}{2}\frac{\theta_{\mathrm{E}}}{\theta^2}\,\mathrm{e}^{\mathrm{i}\varphi} \; ; \quad \mathcal{G} = \frac{3}{2}\frac{\theta_{\mathrm{E}}}{\theta^2}\,\mathrm{e}^{3\mathrm{i}\varphi} \;. \tag{3.20}$$

A further differentiation then shows that

$$\nabla_{\mathrm{c}}\mathcal{F} = \frac{3}{2}\frac{\theta_{\mathrm{E}}}{\theta^3}\,\mathrm{e}^{2\mathrm{i}\varphi} = \nabla_{\mathrm{c}}^{*}\mathcal{G} \;,$$

again confirming that $\mathcal{H} = 0$. The corresponding case for a B-mode shear field is obtained by multiplying all expressions in Eq.(3.20) by $\mathrm{e}^{\mathrm{i}\pi/2} = \mathrm{i}$.

To obtain a similar example for the case that flexion is not derived from a shear field, a particular mode for $\mathcal{H}$ is chosen here

$$\mathcal{H} = \frac{3\theta_{\mathrm{E}}}{\theta^3}\,\mathrm{e}^{2\mathrm{i}\varphi} \;.$$

By appropriately choosing the integration constants in Eq.(3.19), the flexions then become

$$\mathcal{F} = -\frac{\theta_{\mathrm{E}}}{2\theta^2}\,\mathrm{e}^{\mathrm{i}\varphi} \; ; \quad \mathcal{G} = -\frac{3\theta_{\mathrm{E}}}{2\theta^2}\,\mathrm{e}^{3\mathrm{i}\varphi} \;. \tag{3.21}$$

Thus, the flexion fields are very similar to those given in Eq.(3.20), except that the relative signs are different. A graphical illustration of the different cases is provided in Fig. 3.2.

3.3 Critical curves and caustics

As mentioned before, the lensing equation Eq.(3.2) can give rise to multiple images. If the flexion is sufficiently small, all but one of these multiple images will be located at a large distance from the origin, and the central image of an extended source will be isolated. In this case, this central, or primary image (from which we measure shear or flexion) is not crossed by a critical curve, and thus the source is not crossed by a caustic. The multiple images at large distances from the origin then result from the low-order Taylor expansion of the lens equation, which most likely breaks down at these image positions. These additional images are thus of no relevance. However, if the flexion becomes sufficiently large, or the source is large enough, the multiple images of an extended source will merge. If that happens, the shear or flexion cannot be measured, the whole concept will break down. This can be easily seen by considering the caustic curve cutting the source. Different parts of the source will be mapped onto a different number of image points in the lens plane, and the caustic curve introduces a direction into the situation. Hence, the assumption of an isotropic orientation of

sources can no longer be employed. Mathematically, the Jacobian determinant $\det\mathcal{A}$ of the lens equation will change sign over the image. In short, the condition that the central image is isolated (so that locally no multiple images occur) can be expressed solely by the products $G_n \Theta_s$, where Θ_s stands for the size of the source. These products approximately measure the fractional change of the reduced shear across the image of a source.

To find the principal limits of the applicability of the flexion formalism, the full Jacobian needs to be derived, which can easily be obtained from considering θ and θ^* as independent variables, and then use $\partial/\partial\theta_1 = \partial/\partial\theta + \partial/\partial\theta^*$, $\partial/\partial\theta_2 = \mathrm{i}\,(\partial/\partial\theta - \partial/\partial\theta^*)$, which can be inverted to yield $\partial/\partial\theta = \nabla_c^*/2$, $\partial/\partial\theta^* = \nabla_c/2$. With these relations, one finds that $\det\mathcal{A} = (\partial\beta/\partial\theta)(\partial\beta^*/\partial\theta^*) - (\partial\beta/\partial\theta^*)(\partial\beta^*/\partial\theta) = (\nabla_c^*\beta\,\nabla_c\beta^* - \nabla_c\beta\,\nabla_c^*\beta^*)/4$. Carrying out these derivatives, the Jacobian becomes

$$\det\mathcal{A} = 1 - gg^* - \eta^*\theta - \eta\theta^* + A^*\theta^2 + B\theta\theta^* + A(\theta^*)^2 \,, \qquad (3.22)$$

with

$$A = 4\left(\Psi_1^2 - \Psi_1^*\Psi_3\right) \,; \quad B = 4\left(\Psi_1\Psi_1^* - \Psi_3\Psi_3^*\right) \,; \quad \eta = 4\Psi_1 + 2g\Psi_1^* + 2g^*\Psi_3 \,. \qquad (3.23)$$

Note that A is a spin-2 quantity, whereas B is a real scalar, i.e., has spin-0. In the generic case, the critical curves ($\det\mathcal{A} = 0$) are conical sections, which may be degenerate, though. We will now perform a complete classification of cases that can occur, as well as to derive the critical curve(s) in parametric form; the caustics are then obtained by inserting the parametric form of the critical curves into the second-order lens equation. As we shall see, the type of conical section is determined, amongst other parameters, by the discriminant

$$\Delta = B^2 - 4AA^* \,. \qquad (3.24)$$

3.3.1 Zero discriminant

We start with the case that $\Delta = 0$, which implies $B^2 = 4AA^*$, or $B = \pm 2|A|$. The case $A = 0 = B$ either implies that $\Psi_1 = 0 = \Psi_3$, in which case also $\eta = 0$ so that no critical curves occur, or that $\Psi_3 = \Psi_1^2/\Psi_1^*$, for which $\eta \neq 0$ in general. In this case, the critical curve is a straight line, satisfying $\eta^*\theta + \eta\theta^* = 1 - gg^*$. As can be seen by inspection, it reads

$$\theta = \frac{1 - gg^*}{2\eta^*} + \mathrm{i}\lambda\eta \,, \quad -\infty < \lambda < \infty \,. \qquad (3.25)$$

If $A \neq 0$, the phase of A is defined. Since it is a spin-2 quantity, we write $A = |A|\,\mathrm{e}^{2\mathrm{i}\varphi_A}$. Furthermore, we introduce the rotation $\theta = x\,\mathrm{e}^{\mathrm{i}\varphi_A}$. Then the equation $\det\mathcal{A} = 0$ for the critical curve reads, after dividing Eq.(3.22) by $|A|$,

$$(x \pm x^*)^2 = \nu^*x + \nu x^* + \frac{gg^* - 1}{|A|} \,, \quad \text{with } \nu = \frac{\eta\,\mathrm{e}^{-\mathrm{i}\varphi_A}}{|A|} \,, \qquad (3.26)$$

and the sign on the left-hand side of the equation depends on the sign of B, where we used $B = \pm 2|A|$. The parametric form of the critical curve, which takes the form of a parabola, can then be written as

$$\theta = \frac{2\,\mathrm{e}^{\mathrm{i}\varphi_A}}{(\nu^* - \nu)}\left(2\lambda^2 - \lambda\nu + \frac{1 - gg^*}{2|A|}\right) \,; \quad \theta = \frac{2\,\mathrm{e}^{\mathrm{i}\varphi_A}}{(\nu^* + \nu)}\left(\frac{1 - gg^*}{2|A|} - \mathrm{i}\lambda\nu - 2\lambda^2\right) \,, \qquad (3.27)$$

where the first (second) equation applies for $B > 0$ ($B < 0$). Note that the parabola degenerates into a straight line if ν is real (for $B > 0$) or purely imaginary (for $B < 0$).
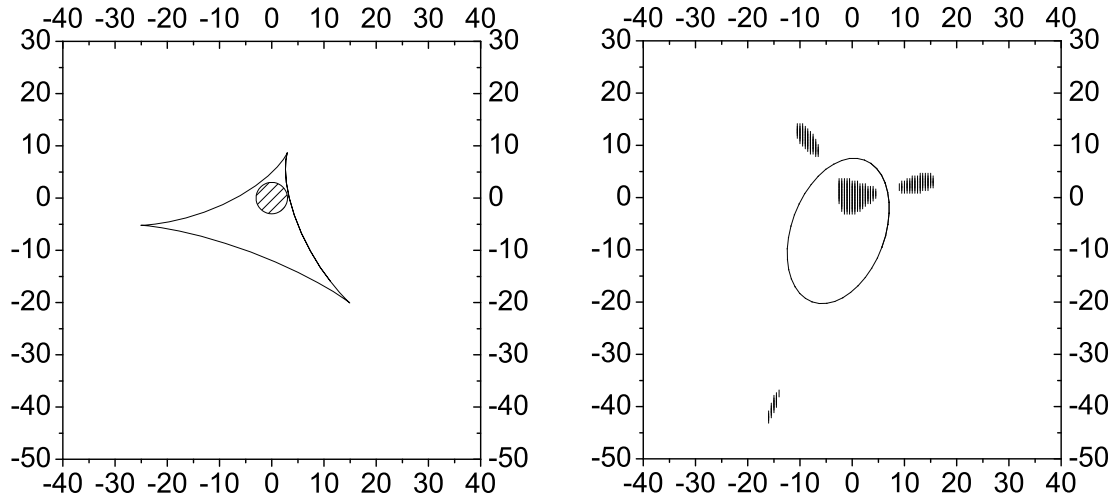
Figure 3.3: The critical curves (right-hand panel) and caustics (left-hand panel) of the lens equation (Eq.3.6) for the cases of elliptical critical curves. The parameters chosen here are $g = 0.05$, $G_1 = 0.015 + 0.035i$, $G_3 = 0.19 + 0.105i$. A circular source is mapped onto four images, as indicated. If the source size were increased, it would hit the caustic, three images would merge, and the flexion concept would break down. The unit of the reduced flexion is the inverse of the unit in which coordinates are measured
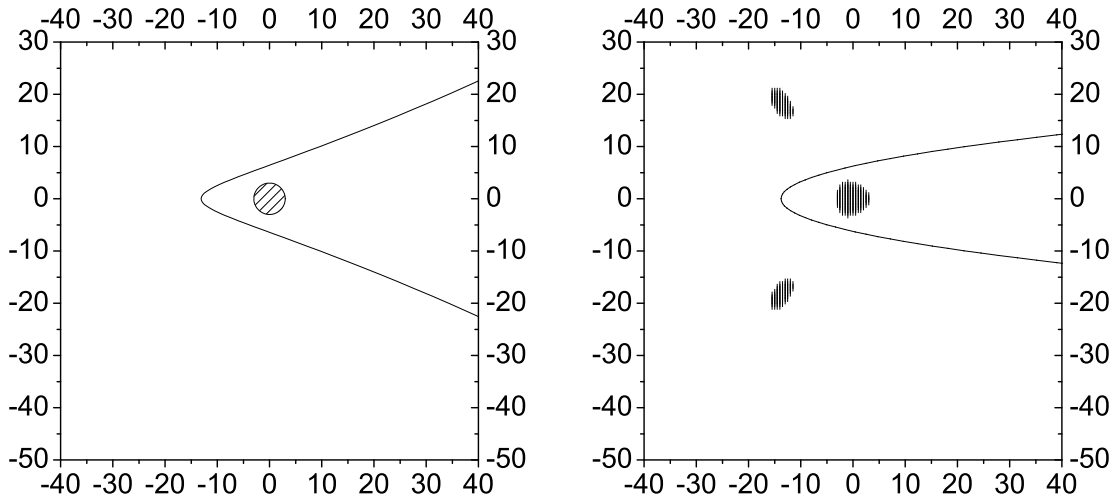

Figure 3.4: Same as Fig. 3.3, but for the parabolic case, with parameters $g = 0.05$, $G_1 = -0.04$, $G_3 = 0.112$
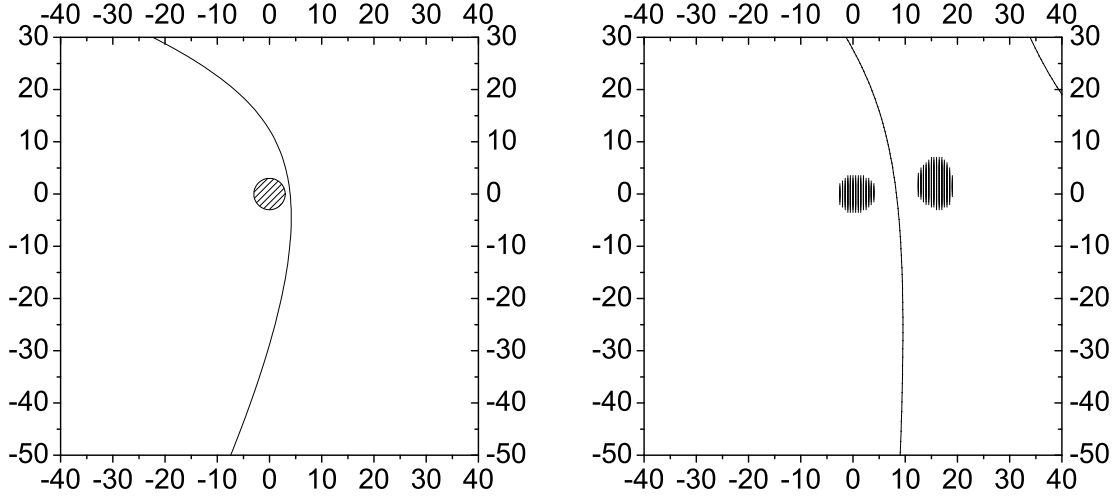
Figure 3.5: Same as Fig. 3.3, but for the hyperbolic case, with parameters $g = 0.05$, $G_1 = 0.07 + 0.015\mathrm{i}$, $G_3 = 0.03 + 0.005\mathrm{i}$

3.3.2   Non-zero discriminant

If $\Delta \neq 0$, we can perform a translation to eliminate the linear term in $\det\mathcal{A}$. Hence we define $\theta = \theta_0 + \vartheta$ and choose θ_0 such that terms linear in ϑ vanish. We then obtain for θ_0 and for the critical curve condition

$$\theta_0 = \frac{B\eta - 2A\eta^*}{\Delta} \; ; \quad A^*\vartheta^2 + B\vartheta\vartheta^* + A(\vartheta^*)^2 = C \; , \tag{3.28}$$

with

$$C = \frac{B\eta\eta^* - A(\eta^*)^2 - A^*\eta^2}{\Delta} + gg^* - 1 = -\frac{1}{\Delta}\left(gA^* + g^*A + B\right)^2 =: -\frac{1}{\Delta}V^2 \; , \tag{3.29}$$

where the second step was obtained by inserting the expression for η in terms of the Ψ's, and in the final one we defined V as the expression in the parenthesis.

As the first case, we consider $A = 0$ and $B \neq 0$ (the case $A = 0 = B$ was treated above), which implies that $\Psi_1 = 0$ and $B = -4\Psi_3\Psi_3^* < 0$. The equation for the critical curve then reduces to $B|\vartheta|^2 = C$. Furthermore, $\Delta = B^2$, and $C = -1$. Thus, the critical curve is a circle of radius $1/(2|\Psi_3|)$ and center θ_0, or $\theta = \theta_0 + \mathrm{e}^{\mathrm{i}\lambda}/(2|\Psi_3|)$, $0 \leq \lambda < 2\pi$.

We now consider the case $A \neq 0$; then the phase φ_A of A is defined, as used before. Introducing a rotation by defining $\vartheta = x\,\mathrm{e}^{\mathrm{i}\varphi_A}$, the equation for the critical curve becomes

$$|A|\left[x^2 + (x^*)^2\right] + Bxx^* = (B + 2|A|)\,x_1^2 + (B - 2|A|)\,x_2^2 = C \; . \tag{3.30}$$

The presence and topology of critical curves now depends on the signs of Δ and C. We first consider the case $C = 0$; then, if $\Delta > 0$, no critical curves occur, except for the isolated point $x = 0$. If $\Delta < 0$, the critical curves are two straight lines, as can be obtained from Eq.(3.28): inserting the ansatz $\vartheta = \lambda\,\mathrm{e}^{\mathrm{i}\zeta}$, one obtains $\mathrm{e}^{2\mathrm{i}(\zeta - \varphi_A)} = (-B \pm \mathrm{i}\sqrt{-\Delta})/(2|A|)$. Thus,

the critical curves are parameterized as

$$\theta = \theta_0 + \lambda \, e^{i\varphi_A} \sqrt{\frac{-B \pm i\sqrt{-\Delta}}{2|A|}} \; ; \quad -\infty < \lambda < \infty \; . \tag{3.31}$$

For the case of $C \neq 0$, the consideration of Eq.(3.30) yields the result that for $\Delta < 0$, the critical curves consist of two hyperbolae. From Eq.(3.29) we see that negative Δ implies $C > 0$. Also note that $\Delta < 0$ implies that $2|A| - B > 0$, $2|A| + B > 0$. The critical curves then read

$$\theta = \theta_0 + \frac{e^{i\varphi_A} V}{\sqrt{-\Delta}} \left(\pm \frac{\cosh \lambda}{\sqrt{2|A| + B}} + i \frac{\sinh \lambda}{\sqrt{2|A| - B}} \right) \; ; \quad -\infty < \lambda < \infty \; . \tag{3.32}$$

For the other case, $\Delta > 0$, we find from Eq.(3.29) that $C < 0$. If $B \pm 2|A| > 0$, we then see from Eq.(3.30) that no critical curves exist. If $B \pm 2|A| < 0$, which in particular implies $B < 0$, the critical curve is an ellipse parameterized as

$$\theta = \theta_0 + \frac{e^{i\varphi_A} V}{\sqrt{\Delta}} \left(\frac{\cos \lambda}{\sqrt{-2|A| - B}} + i \frac{\sin \lambda}{\sqrt{2|A| - B}} \right) \; ; \quad 0 \leq \lambda < 2\pi \; . \tag{3.33}$$

This concludes the classification of critical curves of the lens equation (Eq.3.6). The caustics are obtained by inserting the parameterized form of the critical curves into the lens equation. In order to see whether a critical curve cuts through the primary image of a circular source of outer isophotal radius Θ_s, one can calculate the minimum value β_{min} of $|\beta(\lambda)|$ along the caustics, or control the sign of the Jacobian determint. If $\beta_{min} > \Theta_s$, the image is not cut by a critical curve. When the source size becomes too large, some points in the image will have a negative Jacobian. For an elliptical critical curve, the maximum source size allowed is β_{min}. In the cases where two critical curves exist (e.g., two straight lines or hyperbolae), the situation is slightly more complicated. Consider, e.g., the case of two straight critical curves. Only those sections of them that are closer to the origin are relevant for this consideration, since if the primary image of the source is not cut by these closer sections of critical curves, it will still be an isolated image; the caustics coming from the outer sections of the critical curves correspond to multiply imaged source sections of secondary images. Accounting for this complication, the limit of sources size and reduced flexion have been obtained. Some examples of this are plotted in Fig. 3.6. Each panel shows the dividing line between parameter pairs $(G_1\Theta, G_3\Theta)$ for a circular source of limiting isophotal radius Θ; below the curves, no local multiple images occur, whereas for parameter pairs above the lines, the shear and flexion formalism breaks down. The different lines in each panel correspond to different values of g.

3.4 Flexion measurement

The estimation of flexion on real images is difficult even without the PSF effect or noise, since there is no simple mathematical form to quantify what a 'flexioned' image is. Nevertheless, there are two methods which are developed from shear measurement to estimate flexion, i.e. shapelets and brightness moments.
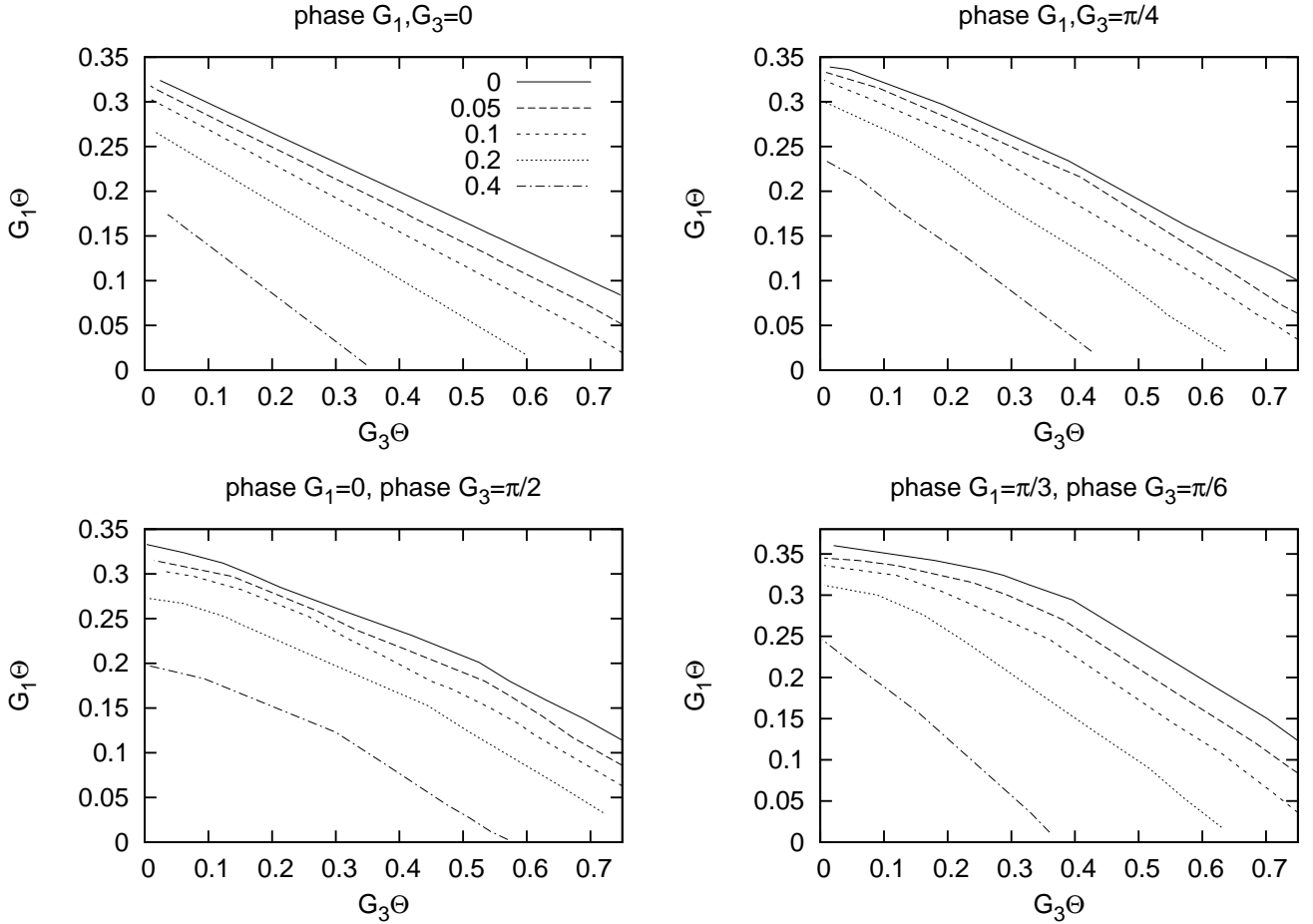
Figure 3.6: Constraints on the combination of source size and reduced flexion for the validity of the concept of flexion. Each curve shows the dividing line between a circular source of limiting isophote Θ being cut by a caustic (above the curve) or not (below the curve); in the former case, the assumptions underlying the flexion concept break down. The different curves in each panel are for different values of g, chosen as $g = 0.4, 0.2, 0.1, 0.05, 0$, as indicated by different line types. Without loss of generality, we choose g to be real and non-negative. The four panels differ in the phase of the reduced flexion, as indicated. E.g., in the upper left panel, the phases of G_1, G_3 are the same as that of g

3.4.1 Shapelets

A measurement of flexion ultimately requires very accurate knowledge of the distribution of light in the image. The shapelets decompose image into shapelet coefficients with 2D orthogonal basis functions. This technique has natural advantages. In the absence of a PSF, all shapelet coefficients will have equal noise. Moreover, the basis set is quite localized, and thus is ideal for modeling galaxies. And the generating 'step-up' and 'step-down' operators for the Hermite polynomials are simply combinations of derivatives (Refregier & Bacon 2003; Goldberg & Bacon 2005).

Here I only briefly introduce the shapelet method in polar coordinates. In Massey et al. (2007b), a polar shapelet approach was introduced to measure flexion, in which the flexion operators $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ are much simpler than corresponding expressions in Cartesian shapelet space. The observed image of galaxy $f(r, \theta)$ can be decomposed into a sum of orthogonal 2D basis functions $\chi_{n,m}(r, \theta)$ weighted by shapelet coefficients $f_{n,m}$

$$f(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{n,m} \chi_{n,m}(r, \theta). \tag{3.34}$$

The basis functions are described in Massey & Refregier (2005); Bernstein & Jarvis (2002). They are Laguerre polynomials in r multiplied by sines and cosines in θ, and a circular Gaussian of width β. (Only in this shapelet section, (r, θ) are polar coordinates, and β is the scale size.) This scale size is chosen to match the observed size of each galaxy. The shape of each galaxy can then be completely described by the array of its shapelet coefficients $f_{n,m}$. These are complex numbers, with $f_{n,-m} = f_{n,m}^*$. The indices n and m correspond to the numbers of radial and tangential oscillations respectively: n can take any non-negative integer, and m can take any integer between $-n$ and n, in steps of two. The expansion surely must be truncated to some $n_{\max}$. The galaxy image flux can be written as

$$F = \int \int f(r, \theta) \, r \mathrm{d}r \, \mathrm{d}\theta = \beta \sqrt{4\pi} \sum \sum f_{n,m}, \tag{3.35}$$

and the galaxy size

$$R^2 = \frac{1}{F} \int \int r^2 f(r, \theta) \, r \mathrm{d}r \, \mathrm{d}\theta = \frac{\beta^3 \sqrt{16\pi}}{F} \sum \sum (n+1) f_{n,m}, \tag{3.36}$$

where the summation is only over some specific conditions.

The action of the flexion operators $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ can be determined in polar shapelet space, and formally are written as

$$\hat{\mathcal{F}} : f_{n,m} \rightarrow f'_{n,m} = f_{n,m} + \mathcal{F}\beta \sum f_{n-i,m-j} + \mathcal{F}^*\beta \sum f_{n-i,m-j};$$
$$\hat{\mathcal{G}} : f_{n,m} \rightarrow f'_{n,m} = f_{n,m} + \mathcal{G}\beta \sum f_{n-i,m-j} + \mathcal{G}^*\beta \sum f_{n-i,m-j}, \tag{3.37}$$

where $f'_{n,m}$ are the coefficients for the unlensed galaxy and $f_{n,m}$ are those for the lensed galaxy. Again the summation is over the specific conditions and $f_{n-i,m-j}$ stands for all these coefficients that are needed to describe flexion. These operators are illustrated graphically in Fig. 3.7. To account for the centroid shift correction, i.e. the shift of the center of $f(r, \theta)$ and

Figure 3.7: The mixing of polar shapelet coefficients under weak lensing transformations. If a galaxy initially contains power in its $f_{6,0}$ coefficient, it will contain additional power in $f_{4,\pm2}$ and $f_{8,\pm2}$ after shear. After both types of flexion, it will contain additional power in eight shapelet coefficients, as illustrated. The directions in which power moves between adjacent coefficients are the same for a given operator wherever there are non-zero coefficients across shapelet space, although the amount of mixing varies. Wherever the pattern would seem to couple coefficients that do not exits, the amount of mixing is zero. Figure is from Massey et al. (2007b).
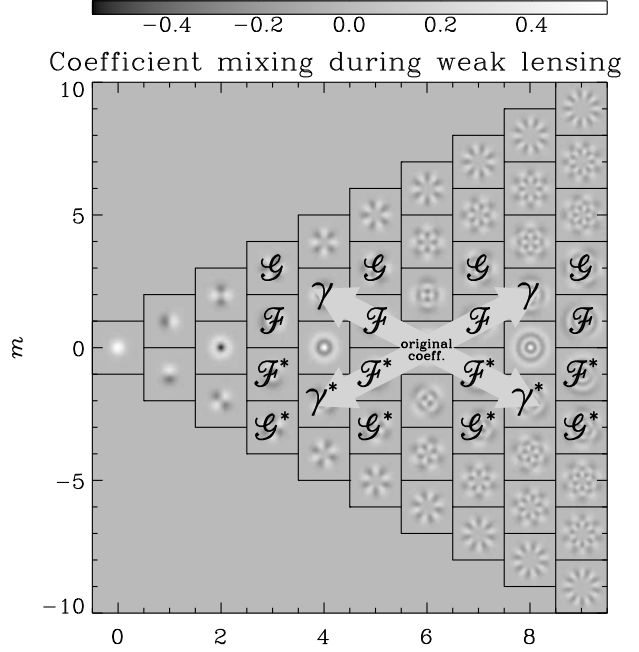
$f'(r,\theta)$, the real flexion operator used for estimation is

$$
\begin{aligned}
\hat{\mathcal{F}}_T &\equiv \hat{\mathcal{F}} - \hat{T}\left(\frac{R^2}{4\beta}(6\mathcal{F} + 5\mathcal{F}^*\epsilon)\right); \\
\hat{\mathcal{G}}_T &\equiv \hat{\mathcal{G}} - \hat{T}\left(\frac{R^2}{4\beta}\mathcal{G}\epsilon^*\right),
\end{aligned}
\tag{3.38}
$$

where $\hat{T}$ is the centroid shift operator. Finally the flexion estimator is given by

$$
\begin{aligned}
\mathcal{F}_T &= \frac{4\beta}{3}\frac{f_{1,1}}{(\beta^2 - R^2)f_{0,0} + R^2 f_{2,0} - \beta^2 f_{4,0}}; \\
\mathcal{G}_T &= \frac{4\sqrt{6}}{3\beta}\frac{f_{3,3}}{f_{0,0} + f2,0 - f_{4,0} - f_{6,0}}.
\end{aligned}
\tag{3.39}
$$

For the detail about polar shapelet for flexion measurement, the reader is referred to Massey et al. (2007b).

The shapelets analysis can produce a good decomposition, and is convenient for the flexion measurement. However, to find the best fit scaling parameter β requires large amount of computational time. This may be a significant limitation of shapelets for large lensing field application.

3.4.2 Brightness moments/ HOLICs

Apart from shapelets, the brightness moment approach provides a more intuitive picture of the effects of flexion. In Okura et al. (2007), an approach named HOLICs (Higher-Order Lensing Image characteristics) was developed to measure flexion, which is very similar to the brightness moments method. In Schneider & Er (2008), the mass-sheet degeneracy was taken into account, and the estimators for reduced flexion by brightness moments is obtained.

The image of a source denote the brightness distribution by $I^s(\beta)$. Since surface brightness is conserved by lensing, the brightness distribution of the image is $I(\theta) = I^s(\beta(\theta))$. Since the scaling of the source is unobservable, the following is shown in terms of the scaled source plane coordinates, and therefore drop the hat on β and $\mathcal{A}$ (Eq.3.6, 3.9).

Recall the definition of brightness moment for a function of the source coordinate $F(\beta)$ (Eq.2.25)

$$\text{Mom}[F(\beta)] = \int \mathrm{d}^2\beta \, F(\beta) \, I^s(\beta) = \int \mathrm{d}^2\theta \, \det\mathcal{A}(\theta) \, F(\beta(\theta)) \, I(\theta) \tag{3.40}$$

$$\approx \int \mathrm{d}^2\theta \, (1 - gg^* - \eta^*\theta - \eta\theta^*) \, F(\beta(\theta)) \, I(\theta) \, , \tag{3.41}$$

where here and in the following, we use the linear approximation for $\det\mathcal{A}$.

The centroid shift correction of the source image is important for the flexion measurement. In particular, the first flexion is a spin-1 quantity, and so is the centroid shift, thus the centroid shift directly affect the measurement result of the spin-1 flexion. The centroid shift is defined as

$$\bar{\beta} = \frac{1}{S_0} \int \mathrm{d}^2\beta \, \beta \, I^s(\beta) \tag{3.42}$$

$$= \frac{1}{S(1 - gg^*)} \int \mathrm{d}^2\theta \, (1 - gg^* - \eta^*\theta - \eta\theta^*) \left[\theta - g\theta^* - \Psi_1^*\,\theta^2 - 2\Psi_1\,\theta\theta^* - \Psi_3\,(\theta^*)^2 \right] I(\theta) \, ,$$

where S/S_0 are the flux of the lensed/unlensed image. Expanding the integrand, we note that terms linear in θ vanish, due to Eq.(2.26). The second-order brightness moments of the image in the form is defined in Chapter 2

$$Q_2 \equiv \frac{1}{S} \int \mathrm{d}^2\theta \, \theta^2 \, I(\theta) \, ; \quad Q_0 \equiv \frac{1}{S} \int \mathrm{d}^2\theta \, \theta\theta^* \, I(\theta) \, . \tag{3.43}$$

The source centroid shift can be written in terms of second-order brightness moments

$$\bar{\beta} = \frac{3G_1 g^* - 5G_1^* - 2gG_3^*}{4(1 - gg^*)} Q_2 + \frac{4gG_1^* + g^2 G_3^* - G_3 g^* - G_1(3 + gg^*)}{2(1 - gg^*)} Q_0$$

$$+ \frac{5gG_1 - 3g^2 G_1^* - (1 - 3gg^*)G_3}{4(1 - gg^*)} Q_2^* \, . \tag{3.44}$$

We now write these equations in a more compact form; for this, we define the matrix $\boldsymbol{G}$ by $\boldsymbol{G}^{\mathrm{T}} = (G_3^*, G_1^*, G_1, G_3)$, where the 'T' denotes the transpose of the matrix. Then,

$$\bar{\beta} = \boldsymbol{B}\boldsymbol{G} \, , \tag{3.45}$$

where the coefficients of $\boldsymbol{B} = (b_1, b_2, b_3, b_4)$ are given by

$$b_1 = \frac{g^2 Q_0 - gQ_2}{2(1 - gg^*)} \, ; \quad b_2 = \frac{8gQ_0 - 5Q_2 - 3g^2 Q_2^*}{4(1 - gg^*)} \, ;$$

$$b_3 = \frac{3g^* Q_2 - 2(3 + gg^*)Q_0 + 5gQ_2^*}{4(1 - gg^*)} \, ; \quad b_4 = \frac{(3gg^* - 1)Q_2^* - 2g^* Q_0}{4(1 - gg^*)} \, . \tag{3.46}$$

The centroid shift in the source plane is thus given by the product of the derivatives of the reduced shear (expressed by G_1 and G_3) and the area of the image, which is proportional to

Q_0 and Q_2. Of course, since the reduced shear and its derivatives are not directly observable, the centroid shift is unobservable as well. To get an order-of-magnitude estimate of $\bar{\beta}$, we assume that the source has a linear angular size Θ_s, consider the reduced shear to be of order unity, and let Θ_c be the angular scale on which the reduced shear varies. Then,

$$G_n = \mathcal{O}\left(\frac{1}{\Theta_c}\right) \; ; \quad Q_n = \mathcal{O}\left(\Theta_s^2\right) \quad \Rightarrow \quad \bar{\beta} = \mathcal{O}\left(\frac{\Theta_s^2}{\Theta_c}\right) . \tag{3.47}$$

The second-order brightness moments of the source is defined as $Q_2^s = \mathrm{Mom}[(\beta-\bar{\beta})^2]/S_0 = \mathrm{Mom}[\beta^2]/S_0 - \bar{\beta}^2$ and $Q_0^s = \mathrm{Mom}[(\beta-\bar{\beta})(\beta-\bar{\beta})^*]/S_0 = \mathrm{Mom}[\beta\beta^*]/S_0 - \bar{\beta}\bar{\beta}^*$. By defining the third-order brightness moments of the image through

$$T_3 \equiv \frac{1}{S}\int \mathrm{d}^2\theta\, \theta^3\, I(\theta) \; ; \quad T_1 \equiv \frac{1}{S}\int \mathrm{d}^2\theta\, \theta^2\, \theta^*\, I(\theta) , \tag{3.48}$$

we obtain

$$
\begin{aligned}
Q_2^s =\ & Q_2 - 2gQ_0 + g^2Q_2^* + \frac{8gG_1^* - (4+3gg^*)G_1 - g^*G_3 + 2g^2G_3^*}{2(1-gg^*)}T_1 \\
& + \frac{(7+gg^*)gG_1 - 7g^2G_1^* + (3gg^*-1)G_3 - g^3G_3^*}{2(1-gg^*)}T_1^* \\
& + \frac{2g^*G_1 - 3G_1^* - gG_3^*}{2(1-gg^*)}T_3 + \frac{(1-2gg^*)gG_3 - 3g^2G_1 + 2g^3G_1^*}{2(1-gg^*)}T_3^* - \bar{\beta}^2 ,
\end{aligned}
\tag{3.49}
$$

$$
\begin{aligned}
Q_0^s =\ & -g^*Q_2 + (1+gg^*)Q_0 - gQ_2^* + \frac{6g^*G_1^* + (3gg^*-1)G_3^* - 4g^{*2}G_1}{4(1-gg^*)}T_3 \\
& + \frac{2g^{*2}G_3 + (11+3gg^*)g^*G_1 - (7+9gg^*)G_1^* - (1+3gg^*)gG_3^*}{4(1-gg^*)}T_1 \\
& + \frac{2g^2G_3^* + (11+3gg^*)gG_1^* - (1+3gg^*)g^*G_3 - (7+9gg^*)G_1}{4(1-gg^*)}T_1^* \\
& + \frac{6gG_1 - 4g^2G_1^* - (1-3gg^*)G_3}{4(1-gg^*)}T_3^* - \bar{\beta}\bar{\beta}^*
\end{aligned}
\tag{3.50}
$$

Note that Q_0^s is real. In a more compact notation, Eq.(3.49) reads

$$Q_2^s = Q_2 - 2gQ_0 + g^2Q_2^* + A\boldsymbol{G} - \bar{\beta}^2 , \tag{3.51}$$

where the matrix $\boldsymbol{A} = (a_1, a_2, a_3, a_4)$ has coefficients

$$
\begin{aligned}
a_1 &= \frac{-g^3T_1^* + 2g^2T_1 - gT_3}{2(1-gg^*)} \; ; \quad a_3 = \frac{-3g^2T_3^* + g(7+gg^*)T_1^* - (4+3gg^*)T_1 + 2g^*T_3}{2(1-gg^*)} \; ; \\
a_2 &= \frac{2g^3T_3^* - 7g^2T_1^* + 8gT_1 - 3T_3}{2(1-gg^*)} \; ; \quad a_4 = \frac{g(1-2gg^*)T_3^* - (1-3gg^*)T_1^* - g^*T_1}{2(1-gg^*)} .
\end{aligned}
\tag{3.52}
$$

In order to get estimators for flexion, we write down the third-order brightness moments of the source, separated into a spin-3 and a spin-1 component,

$$T_3^s = \frac{\mathrm{Mom}[(\beta-\bar{\beta})^3]}{S_0} = \frac{\mathrm{Mom}[\beta^3]}{S_0} - 3\bar{\beta}\,\frac{\mathrm{Mom}[\beta^2]}{S_0} + 3\bar{\beta}^2\,\frac{\mathrm{Mom}[\beta]}{S_0} - \bar{\beta}^3 = \frac{\mathrm{Mom}[\beta^3]}{S_0} - 3\bar{\beta}Q_2^s - \bar{\beta}^3 , \tag{3.53}$$

where we used that $\mathrm{Mom}[\beta^2]/S_0 = Q_2^{\mathrm{s}} + \bar{\beta}^2$ and $\mathrm{Mom}[\beta\beta^*]/S_0 = Q_0^{\mathrm{s}} + \bar{\beta}\bar{\beta}^*$. Similarly, we obtain

$$T_1^{\mathrm{s}} = \frac{\mathrm{Mom}[\left(\beta - \bar{\beta}\right)^2 \left(\beta^* - \bar{\beta}^*\right)]}{S_0} = \frac{\mathrm{Mom}[\beta^2\beta^*]}{S_0} - 2Q_0^{\mathrm{s}}\bar{\beta} - Q_2^{\mathrm{s}}\bar{\beta}^* - \bar{\beta}^2\bar{\beta}^* . \tag{3.54}$$

Defining the fourth-order brightness moments of the image by

$$F_0 = \frac{1}{S} \int \mathrm{d}^2\theta \, (\theta\theta^*)^2 I(\theta) \; ; \quad F_2 = \frac{1}{S} \int \mathrm{d}^2\theta \, \theta^3\theta^* \, I(\theta) \; ; \quad F_4 = \frac{1}{S} \int \mathrm{d}^2\theta \, \theta^4 \, I(\theta) \; , \tag{3.55}$$

where F_n is a spin-n quantity, we obtain for the third-order moments of the source:

$$\mathcal{T}^{\mathrm{s}} = \tau + \boldsymbol{C}\,\boldsymbol{G} + \mathcal{O}(\bar{\beta}^3) \; , \tag{3.56}$$

where the matrix $\mathcal{T}^{\mathrm{s}}$ is defined by its transpose $\mathcal{T}^{\mathrm{s,T}} = (T_3^{s*}, T_1^{s*}, T_1^s, T_3^s)$. The elements of τ are

$$\begin{aligned}
\tau_1 &= T_3^* - 3g^*T_1^* + 3g^{*2}T_1 - g^{*3}T_3 \; ; \quad \tau_2 = -gT_3^* + (1 + 2gg^*)T_1^* - g^*(2 + gg^*)T_1 + g^{*2}T_3 \; ; \\
\tau_3 &= \tau_2^* \; ; \quad \tau_4 = \tau_1^* \; ,
\end{aligned} \tag{3.57}$$

where the last two relations are obvious. The 4×4 matrix $\boldsymbol{C}$ is given explicitly in Appendix B; each of its elements consists of a sum of terms proportional to fourth-order brightness moments, F_n, and terms proportional to squares of second-order brightness moments. Okura et al. (2007a) and Goldberg & Leonard (2007) have derived expressions similar to Eq.(3.56), though using a number of simplifying assumptions (such as $|g| \ll 1$) and not considering the reduced flexion.

There are a large number of terms in Eq.(3.56), and there is no need to take all of them into account for the flexion estimator. Therefore the order-of-magnitudes of the various terms appearing in Eq.(3.51) and Eq.(3.56) need to be considered. Assuming that the third-order moments of the sources are small, then the third-order moments of the image are given by the product of $\boldsymbol{C}$ and $\boldsymbol{G}$. With $\boldsymbol{G} = \mathcal{O}\left(1/\Theta_{\mathrm{c}}\right)$ and $\boldsymbol{C} = \mathcal{O}\left(\Theta_{\mathrm{s}}^4\right)$, we find that $T = \mathcal{O}\left(\Theta_{\mathrm{s}}^4/\Theta_{\mathrm{c}}\right) = \mathcal{O}\left(\Theta_{\mathrm{s}}^3\right)\left(\Theta_{\mathrm{s}}/\Theta_{\mathrm{c}}\right)$. To get an estimate of the size of the various terms in Eq.(3.51), we note that the first three terms on the right-hand side (those proportional to the Q_n) are of order $\mathcal{O}\left(\Theta_{\mathrm{s}}^2\right)$, whereas $\boldsymbol{AG} = \mathcal{O}\left(\Theta_{\mathrm{s}}^4/\Theta_{\mathrm{c}}\right) \mathcal{O}\left(1/\Theta_{\mathrm{c}}\right)$ and $\bar{\beta}^2 = \mathcal{O}\left(\Theta_{\mathrm{s}}^4/\Theta_{\mathrm{c}}^2\right)$. Hence, the last two terms are of equal magnitude in general, each of them being smaller than the first three terms by a factor $\left(\Theta_{\mathrm{s}}/\Theta_{\mathrm{c}}\right)^2$. Only if the source is of the same order as the scale over which the reduced shear varies do the last two terms in Eq.(3.51) contribute. In Eq.(3.56), we have neglected the terms $\bar{\beta}^3$, since they are two powers of $\left(\Theta_{\mathrm{s}}/\Theta_{\mathrm{c}}\right)$ smaller than the terms written down.

3.4.3 Reduced shear and flexion estimates

Eq.(3.56) is a linear equation for $\boldsymbol{G}$, which can thus be solved,

$$\boldsymbol{G} = \boldsymbol{C}^{-1}\left(\mathcal{T}^{\mathrm{s}} - \tau\right) \; . \tag{3.58}$$

Inserting this into Eq.(3.51) then yields

$$Q_2^{\mathrm{s}} = Q_2 - 2gQ_0 + g^2Q_2^* + \boldsymbol{A}\,\boldsymbol{C}^{-1}\left(\mathcal{T}^{\mathrm{s}} - \tau\right) - \left(\boldsymbol{BG}\right)^2 \; . \tag{3.59}$$

We are thus left with a single complex equation for g, which contains the observable brightness moments of the image, as well as the unobservable brightness moments of the source. This equation can be used to estimate the reduced shear if we make assumptions concerning the properties of the source brightness moments. We assume that the sources are oriented randomly, which implies that all quantities with spin unequal zero have a vanishing expectation value. That is, we set $Q_2^{\mathrm{s}} = 0$, $\mathcal{T}^{\mathrm{s}} = 0$, to arrive at

$$Q_2 - 2gQ_0 + g^2 Q_2^* = \boldsymbol{A}\,\boldsymbol{C}^{-1}\tau + \left(\boldsymbol{B}\boldsymbol{C}^{-1}\tau\right)^2 =: Y(g)\,, \tag{3.60}$$

where we have indicated that the right-hand side depends on the reduced shear (in fact it does so in a very complex manner). However, since we have argued above that the terms on the left-hand are much larger than those on the right-hand side, an iterative solution of this equation is suggested. Assume the right-hand side is given, then we get the solutions

$$g = \frac{\chi}{|\chi|^2}\left(1 \pm \sqrt{1 - |\chi|^2 + \frac{Y\chi^*}{Q_0}}\right)\,, \quad \text{where} \ \ \chi = \frac{Q_2}{Q_0} \tag{3.61}$$

is the complex ellipticity of the image. Obviously, there are two solutions g for a given value of Y. This situation is similar to that of 'ordinary' weak lensing, where this ambiguity also occurs: as shown by Schneider and Seitz (1995), from shape measurements of background galaxies, one cannot distinguish locally between an estimate g and $1/g^* = g/|g|^2$. The same occurs here; we therefore assume that we pick one of the two solutions, say the one corresponding to the '$-$' sign; this then yields for small shear $g \approx \chi/2$. It should be stressed that flexion impacts the determination of shear from the second-order brightness moments, due to its impact on higher-order brightness moments; hence, in general the determination of shear and flexion are coupled.

We start the iteration by setting $Y_0 = 0$. This yields a first-order solution for the estimate of g,

$$g_0 = \frac{\chi}{|\chi|^2}\left(1 - \sqrt{1 - |\chi|^2}\right)\,. \tag{3.62}$$

We then use the iteration equations

$$Y_n = Y(g_{n-1})\,; \ \ g_n = \frac{\chi}{|\chi|^2}\left(1 - \sqrt{1 - |\chi|^2 + \frac{Y_n\chi^*}{Q_0}}\right)\,. \tag{3.63}$$

This procedure converges quickly to one of the two solutions (g, G_1, G_3); the other solution is obtained by taking the '$+$' sign in the above equations.

Of course, our approach of setting $Q_2^{\mathrm{s}} = 0$ yields a biased estimator for g; this is true even in the absence of flexion (e.g., Schneider & Seitz 1995). The reason is that, although the expectation value of Q_2^{s} vanishes, the resulting estimator for g is a non-linear function of $\chi^{\mathrm{s}} = Q_2^{\mathrm{s}}/Q_0^{\mathrm{s}}$ and thus biased. The bias depends on the ellipticity distribution of the sources. It should be stressed, however, that a modified definition of image ellipticity exist such that its expectation value is an unbiased estimate of the reduced shear (Seitz & Schneider 1997).

The flexion estimator is given by Eq.(3.58). Since the matrix $\boldsymbol{C}$ contains many terms, this is a fairly complicated equation in general. A simpler estimate is obtained if we assume that the reduced shear is small, $|g| \ll 1$, in which case the matrix $\boldsymbol{C}$ simplifies considerably

– see Appendix B. Furthermore, if we assume that the brightness moments of spin $\neq 0$ are much smaller than the corresponding ones with spin 0, then we find the simple relations

$$T_1^{\mathrm{s}} \approx T_1 - \frac{9F_0 - 12Q_0^2}{4}G_1 \; ; \quad T_3^{\mathrm{s}} \approx T_3 - \frac{3F_0}{4}G_3 \; . \tag{3.64}$$

If we then set the $T_n^{\mathrm{s}} = 0$, as would be true for the expectation value, then we obtain as estimates for the reduced flexion

$$G_1 \approx \frac{4}{9F_0 - 12Q_0^2}T_1 \; ; \quad G_3 \approx \frac{4}{3F_0}T_3. \tag{3.65}$$

Thus, the flexion is then given by the third-order brightness moments of the image, divided by a quantity that just depends on the size of the image. Similar relations to Eq.(3.65) have been given in Goldberg & Leonard (2007) and Okura et al. (2007a).

3.4.4 Accuracy of brightness moment estimators

Some simulations were performed in order to test the behavior of the estimators by brightness moments (Eq.3.65). We model the sources as elliptical Gaussians, truncated at three times the scale 'radius' Θ_{s} chosen such that the area of a source is independent of its ellipticity. The ellipticity of the sources follows a Gaussian distribution, with a dispersion of χ^{s} of $R = 0.4$ (i.e., we use the same ellipticity distribution as in Schneider & Seitz 1995). However, the intrinsic ellipticity distribution is truncated at $|\chi^{\mathrm{s}}| \leq 0.9$, since in the limit of $|\chi^{\mathrm{s}}| \to 1$, there will be orientation angles for which the source will hit a caustic. For each source, we map a grid of pixels from the lens plane to the source plane using the lens equation to obtain the brightness distribution in the lens plane. From this distribution, the brightness moments of the image are measured. A shift in the lens plane coordinates is applied as to satisfy Eq.(2.26). We then apply the shear and flexion estimators described above to the resulting brightness moments Q_n, T_n and F_n. The shear and flexion estimates are then averaged over the Gaussian ellipticity distribution of the sources, in particular over their random orientation.

It should be noted that flexion is a dimensional quantity $\propto \Theta_{\mathrm{c}}^{-1}$. As can be checked explicitly from its definition, the way flexion appears in the equations is always with one order higher in the source (or image) size than the other terms in the equations. As an example, we consider Eq.(3.56); the left-hand side and the first term on the right-hand side are $\propto \Theta_{\mathrm{s}}^3$, whereas the coefficients of the matrix $C \propto \Theta_{\mathrm{s}}^4$. This then implies that the accuracy of the flexion estimates does not depend on the magnitude of the flexion and the source size individually, but only on the product $G_n \Theta_{\mathrm{s}}$. Therefore, the following results are quoted always in terms of this product.

We now present some results of our numerical simulation regarding the accuracy with which the reduced shear and flexion can be obtained with our moment approach. For given input values of g, G_1 and G_3, we either measure the brightness moments for a single circular source, or average the results over an ellipticity distribution, as described above. It should be noted that we have to deal with a 5-dimensional parameter space, namely the 3 complex parameters g, G_1 and G_3, minus one overall phase that can be chosen, e.g., to make g real and positive. Thus, instead of sampling the parameter space comprehensively, we only give a few selected results.

We start by considering a circular source, and determine the effect of flexion on the determination of the reduced shear. The left-hand panel of Fig. 3.9 shows contours of constant

fractional deviation $\Delta g/g$, in the flexion parameter plane. Here it is assumed that the phase of both flexion components is the same as that of g (as would be the case in an axially-symmetric lens potential). Errors of order 5% occur already for $\sqrt{|G_1^2 + G_3^2|}\Theta_s \sim 0.03$, and the fractional error increases approximately linearly with the strength of flexion (or with the source size), although it does not scale equally with both flexion components. The reason for this effect has been mentioned before – flexion affects the transformation between source and image quadruple moments, as can be seen in Eq.(3.49).

In Fig. 3.8, we show the expectation value of the reduced flexion components, as a function of the input flexion. The expectation value has been determined by averaging over an isotropic ensemble of elliptical sources, as described before. The left and right panel show the behavior of the expectation value of G_1 and G_3, respectively, where the other flexion component was set to zero. The dashed curve shows the identity, the plus symbols were obtained by using the approximate estimator Eq.(3.65), whereas the crosses show the expectation values as obtained by employing the full expression Eq.(3.58), where the corresponding value of g was obtained by the iterative process described in last section. It is reassuring that the expectation value closely traces the input value, i.e., that the estimates have a fairly small bias. Furthermore, we see that the approximate estimator Eq.(3.65) performs remarkably well. It is seen that the estimates for G_3 behave better than those for G_1. This can also be seen from the right-hand panel of Fig. 3.9, where we plot contours of constant fractional error

$$\Delta G := \sqrt{\left|\frac{\Delta G_1}{G_1}\right|^2 + \left|\frac{\Delta G_3}{G_3}\right|^2}\,, \tag{3.66}$$

where ΔG_n is the deviation of the estimate of G_n from its input value. For simplicity, we have assumed a circular source. We see that the accuracy decreases much faster with increasing G_1 than with increasing G_3.

In addition, the moment approach for flexion as presented in this chapter must be modified in several ways to be applicable to real data. First, brightness moments must be weighted in order to not be dominated by the very noisy outer regions of the image. As is known from shear measurements, such a weighting affects the relation between source and image brightness moments. Secondly, one needs to account for the effects of a point spread function. I will present a preliminary result on that in Appendix A.

A word on the two approaches of flexion measurement. In Goldberg & Leonard (2007), the Cartesian shapelets and the HOLICs are compared. And they give similar measurement errors at fixed sky noise. However, there is strong correlation of the errors for spin-1 flexion by the two approaches and much lower correlation for spin-3 flexion. The two techniques have different response to the signal, and thus noise. In other words, the two techniques may measure 'different' spin-3 flexion. One might ask, which one is better? This is a difficult question (at least you cannot trust the people who is in shapelets or brightness moment neither). It may turn out that the shapelet is more convenient with measurement and PSF corrections, the moment approach provides a more intuitive picture of the effects of flexion. Besides, the shapelets approach is very expensive in computing time. The shapelet decomposition runs several orders of magnitude slower than HOLICs. For large lensing survey, this may be a limitation of shapelets.
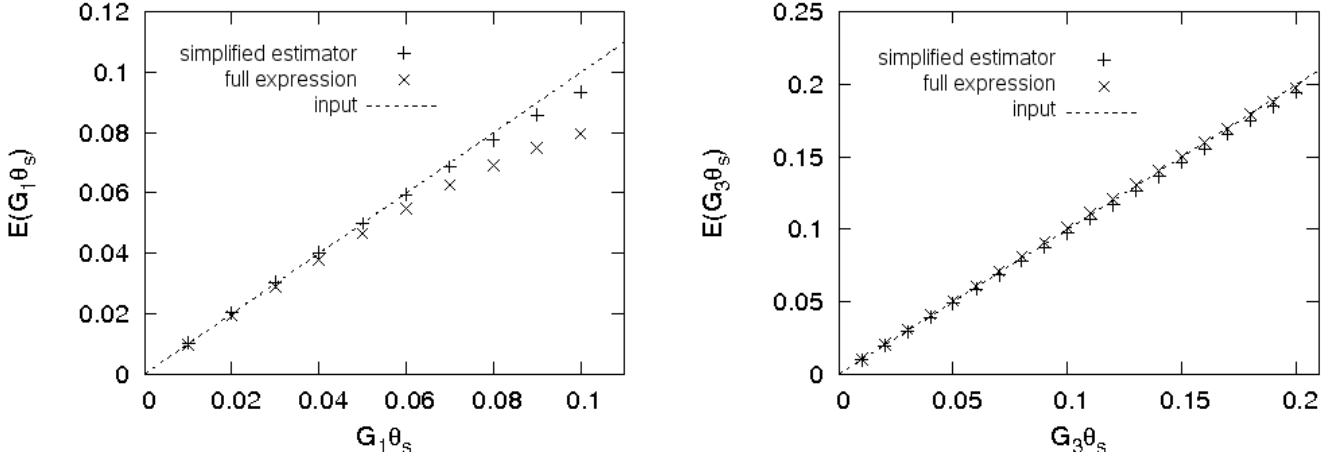
Figure 3.8: Comparison of the reduced flexion estimators Eq.(3.65) with the full expression Eq.(3.58) and the input value. The horizontal and vertical axis show $G_i\Theta_s, i = 1, 3$. For both panels, we take $g = 0.05$, and $G_3 = 0$ ($G_1 = 0$) for the left (right) panel. The line indicates the input value, the plus symbols show the simplified reduced flexion estimate Eq.(3.65), and the crosses result from the full expression of reduced flexion Eq.(3.58). As can be seen from the left-hand panel, the full estimator for the reduced flexion yields a more biased result that the approximate expression Eq.(3.65); we have not found a reasonable explanation for this behavior
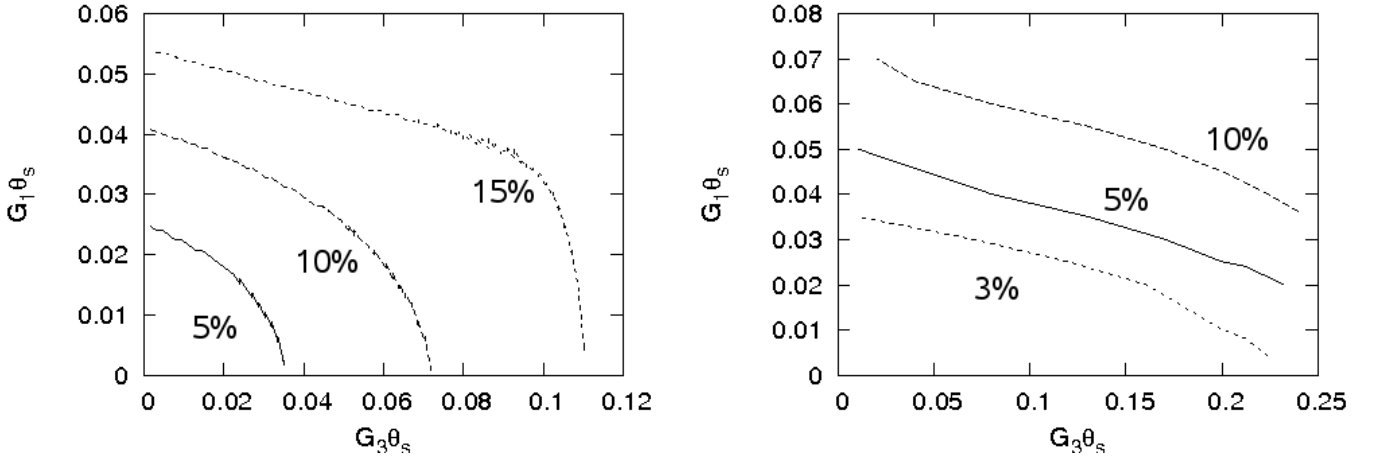

Figure 3.9: Accuracy of the estimates for reduced shear and flexion. The left panel shows contour of constant fractional error of 5%, 10% and 15%, on the estimate of the reduced shear g, as a function of $G_i\Theta_s$, where we chose $g = 0.05$ as input value, and assumed the phases of G_1, G_3 to be the same as that of g. The estimate was obtained by solving the iteration equations given in Sect. 3.4.3. The right panel shows the fractional error levels at 3, 5, and 10% for the reduced flexion, as quantified by Eq.3.66, where the estimate was obtained again with the iterative procedure. In both cases, we assumed circular sources

Chapter 4

Mass reconstruction of galaxy clusters

Galaxy clusters have been the focus of a very intense ongoing research in the past decades, since they are the most massive bound structures in the universe. Moreover, a cluster contains hundreds or thousands of galaxies, which do not cover a large fraction of cluster mass. The galaxy clusters thus are believed to be a probe of the matter content of the universe: the ratio of their luminous mass to their total mass resembles the overall baryonic fraction in the universe. And this can be measured only if reliable masses of clusters can be obtained. The gravitational lensing effects offer an excellent tool to study the matter in the universe, since it is independent of the nature (luminous or dark) or state of the matter. We therefore can use gravitational lensing to constrain the projected mass distribution of galaxy clusters. In this chapter, I discuss this problem on how flexion can contribute to the cluster mass reconstruction.

4.1   Mass reconstruction by flexion

4.1.1   Kaiser-Squires inversion

The 2D projected matter density can be obtained from both flexions $\mathcal{F}$ and $\mathcal{G}$ (BGTR06) directly, following the strategy of Kaiser & Squires (1993). The Fourier transform of the relation $\mathcal{F}_i = \partial_i \kappa$ is

$$\hat{\mathcal{F}}_1 = -\mathrm{i} l_1 \hat{\kappa}(\boldsymbol{l}); \quad \hat{\mathcal{F}}_2 = -\mathrm{i} l_2 \hat{\kappa}(\boldsymbol{l}). \tag{4.1}$$

We can invert both of these terms to obtain an estimate for κ, and perform the similar calculation for $\mathcal{G}$ and κ. An explicit representation for the inversion equations is obtained in Fourier space

$$
\begin{aligned}
\hat{\kappa}_{\mathcal{F}}(\boldsymbol{l}) &= -\mathrm{i}\frac{l_1 \hat{\mathcal{F}}_1 + l_2 \hat{\mathcal{F}}_2}{l_1^2 + l_2^2}; \\
\hat{\kappa}_{\mathcal{G}}(\boldsymbol{l}) &= -\mathrm{i}\frac{(l_1^3 - 3 l_1 l_2^2)\hat{\mathcal{G}}_1 + (3 l_1^2 l_2 - l_2^3)\hat{\mathcal{G}}_2}{(l_1^2 + l_2^2)^2},
\end{aligned} \tag{4.2}
$$

for $\boldsymbol{l} \neq 0$. This provide us with the mass-mapping equations. We measure the flexion field, take the Fourier transform, calculate $\hat{\kappa}$ according to these equations and then take the inverse

Fourier transform to find κ. These mapping relations can be expressed in the complex notation (BGTR06)

$$
\begin{aligned}
(\kappa^{\mathrm{E}} + \mathrm{i}\kappa^{\mathrm{B}})_{\mathcal{F}} &= \nabla^{-2}\nabla_{\mathrm{c}}^{*}\mathcal{F}; \\
(\kappa^{\mathrm{E}} + \mathrm{i}\kappa^{\mathrm{B}})_{\mathcal{G}} &= \nabla^{-4}\nabla_{\mathrm{c}}^{*}\nabla_{\mathrm{c}}^{*}\nabla_{\mathrm{c}}^{*}\mathcal{G},
\end{aligned}
\tag{4.3}
$$

where the imagine part again gives us the B-mode component (Eq.2.49), which can be used as a test of systematics, and ∇^{-2} is the 2D inverse Laplacian, the κ^{B} field is included as the complex part of the field.

Now the mass maps can be obtained with independent noise for γ, $\mathcal{F}$ and $\mathcal{G}$. Further we can combine these mass reconstructions $\hat{\kappa}_a$ ($a = \gamma, \mathcal{F}, \mathcal{G}$) linearly in Fourier space with minimum noise variance weighting in order to obtain a best mass map (Okura et al. 2007)

$$
\hat{\kappa}(\boldsymbol{k}) = \frac{\Sigma_a \hat{W}_{\kappa a}(\boldsymbol{k})\hat{\kappa}_a(\boldsymbol{k})}{\Sigma_a \hat{W}_{\kappa a}(\boldsymbol{k})},
\tag{4.4}
$$

where $\hat{W}_{\kappa a}(\boldsymbol{k}) = 1/P_{\kappa a}^{(N)}(\boldsymbol{k})$. The noise power spectrum $P_{\kappa a}^{(N)}(\boldsymbol{k})$ of a κ map reconstructed using ath observable is defined as

$$
\begin{aligned}
P_{\kappa\gamma}^{(N)} &= \frac{P_{\gamma}^{(N)}(\boldsymbol{k})}{2} = \frac{\sigma_{\gamma}^2}{8\pi n_{\mathrm{g}}}; \\
P_{\kappa\mathcal{F}}^{(N)} &= \frac{P_{\mathcal{F}}^{(N)}(\boldsymbol{k})}{2k^2} = \frac{\sigma_{\mathcal{F}}^2}{8\pi n_{\mathrm{gf}}k^2}; \\
P_{\kappa\mathcal{G}}^{(N)} &= \frac{P_{\mathcal{G}}^{(N)}(\boldsymbol{k})}{2k^2} = \frac{\sigma_{\mathcal{G}}^2}{8\pi n_{\mathrm{gf}}k^2},
\end{aligned}
\tag{4.5}
$$

where $P_a^{(N)}(\boldsymbol{k})$ is the shot noise power of the ath observable, σ_a is the intrinsic dispersion of the ath observable and n_{g} is the surface number density of background galaxies. Here one should notice that the number density of flexion n_{rmgf} may not be the same as that of shear.

In Okura et al. (2008), the KS inversion flexion mass reconstruction has been put into use for real data. The cluster A1689 has been analyzed and the central density peaks are identified which clearly correspond to peaks in galaxy counts (Fig. 4.1).

4.1.2 Finite-field inversion

There is a natural way to use flexion for the finite-field mass reconstructions in weak lensing. Kaiser (1995) derived a relation between the gradient of $K = \ln(1 - \kappa)$ and combinations of first derivatives of g

$$
\nabla K = \frac{-1}{1 - gg^*} \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix} \begin{pmatrix} g_{1,1} + g_{2,2} \\ g_{2,1} - g_{1,2} \end{pmatrix} \equiv \boldsymbol{u}(\boldsymbol{\theta}).
\tag{4.6}
$$

Here the two components in the second bracket are the real and imagine part of spin-1 reduced flexion. The right side of this equation thus can be obtained from observation. Seitz & Schneider (2001) formulated the finite-field method from measured reduced shear in terms of a von Neumann boundary value problem, whose solution determines K up to an additive constant. Eq.(4.6) can thus be solved by line integration, but different schemes yield different
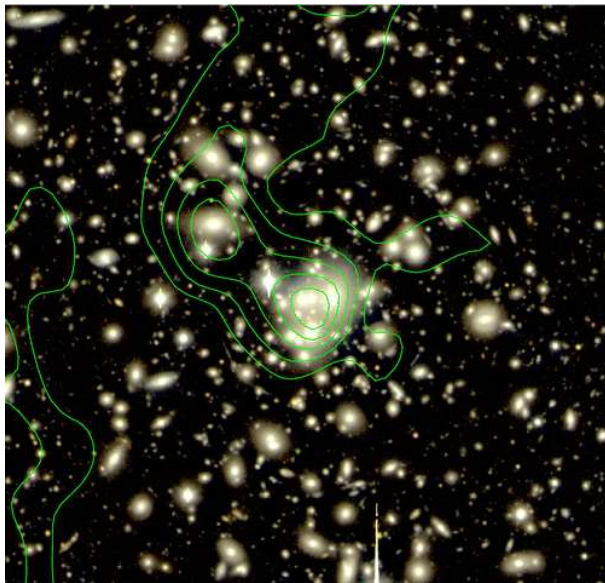
Figure 4.1: Image of cluster A1689 ($z = 0.183$) in the central $4' \times 4'$ region, composed of the Subaru/Suprime-Cam V and i' images. Contours are the lensing κ-field reconstructed from the spin-1 flexion measurements using the i'-band data with a Gaussian filter of FWHM ($= 0'.33$). The rms dispersion in the Gaussian smoothed B-mode κ map is obtained as $\sigma \approx 0.51$. The contours are spaced in units of 1σ. The peak has a value of $\kappa_E = 2.66$, and is detected at 5.2σ significance. Figure is from Okura et al. (2008).

results, since the vector field $\boldsymbol{u}$ comes from noisy observations. Moreover, $K_2 = \nabla^2 K$ is determined by the reduced shear and its derivatives, and is given by Eq.(3.17). If flexion is measured, one can replace the 'source' for $\nabla^2 K$ by a weighted sum of the differentiated reduced shear field and the combination $(K_2 + K_2^*)/2$ of the flexion field, with the weights chosen according to the estimated noise properties of both contributions.

4.2 Mass reconstruction by shear and flexion

Although flexion can be used for mass reconstruction equally as shear, one should notice that flexion is very sensitive to the small-scale variations in the gravitational potential, such as substructure. Hence, the better way of making use of flexion for mass reconstruction is combined with shear and even strong lensing information. The idea of this method is similar with strong and weak lensing united (Bradač et al. 2005b,a), the difference here is that we also include the flexion information. There is a strong flexion signal near the center of the cluster and substructures. The result shows that flexion can significantly improve the cluster mass profile to small radii and resolving of substructure.

4.2.1 The χ^2-function

The method is to describe the cluster mass distribution by the deflection potential ψ on a regular grid field, use the finite differencing method to calculate the deflection angle, the reduced shear and flexion. All these quantities on the grids considered as a model compare with data. Our aim is to seek a potential field $\psi(\boldsymbol{\theta})$ that minimizes the difference between model (g, G_1, G_3) and data (ϵ, t_1, t_3). Here t_1, t_3 stand for the estimators of flexion using brightness moments. We therefore define a χ^2-function

$$\chi^2(\psi) = \chi_s^2(\psi) + \chi_\epsilon^2(\psi) + \chi_f^2(\psi) + \eta R(\psi), \tag{4.7}$$

χ_s^2, $\chi_\epsilon^2(\psi)$ and $\chi_f^2(\psi)$ contain the information from strong lensing multiple image system, shear and flexion, respectively. ηR is a regularization term, which is a function to smooth out

the small-scale numerical fluctuations. For a multiple image system with N_m images located at $\boldsymbol{\theta}_m$, the χ_s^2 term is well defined by Bradač et al. (2005b)

$$\chi_s^2 = \sum \sum_{m=1}^{N_m} b_m^T S^{-1} b_m, \tag{4.8}$$

where the first $\sum$ is sum over all strong lens system, and $\boldsymbol{b}_m = \boldsymbol{\theta}_m - \boldsymbol{\alpha}(\boldsymbol{\theta}_m) - \boldsymbol{\beta}_s$. $\boldsymbol{\beta}_s$ is the average source position, and S is the covariance matrix $S = \mathrm{diag}(\sigma_{s1}^2, \sigma_{s2}^2)$, where σ_{s1} and σ_{s2} are the errors on image positions, projected onto the source plane. The shear term χ_ϵ^2 is

$$\chi_\epsilon^2 = \sum_{i=1}^{N_{\mathrm{g}}} \frac{|\epsilon_i - g(\boldsymbol{\theta}, z_i)|^2}{\sigma_\epsilon^2}, \tag{4.9}$$

where N_{g} is the number of background galaxies, and

$$\sigma_\epsilon^2 = \left(1 - |g|^2\right)^2 \sigma_{\epsilon^s}^2 + \sigma_{\mathrm{err}}^2, \tag{4.10}$$

with $\sigma_{\epsilon^s} \approx 0.3$ is the standard deviation of intrinsic ellipticity of galaxies (Brainerd et al. 1996), σ_{err} is the measurement error, which we take 0.1 in this chapter, and $|g|$ refers to the reconstructed value.

The flexion term is defined in a way similar to the shear

$$\chi_f^2 = \sum_{i=1}^{N_f} \left(\frac{|t_{1i} - G_1(\boldsymbol{\theta}_i, z_i)|^2}{\sigma_{t1}^2} + \frac{|t_{3i} - G_3(\boldsymbol{\theta}_i, z_i)|^2}{\sigma_{t3}^2} \right), \tag{4.11}$$

where the $G_{1,3}$ are the reduced flexions at position $\boldsymbol{\theta}_i$ and redshift z_i in our model. Note that what we can measure is not flexion but the reduced flexion. There is significant difference between the reduced flexion and its approximation $\mathcal{F}/(1-\kappa)$ or $\mathcal{G}/(1-\kappa)$ in the region where the shear is not small. On the other hand, the definition of reduced flexion (Eq.3.5) renders the χ_f^2 function (Eq.4.11) complicated and the equations become difficult to solve. We thus define

$$G_1' \equiv \frac{\mathcal{F}}{1-\kappa}; \quad G_3' \equiv \frac{\mathcal{G}}{1-\kappa}. \tag{4.12}$$

From Eq.(3.5) it is easy to obtain

$$G_1' = \frac{G_1 - gG_1^*}{1 - gg^*}; \quad G_3' = G_3 - gG_1'. \tag{4.13}$$

We thus use corrective estimators for the observed reduced flexion,

$$t_1' = \frac{t_1 - \epsilon t_1^*}{1 - \epsilon\epsilon^*}; \quad t_3' = t_3 - \epsilon t_1', \tag{4.14}$$

and replace t_1, t_3 by t_1', t_3' in (Eq.4.11). The flexion term χ_f^2 is thus redefined as

$$\chi_f^2 = \sum_{i=1}^{N_f} \left(\frac{|\frac{\mathcal{F}_i}{1-\kappa_i} - t_{1i}'|^2}{\sigma_{t1'}^2} + \frac{|\frac{\mathcal{G}_i}{1-\kappa_i} - t_{3i}'|^2}{\sigma_{t3'}^2} \right). \tag{4.15}$$

Here these $\sigma_{t1'}^2$ and $\sigma_{t3'}^2$ are different from the dispersion in flexion measurement or intrinsic flexion variance, which are both difficult to obtain from current observation. In Goldberg & Leonard (2007), the estimate of the intrinsic scatter of flexion are $\sigma_{a|\mathcal{F}|} = 0.03$ and $\sigma_{a|\mathcal{G}|} = 0.04$, where a is size of the image, thus the combination $a|\mathcal{F}|$ represents a dimensionless term. Hawken & Bridle (2009); Leonard et al. (2009) used a conservative estimate $\sigma_{\mathcal{F}} = 0.1''$. Here we use

$$\sigma_{t1'}^2 = \left|\frac{\partial G_1'}{\partial g}\right|^2 \sigma_\epsilon^2 + \sigma_{t1}^2; \quad \sigma_{t3'}^2 = \left|\frac{\partial G_3'}{\partial g}\right|^2 \sigma_\epsilon^2 + \sigma_{t3}^2, \tag{4.16}$$

where σ_{t1}, σ_{t3} are the dispersions of flexion data. The first term on the right is scatter due to intrinsic ellipticity and shear measurement error.

In principle, the total number of flexion is the same as the number of galaxy images $N_f = N_g$. But in reality, not all the 3rd-order brightness moments are measurable, besides that, we discard the low signal-to-noise images. Thus the number of images from which flexion can be estimated is usually smaller than the number of galaxy images.

To find the minimum χ^2-function, we solve the equations

$$\frac{\partial \chi^2(\psi_i)}{\partial \psi_i} = \frac{\partial \chi_s^2(\psi_i)}{\partial \psi_i} + \frac{\partial \chi_\epsilon^2(\psi_i)}{\partial \psi_i} + \frac{\partial \chi_f^2(\psi_i)}{\partial \psi_i} + \eta\frac{\partial R(\psi_i)}{\partial \psi_i} = 0, \tag{4.17}$$

which is in general a non-linear set of equations. This problem is solved by an iterative procedure (Bradač et al. 2005b). We perform a three-level iteration process, start with initial model κ^0 at a fixed regularization, linearize the system (Sect. 4.2.3), solve the linear system of equations as the inner-level, the steps are repeated until convergence of κ. The middle-level is to repeat the inner level with new regularization κ from previous step result, until we can get $\chi_{\mathrm{red}}^2 \sim 1$. At last we increase the number of grid points in the field and reperform the first two level iterations on the new grids with new κ until we reach the final grid size.

For the regularization we choose

$$R = \eta \sum_{i,j=1}^{N_x,N_y} \left(\kappa_{ij}^{(n)} - \kappa^{(n-1)}(\boldsymbol{\theta}_{ij})\right)^2, \tag{4.18}$$

where $\kappa^{(n)}$ and $\kappa^{(n-1)}$ are the current and previous resulting κ map in every middle-level. The parameter η is chosen due to the quality of the data (Bradač et al. 2005b).

For the first step, we can use an initial model $\kappa^{(0)}$ which is obtained from other methods, or simple set $\kappa^{(0)} = 0$ across the whole field.

4.2.2 Grid point potential field

Finite differencing techniques (Abramowitz & Stegun 1972) provides a way to calculate κ and γ on the potential ψ grid field. We need 9 grid points for second order quantities κ and γ, and 16 points for flexion. But by using 4×4 points, we obtain the flexion value at the center of the grid. Thus we increase to 5×5 points to get the flexion value at the grid. For instance,

$$\mathcal{F}_1(i,j) = \frac{1}{12\Delta^3} \sum_{\Delta_i,\Delta_j=-2}^{+2} W_{\Delta_i,\Delta_j}\psi(i+\Delta_i,j+\Delta_j), \tag{4.19}$$

where the coefficients W_{Δ_i,Δ_j} used for spin-1 flexion are given in the left panel of Fig. 4.2. Bilinear interpolation can be used for the points at arbitrary position in the field. Therefore,
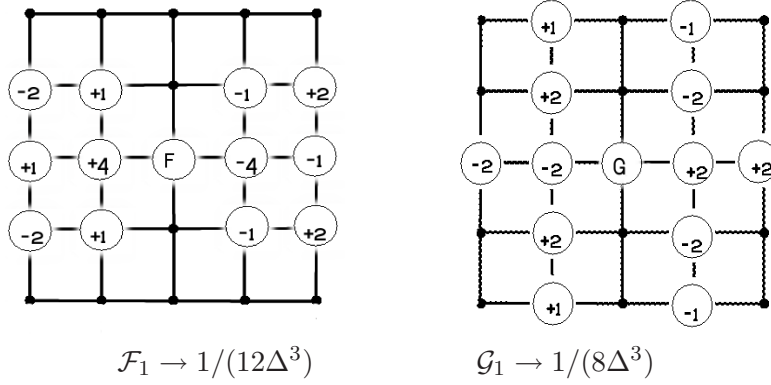
$$\mathcal{F}_1 \to 1/(12\Delta^3) \qquad\qquad \mathcal{G}_1 \to 1/(8\Delta^3)$$

Figure 4.2: The finite differencing coefficients of $\mathcal{F}_1$(left) and $\mathcal{G}_1$(right). And the coefficients of $\mathcal{F}_2(\mathcal{G}_2)$ are the same as that of $\mathcal{F}_1(\mathcal{G}_1)$ after rotating $\pi/2$ anticlockwise(clockwise).

we need to extend two grid rows and columns on each boundaries of the entire field, which means we use $(N+4)^2$ grids to perform the method for inner N^2 grids. But the flexion drops quickly with the increasing distance to the cluster, in our case there are few images with flexion signal near the boundary.

Since shear and flexion are the second- and third-order derivatives of ψ, the potential field is not fixed under the shear and flexion constrain. This is not important to us because κ is also an invariant under the transformation $\psi(\boldsymbol{\theta}) \to \psi(\boldsymbol{\theta}) + \psi_0 + \boldsymbol{\alpha} \cdot \boldsymbol{\theta}$, where ψ_0 and $\boldsymbol{\alpha}$ are arbitrary constants. We leave the constants free to simplify the process of solving equations. The mass-sheet degeneracy transformation of the potential is given by $\psi \to \psi' = (1-\lambda)\theta^2/2 + \lambda\psi$. This transformation doesn't change the reduced shear but affects the κ result by

$$\kappa' = (1 - \lambda) + \lambda\kappa, \tag{4.20}$$

which we called λ transformation. We will take advantage of this transformation for adjusting κ later.

4.2.3   The linearization of the equations

We present here the details on how to linearize and solve the equation

$$\frac{\partial \chi_s^2(\psi_i)}{\partial \psi_i} + \frac{\partial \chi_\epsilon^2(\psi_i)}{\partial \psi_i} + \frac{\partial \chi_f^2(\psi_i)}{\partial \psi_i} + \eta \frac{\partial R(\psi_i)}{\partial \psi_i} = 0. \tag{4.21}$$

The lensing quantities are calculated by finite differencing, and are thus linear combinations of ψ at each position. They are expressed in the following matrix notations (Fig. 4.2)

$$\kappa(\theta_i) = M_{ik}^\kappa \psi_k; \; \gamma_1(\theta_i) = M_{ik}^{\gamma 1} \psi_k; \; \gamma_2(\theta_i) = M_{ik}^{\gamma 2} \psi_k; \tag{4.22}$$

$$\mathcal{F}_1(\theta_i) = M_{ik}^{\mathrm{f1}} \psi_k; \; \mathcal{F}_2(\theta_i) = M_{ik}^{\mathrm{f2}} \psi_k; \; \mathcal{G}_1(\theta_i) = M_{ik}^{\mathrm{g1}} \psi_k; \; \mathcal{G}_2(\theta_i) = M_{ik}^{\mathrm{g2}} \psi_k. \tag{4.23}$$

Then we plug these into Eq.(4.17) and obtain the entire expression of the equations. Here we show the first flexion term as an example. The strong lens multiple images, shear and regularization part can be found in Bradač et al. (2005b) and the second flexion term will be the same as the first one.

$$\chi_{\mathrm{f1}}^2(\psi) = \sum_{i=1}^{N_{\mathrm{f}}} \frac{|(1-\kappa)t_1 - F|^2}{(1-\kappa)^2 \sigma_f^2}, \tag{4.24}$$

where t_1 is already modified by Eq.(4.14), since in this section we only work with the modified ones in the following, therefore drop the $'$ of t_1. We omit index i to all parameters of every galaxy for simplicity, since they are all different. We fix the denominator $\hat{\sigma}_f^2 = (1 - \kappa)^2 \sigma_f^2$ as constant at each step. Then they will not appear in the derivative

$$
\begin{aligned}
\frac{\partial \chi_{\mathrm{f1}}^2}{\partial \psi} &= \sum_{i=1}^{N_{\mathrm{f}}} \frac{-2}{\hat{\sigma}_f^2} \sum_{r=1,2} \left[((1 - \kappa)t_{1r} - F_r)(t_{1r}\frac{\partial \kappa}{\partial \psi} + \frac{\partial F_r}{\partial \psi}) \right] \\
&= \frac{-2}{\hat{\sigma}_f^2} [M_{ij}^{\mathrm{f1}} M_{ik}^{\mathrm{f1}} + M_{ij}^{\mathrm{f2}} M_{ik}^{\mathrm{f2}} + (t_{11}^2 + t_{12}^2)M_{ij}^{\kappa} M_{ik}^{\kappa} \\
&\quad + t_{11}(M_{ij}^{\mathrm{f1}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{f1}} M_{ij}^{\kappa}) + t_{12}(M_{ij}^{\mathrm{f2}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{f2}} M_{ij}^{\kappa})]\psi_k \\
&\quad + \frac{2}{\hat{\sigma}_f^2} \left[t_{11}M_{ij}^{\mathrm{f1}} + t_{12}M_{ij}^{\mathrm{f2}} + (t_{11}^2 + t_{12}^2)M_{ij}^{\kappa} \right],
\end{aligned} \tag{4.25}
$$

where t_{11} and t_{12} are the two components of the spin-1 flexion estimator of galaxy image. The same calculation can be performed to spin-3 flexion, and a similar formulation is obtained.

It is easy to separate the terms with or without ψ, and written Eq.4.17 in the form

$$
B_{ij}\psi_i = V_j, \tag{4.26}
$$

where the matrix B_{ij} and vector V_j contain the contributions from the nonlinear part. The regularization term is already linear in ψ_k, therefore the full expression of matrix is given

$$
\begin{aligned}
B_{ij} &= \sum_{i=1}^{N_{\mathrm{f}}} \frac{1}{\hat{\sigma}_{t1}^2} [M_{ij}^{\mathrm{f1}} M_{ik}^{\mathrm{f1}} + M_{ij}^{\mathrm{f2}} M_{ik}^{\mathrm{f2}} + t_{11}(M_{ij}^{\mathrm{f1}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{f1}} M_{ij}^{\kappa}) \\
&\quad + t_{12}(M_{ij}^{\mathrm{f2}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{f2}} M_{ij}^{\kappa}) + (t_{11}^2 + t_{12}^2)M_{ij}^{\kappa} M_{ik}^{\kappa}] \\
&\quad + \frac{1}{\hat{\sigma}_{t3}^2} [M_{ij}^{\mathrm{g1}} M_{ik}^{\mathrm{g1}} + M_{ij}^{\mathrm{g2}} M_{ik}^{\mathrm{g2}} + t_{31}(M_{ij}^{\mathrm{g1}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{g1}} M_{ij}^{\kappa}) \\
&\quad + t_{32}(M_{ij}^{\mathrm{g2}} M_{ik}^{\kappa} + M_{ik}^{\mathrm{g2}} M_{ij}^{\kappa}) + (t_{31}^2 + t_{32}^2)M_{ij}^{\kappa} M_{ik}^{\kappa}] \\
&\quad + \sum_{i=1}^{N_{\mathrm{Gal}}} \frac{1}{\hat{\sigma}_{\epsilon}^2} [M_{ij}^{\gamma 1} M_{ik}^{\gamma 1} + M_{ij}^{\gamma 2} M_{ik}^{\gamma 2} + \epsilon_1(M_{ij}^{\gamma 1} M_{ik}^{\kappa} + M_{ik}^{\gamma 1} M_{ij}^{\kappa}) \\
&\quad + \epsilon_2(M_{ij}^{\gamma 2} M_{ik}^{\kappa} + M_{ik}^{\gamma 2} M_{ij}^{\kappa}) + (\epsilon_1^2 + \epsilon_2^2)M_{ij}^{\kappa} M_{ik}^{\kappa}] \\
&\quad + \sum_{m=1}^{N_M} \frac{1}{\sigma_s^2} \left[b_{11}(m)D_{mj}^{(1)} D_{mk}^{(1)} + b_{22}(m)D_{mj}^{(2)} D_{mk}^{(2)} \right] + \eta \sum_g M_{gj}^{\kappa} M_{gk}^{\kappa},
\end{aligned} \tag{4.27}
$$

where the sums over i, g and m denote summation over all galaxies images, all grid points, and all images in the multiple imaged system, respectively. $D^{(1)}$ and $D^{(2)}$ are the matrices for the deflection angle, which can be found in Bradač et al. (2005b). The V_j vector is in form

of

$$
\begin{aligned}
V_j &= \sum_{i=1}^{N_{\mathrm{f}}} \frac{1}{\hat{\sigma}_{t1}^2} [t_{11} M_{ij}^{\mathrm{f1}} + t_{12} M_{ij}^{\mathrm{f2}} + (t_{11}^2 + t_{12}^2) M_{ij}^{\kappa}] \\
&\quad + \frac{1}{\hat{\sigma}_{t3}^2} [t_{31} M_{ij}^{\mathrm{g1}} + t_{32} M_{ij}^{\mathrm{g2}} + (t_{31}^2 + t_{32}^2) M_{ij}^{\kappa}] \\
&\quad + \sum_{i=1}^{N_{\mathrm{Gal}}} \frac{1}{\hat{\sigma}_{\epsilon}^2} \left[\epsilon_1 M_{ij}^{\gamma 1} + \epsilon_2 M_{ij}^{\gamma 2} + (\epsilon_1^2 + \epsilon_2^2) M_{ij}^{\kappa} \right] \\
&\quad + \sum_{m=1}^{N_M} \frac{1}{\sigma_s^2} \left[b_1(m) D_{mj}^{(1)} + b_2(m) D_{mj}^{(2)} \right] + \eta \sum_{g} \kappa^{(n-1)} M_{gj}^{\kappa}.
\end{aligned} \tag{4.28}
$$

4.2.4 Numerical test with NIS toy model

First we perform this method on a Non-singular Isothermal Sphere model. The data is generated on a 40×40 grid potential field

$$
\psi(\theta) = \frac{\theta_{\mathrm{E}}}{2} \sqrt{\theta^2 + \theta_c^2} + \frac{\theta_{\mathrm{E}} \theta_c}{2} \ln \left(\frac{\sqrt{\theta^2 + \theta_c^2} - \theta_c}{\theta} \right), \tag{4.29}
$$

where $\theta_{\mathrm{E}} = 1$ arcmin and $\theta_c = 0.2$ arcmin. The reduced shear and reduced flexion are created by finite differencing. We only use the data points of which the absolute value of reduced shear is smaller than 0.9. Fig. 4.3 shows the original and reconstructed radial κ and $|g|$ profiles. The value is obtained by annular bin average on the grid field, the input one and the result one. The small fluctuations of the result point to the input line is mainly due to noise and low resolution. In the left panel, we plot convergence κ after applying the λ transformation Eq.(4.20). We can see that convergence is well recovered for large θ until the vertical line. For the small θ, the plus points from shear diverge from the input line. However, the result from shear and flexion is significantly improved. In the right panel, the absolute value of the reduced shear is shown, which is not affected by mass-sheet degeneracy. Both results by shear or shear and flexion show agreement to the input model. Again for small θ, the result from shear and flexion combination gives better agreement with the input line.

4.2.5 Simulated cluster data

We also test our method on two simulated cluster. Our mock data uses clusters from N-body simulation by Jing & Suto (2002); Jing (2002), the cluster is simulated in the framework of the ΛCDM model with cosmological parameters $\Omega_{\Lambda} = 0.7$, $\Omega_{\mathrm{m}} = 0.3$, the normalization of power spectrum 0.9 and the Hubble constant $H_0 = 70 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$. Dark matter halos are identified with the friends-of-friends method using a linking length equal to 0.2 times the mean particle separation. The halo mass M is defined as the virial mass enclosed within the virial radius according to the spherical collapse model (Kitayama & Suto 1996; Bryan & Norman 1998; Jing & Suto 2002) The virial mass of the cluster which we used here is $3.4 \times 10^{14} \mathrm{h^{-1} M_{\odot}}$ and the redshift is at 0.326. The particles within a box with side length of 2 virial radii were projected onto the lens plane. The surface densities are calculated using the smoothed particle hydrodynamics smoothing algorithm (Monaghan 1992) on a 4096×4096
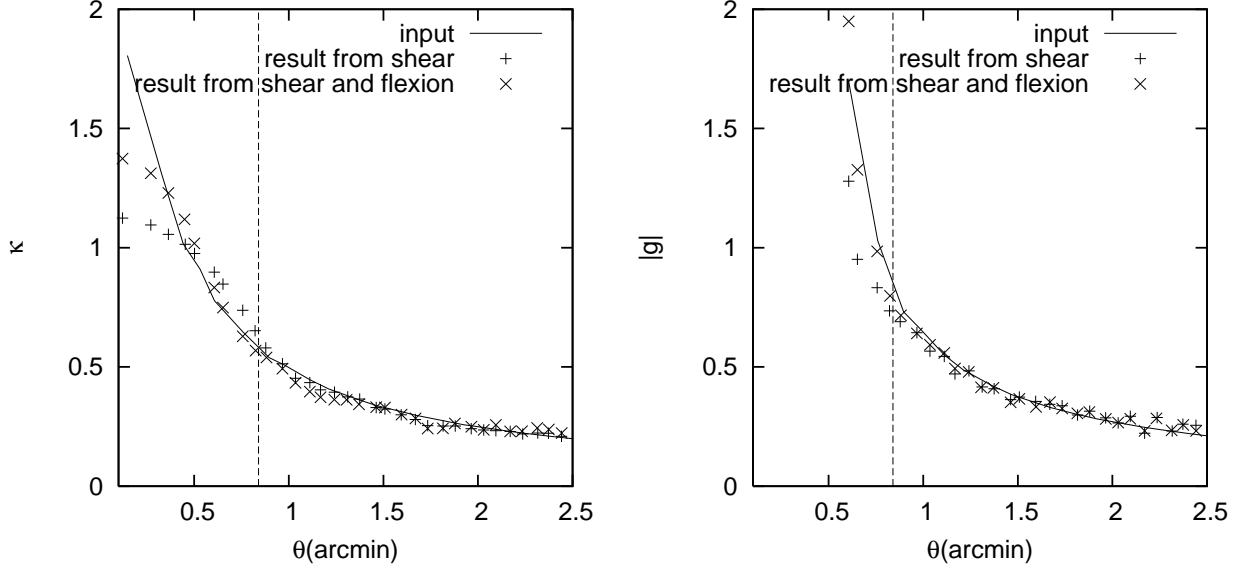
Figure 4.3: Radial profile of the NIS cluster, the solid line is the input model, the plus points are the reconstructed results with weak lensing shear only, and the cross points are the results with weak lensing shear and flexion. The mock data is generated only in the region to the right side of the vertical line. Left panel: convergence κ after λ transformation, Right panel: the absolute value of reduced shear $|g|$, all for a source at $z_s \to \infty$.

grid. The lensing potential is obtained using the fast Fourier transform method (Bartelmann et al. 1998).

A finer grid 0.1 arcsec resolution is created at the cluster center with side length of 4 arcmin. We obtain the source position by lensing equation for each grid point and the corresponding second-order derivatives using cubic spline interpolation. We perform cubic spline interpolation again to the second-order derivatives to get the third-order derivatives. The background sources are located at different source planes randomly and their redshift follows a Gamma distribution

$$p(z) = \frac{z^2}{2\,z_0^3} \exp\left(-z/z_0\right) \; , \tag{4.30}$$

where $z_0 = 1/3$. The peak is at $z = 2/3$ and the mean redshift is $\langle z \rangle = 3z_0 = 1$. Newton-Raphson method is used to find the corresponding images position on the image plane. The reduced shear and flexion on each image were linear interpolated using the four nearest grid points.

One simulated cluster with two different projected directions is used to generate the two different sets of mock data (Fig. 4.4). We name the two sets of data as d02 and d03. All of the reduced shear data is used in our calculation but for flexion, we only consider the absolute value of the reduced flexion range from 0.01 to 0.5. Since the third-order brightness moments cannot be measured when the flexion becomes sufficient large as we discussed in Chapter 3. On the other hand, very small values of the reduced flexion are not easy to measure, neither would they contribute any information due to the low signal-to-noise.
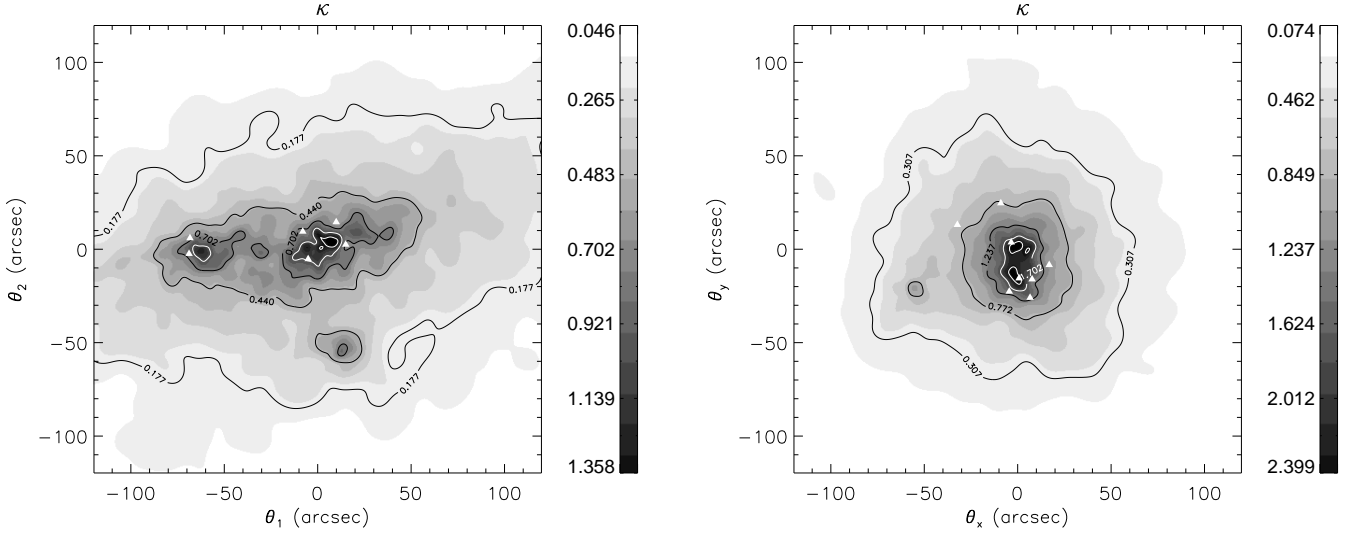
Figure 4.4: The convergence map of two simulated clusters used for generating mock strong lensing multiple images, weak lensing shear and flexion data. The triangles are the strong lensing multiple images. Both are plotted for a source at $z_{\mathrm{s}} \to \infty$. We name the left panel cluster d02 and the right one cluster d03.

4.2.6   Reconstructed κ map

The two mock catalogues are used to test the performance of our method. We start with an initial 20×20 grid, increase N_x and N_y by one each time, to a 30×30 grid. We use $N_{\mathrm{g}} = 1000$ weak lensing galaxy images in each reconstruction, which is an accessible background galaxy number density, ~ 60 images arcmin^{-2}. The result of the reconstructions are shown in Fig. 4.5 for d02 and Fig. 4.6 for d03. The initial regularization parameter is set to $\eta = 200$ for cluster d02, $\eta = 300$ for cluster d03, and increased by 10 for each outer-level iteration to slow the process and make sure we would overfit the noise. It is usually better to set high η and allowing κ to change slowly. Since our reconstruction is done in a three-level iteration, and in each step we ensure $\chi^2/N_{\mathrm{dof}} \sim 1$, the method can successfully adapt to the data and the results are not sensitive to the value of η. We need also an initial κ^0 field for the regularization. A simple model $\kappa^0 = 0.01$ is used here. We have also performed additional reconstructions with different initial models, and found the results are nearly independent of the initial κ^0. But a realistic model allows for a faster convergence.

The number of flexion that we used for the two cluster is different. In reality what we measure are the higher-order brightness moments. We exclude the low signal-to-noise data, thus only use the flexion of which absolue value range from 0.01 to 0.5. We obtained 89 reduced flexion data in Fig. 4.5, and 60 reduced flexion data in Fig. 4.6. Since there is more significant extended structure in cluster d02 than in d03, we found that there are more data points of which the absolute value of spin-1 reduced flexion is between 0.01 and 0.5 in cluster d02 than in cluster d03 in different realizations. We found that there are about 10% images of all the galaxy images which have the suitable flexion value in the cluster d02 field, while this number drops to about 5% in the cluster d03 field.

The results show that our method can reproduce the main properties of the projected mass distribution of both clusters, and is especially powerful in resolving the substructure. Fig. 4.5 shows that our method can reconstruct the κ map by combining strong, weak lensing shear and flexion data for cluster d02. Unfortunately we cannot distinguish the sub-clumps which is due to noise from the true ones even with the help of flexion, like the one in the bottom corner of Fig. 4.5. However we can see that in the right panel, which is the result using strong lensing, weak lensing shear and flexion information, the halos become peaky and the substructure are resolved with correct positions. In Fig. 4.6 it is encouraging to see that besides the main properties of the big dark halo, our method can resolve the small clump after combining the flexion data with shear information. The small clump is not very significant even in the input convergence (Fig. 4.4), and that is with a resolution of 400×400. We also have performed additional tests in which we use different sets of weak lensing shear and flexion data for both clusters. This confirms the validity of our method. In some cases of cluster d03 data, the shape properties can be better reconstructed and the small clump can be clearly resolved, but the position of the small clump can have an up to 10 arcsec offset from its real position. In some other cases the small clump cannot be clearly resolved, which is due to the noise and local low background images density around the small clump.

As an additional test, we calculate the difference D^2 between our result and input κ, which is defined as

$$D^2 = \frac{1}{N'} \sum_{i,j}^{N'} |\kappa_{ij} - \kappa_{ij}^{(\text{input})}|^2, \tag{4.31}$$

where N' is number of grid points which are not inside the critical curves. We use D^2_{sw} for the mean difference between the result from strong lensing and weak lensing shear constraints and input κ, and D^2_{swf} for that from strong lensing, weak lensing shear and flexion constraints. We apply the λ transformation (Eq. 4.20) to the convergence results. The λ is fixed for the smallest difference D^2. For cluster d02, the D^2_{sw} is 0.0086 and D^2_{swf} is 0.0085. For cluster d03, $D^2_{\text{sw}} = 0.0077$ and $D^2_{\text{swf}} = 0.0066$. Since we have strong constraints from strong lensing, the improvement that from flexion is not such strong. In other cases of no strong lensing systems, there is more significant improvement by flexion.

We also change the threshold of flexion, i.e. we use the reduced flexion of which absolute value between $[0.001, 0.5]$. There are more flexion data with low signal-to-noise used in the mass reconstruction. We find that in some cases the D^2 becomes larger after combining the flexion. After checking our result κ on all the grids, we find out the points which gives larger D^2 after combining flexion signal. We argue that the relation of the reduced flexion and the reduced shear (Eq. 4.14) can introduce extra noise from shear into flexion, especially in the case of large intrinsic shear noise. In principle this is not important where flexion is strong, since the shear is also strong, hence the signal-to-noise is high. However in regions where the shear and flexion signals are weak, for instance much lower than the intrinsic ellipticity of galaxies, the intrinsic noise dominates the signal, which means we perform mass reconstruction using noise rather than signal, and combining flexion would reinforce the noise.

Finally a word on the dispersion of flexion σ'_t. This is a difficult parameter to determine at the moment, since we have little knowledge about the noise behavior in flexion measurement. The one we used in this chapter (Eq. 4.16) has a problem: as pointed out by Bacon et al. (2009), the flexion variance is biased by the content of substructure. Moreover, we noticed that the $\sigma_{t'}$ we used is underestimated, it can be seen from following: χ_f^2/N_f is significant smaller than 1, which means that $\chi_\epsilon^2/N_{\text{g}}$ is larger than that without flexion. If we apply more
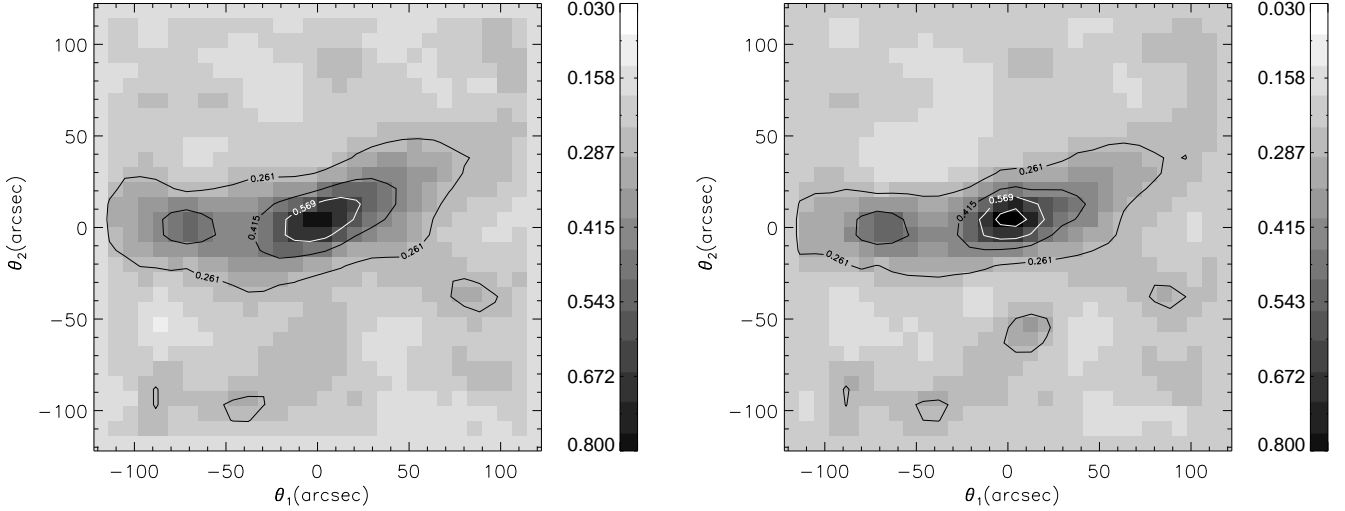
Figure 4.5: κ-maps obtained from shear and flexion reconstruction of the mock data cluster d02 after the λ transformation (Eq.4.20), where λ is choose for the smallest D^2. Left panel shows the result using $N_\text{g} = 1000$ galaxies shear and 3 strong lensing system, while in the right panel we add 89 galaxies flexion. Both figures are given in linear gray-scale and contours.

steps of iteration, the flexion noise might be over fitted. In that case, the cluster becomes very peaky, and looks like being truncated at some edge region. Some other forms for σ'_t have been also tested, e.g. in analogy of shear

$$\sigma_{t'}^2 = \left[1 - (\theta_0|t'|)^2\right]^2 \sigma_{t^s}^2 + \sigma'^2_\text{err}, \tag{4.32}$$

where θ_0 is size of the image. We can easily see that this $\sigma_{t'}$ is not independent of image size and σ'_err is different from that of shear (Eq.4.10), since it is not dimensionless. The result turns out that the improvement by flexion is underestimated, of which the parameter we used is $\sigma_{t^s} = 0.1/''$ and $\sigma'_\text{err} = 0.1/''$.

4.3 Conclusions

In this chapter we proposed a method for projected cluster mass reconstruction, which combines strong, weak lensing shear and flexion data. The method is based on a least-χ^2 fitting of the lensing potential ψ. The particular strength of this method is that the flexion data provides more information to the inner parts of the cluster and substructure.

We test the performance of method on our mock clusters, compare the results with and without flexion. In the NIS cluster, our method can reproduce the radial profile of the convergence and the reduced shear. And flexion can significantly improve the result to the inner part of the cluster. In the other test, we generate our mock data from simulated clusters. We are able to reconstruct the main properties of the cluster mass distribution, especially when the flexion data is included, our method can successfully resolve the cluster
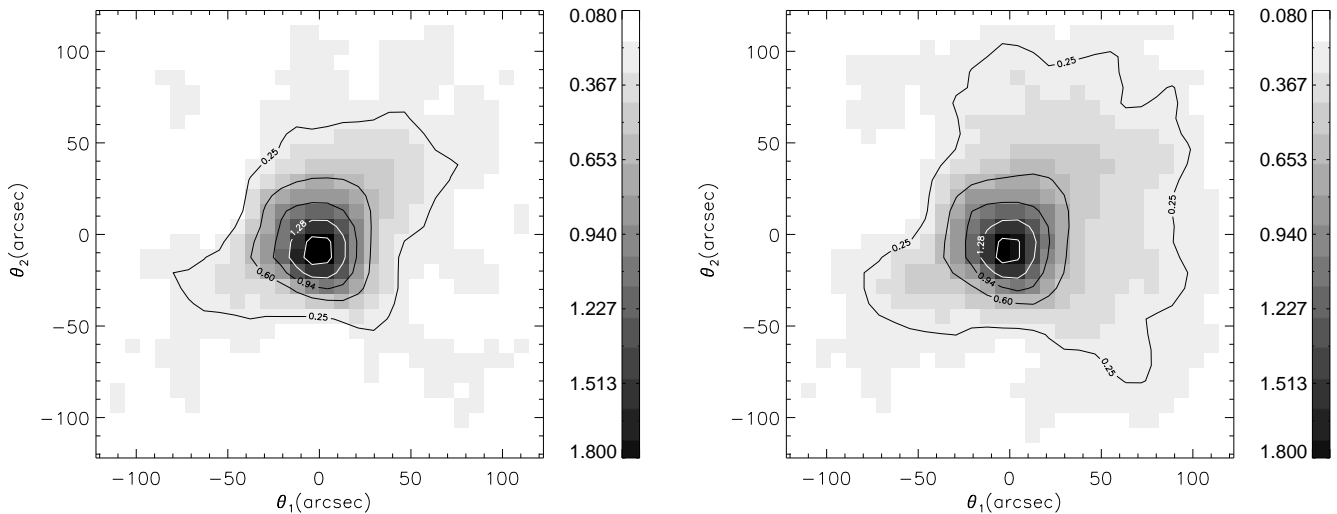
Figure 4.6: Reconstructed κ-maps from mock data cluster d03. Left panel shows result using $N_{\mathrm{g}} = 1000$ galaxies shear, and 3 strong lens. Right panel shows result after combining 60 flexion.

and substructure. In addition, our result is almost independent of the initial model κ^0 and the regularization parameter η.

We have assumed that the intrinsic flexion is small. However, the correction for the reduced flexion introduces extra noise from shear to flexion, especially in the case that the intrinsic galaxy image is highly elliptical. We can ignore the noise where the shear and flexion is strong, since the intrinsic noise is relatively small. This is however not the case in the region where the shear and flexion is not significant. It is of interest to study the relation of intrinsic noise and flexion in detail.

In Leonard et al. (2007); Okura et al. (2008), the result of mass reconstruction by flexion has shown that flexion is sensitive to substructure, and insensitive to the smooth component of the cluster. Our method of combining shear and flexion takes the advantages of shear on the clusters and flexion on the substructure.

Chapter 5

Galaxy-Galaxy Lensing

Galaxy-galaxy lensing is concerned with background galaxies and foreground dark matter halos which host galaxies. It can thus be used as a measure to determine the mass properties of galaxies. In particular it directly probes the matter distribution in the halo. Other techniques, such as rotation curves of spiral galaxies, are less direct since they require assumptions about the dynamical state of the system, and only trace the dark matter halo out to distances of $100 \, h^{-1}$kpc. On the other hand, GGL flexion allows better resolution on small scales than the standard GGL technique with shear measurements; we have already seen that flexion is sensitive to the small-scale variation of the potential field. It is thus an ideal tool for measuring GGL at small separations. In addition, the information obtained from flexion is complementary to that from shear, and thus, flexion can provide additional information about the galaxy halos.

As in the traditional GGL shear measurement, the flexion induced by a single galaxy is not strong enough to be detected, and the flexion signal around many galaxies must be superposed statistically. This can be done for each foreground and background galaxy pair, and the averaging result leads to a circularly symmetric flexion profile

$$
\begin{aligned}
\langle \mathcal{F}(\theta) \rangle &= \frac{1}{N_{\mathrm{fg}}} \int \mathrm{d}^2 \vartheta \, n_{\mathrm{fg}}(\boldsymbol{\vartheta}) \, \mathcal{F}(|\boldsymbol{\theta} - \boldsymbol{\vartheta}|); \\
\langle \mathcal{G}(\theta) \rangle &= \frac{1}{N_{\mathrm{fg}}} \int \mathrm{d}^2 \vartheta \, n_{\mathrm{fg}}(\boldsymbol{\vartheta}) \, \mathcal{G}(|\boldsymbol{\theta} - \boldsymbol{\vartheta}|),
\end{aligned}
\tag{5.1}
$$

where $n_{\mathrm{fg}}(\boldsymbol{\vartheta})$ is the number density of foreground galaxies, and N_{fg} is the total number over the survey. If one defines the fractional galaxy number density $\kappa_{\mathrm{g}} = n_{\mathrm{fg}}/\bar{n} - 1$, where $\bar{n}$ is the galaxy mean number density, this can be written as (showing only the spin-1 flexion for example)

$$
\langle \mathcal{F}(\theta) \rangle = \frac{1}{A} \int \mathrm{d}^2 \vartheta \, \kappa_{\mathrm{g}}(\vartheta) \mathcal{F}(|\boldsymbol{\theta} - \boldsymbol{\vartheta}|),
\tag{5.2}
$$

where A is the area of the survey from which N_{fg} is computed. In reality, it is calculated through $\langle \mathcal{F}(\theta) \rangle = \langle \kappa_{\mathrm{g}}(\vartheta) \mathcal{F}(|\boldsymbol{\theta} - \boldsymbol{\vartheta}|) \rangle$. Under the assumption that the galaxy distribution is homogeneous, the galaxies can be shifted to the origin, thus $\langle \mathcal{F}(\theta) \rangle = \langle \kappa_{\mathrm{g}}(0) \mathcal{F}(|\boldsymbol{\theta}|) \rangle$. Furthermore, in Fourier space, one has $\hat{\mathcal{F}}(\boldsymbol{l}) = -\mathrm{i} \boldsymbol{l} \hat{\kappa}(\boldsymbol{l})$. Therefore the GGL flexion can also be written as

$$
\langle \mathcal{F}(\theta) \rangle = - \int \frac{\mathrm{d}^2 l}{(2\pi)^2} \int \frac{\mathrm{d}^2 l'}{(2\pi)^2} \, \mathrm{e}^{-\mathrm{i} \boldsymbol{\theta} \cdot \boldsymbol{l}'} \, \mathrm{i} \, \langle \hat{\kappa}_{\mathrm{g}}(\boldsymbol{l}) \hat{\kappa}(\boldsymbol{l}') \boldsymbol{l}' \rangle,
\tag{5.3}
$$

which shows explicitly the dependence on the convergence. Since the cross power spectrum of galaxy and convergence is $P_{\kappa g} = (2\pi)^2 \delta_{\mathrm{D}}(\boldsymbol{l} - \boldsymbol{l}')\langle \hat{\kappa}_{\mathrm{g}}(\boldsymbol{l})\hat{\kappa}(\boldsymbol{l}')\rangle$, the extra factor of $\boldsymbol{l}'$ implies that GGL flexion has a greater power at large $\boldsymbol{l}$ than that of GGL shear. In other words, GGL flexion is more sensitive on small scales.

5.1 Circular Halo profiles

The matter distribution in a halo may be approximated by a radial mass profile. The GGL shear and flexion measurements can be used to estimate the radial profile parameters. Such estimates are still useful to constrain halo properties even in the case of non-circularly symmetric halos. Some of the most popular models are described below. Most results of circular halos in this section can be found in Bacon et al. (2006).

5.1.1 Singular Isothermal Sphere

The simplest model for an approximately flat rotation curve is given by $\rho \propto r^{-2}$. Such a profile can be obtained by assuming a constant velocity dispersion for the dark matter throughout the halo, and so it is known as the singular isothermal sphere (hereafter SIS). The projected surface mass density of an SIS is

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi}, \tag{5.4}$$

where ξ is the distance from the center of the lens in the projected lens plane and σ_v is the one-dimensional velocity dispersion of dark matter 'particles' within the gravitational potential. In GGL studies, it is more convenient to use the dimensionless surface mass density $\kappa = \Sigma/\Sigma_{\mathrm{cr}}$, which reads

$$\kappa(\theta) = \frac{\theta_{\mathrm{E}}}{2\theta}, \tag{5.5}$$

where $\theta = \xi/D_{\mathrm{d}}$ is the angular distance from the lens center and $\theta_{\mathrm{E}} = 4\pi(\sigma_v/c)^2 D_{\mathrm{ds}}/D_{\mathrm{s}}$ is the Einstein radius. The shear caused by the SIS halo is

$$\gamma(\boldsymbol{\theta}) = -\frac{\theta_{\mathrm{E}}}{2\theta}\mathrm{e}^{2\mathrm{i}\phi}, \tag{5.6}$$

(Bartelmann & Schneider 2001), where ϕ is the position angle around the lens. Applying Eq.(3.18), the spin-1 flexion is given by

$$\mathcal{F}(\boldsymbol{\theta}) = -\frac{\theta_{\mathrm{E}}}{2\theta^2}\mathrm{e}^{\mathrm{i}\phi}, \tag{5.7}$$

and one can see that the spin-1 flexion for a circularly symmetric halo is directed radially inwards towards the center of the lens (Fig. 3.2). The spin-3 flexion is given by

$$\mathcal{G}(\boldsymbol{\theta}) = \frac{3\theta_{\mathrm{E}}}{2\theta^2}\mathrm{e}^{3\mathrm{i}\phi}. \tag{5.8}$$

It has an amplitude 3 times larger than the spin-1 flexion, drops off with the same power-law index away from the lens, and oscillates around the lens as a spin-3 quantity.
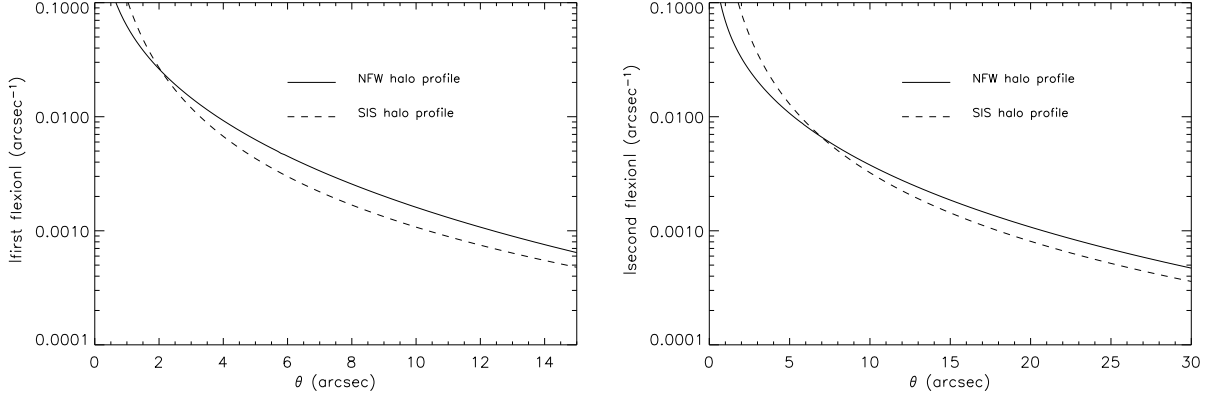
Figure 5.1: Comparison of the magnitude of flexion due to an NFW and SIS halo of $M_{200} = 1 \times 10^{12} \, h^{-1} \, M_\odot$. Left is for spin-1 flexion, right is for spin-3 flexion. The concentration parameter c of the NFW halo is 7.2 and $D_{\rm ds}/D_{\rm s} = 0.5$. Figure is from BGRT06.

5.1.2 Nonsingular Isothermal Sphere

In the SIS model, the mass density becomes infinity for $\theta \to 0$. One simple modification is the non-singular isothermal sphere (NIS) surface mass density

$$\kappa(\theta) = \frac{\theta_{\rm E}}{2\sqrt{\theta^2 + \theta_c^2}}, \tag{5.9}$$

where θ_c is a core radius within which the surface mass density flattens off to a value $\kappa_0 = \theta_E/2\theta_c$. The flexion due to NIS is

$$\begin{aligned}
\mathcal{F}(\theta) &= -\theta \left[\frac{\theta_E}{2(\theta^2 + \theta_c^2)^{3/2}} \right] {\rm e}^{{\rm i}\phi}; \\
\mathcal{G}(\theta) &= \frac{\theta_{\rm E}}{2\theta^3} \left[-8\theta_c + \frac{3\theta^4 + 12\theta^2\theta_c^2 + 8\theta_c^4}{(\theta^2 + \theta_c^2)^{3/2}} \right] {\rm e}^{3{\rm i}\phi}.
\end{aligned} \tag{5.10}$$

For $\theta \gg \theta_c$ the flexion is approximately equal to that of the SIS. At small separations, the flexion goes to zero since the convergence profile tends to be flat for $\theta < \theta_c$.

5.1.3 Navarro-Frenk-White density profile

In N-body simulations of a ΛCDM universe, the density profile of halos was found to have an universal form, i.e., independent of halo mass and size. This profile is shallower in the inner region than the SIS one (Navarro et al. 1996, 1997). We have already presented this NFW profile in Chapter 1. The NFW model is characterized by two parameters. Here we use the scaling radius r_s and the concentration c. The dimensionless surface mass density of an NFW profile reads

$$\kappa(x) = 2\kappa_s \frac{f(x)}{x^2 - 1}, \tag{5.11}$$

with

$$f(x) = \begin{cases} 1 - \dfrac{2}{\sqrt{x^2 - 1}}\arctan\sqrt{\dfrac{x - 1}{x + 1}} & (x > 1) \\[3mm] 1 - \dfrac{2}{\sqrt{1 - x^2}}\operatorname{artanh}\sqrt{\dfrac{1 - x}{1 + x}} & (x < 1), \end{cases} \tag{5.12}$$

where we define $x = r/r_s$, and $\kappa_s = \rho_{\mathrm{crit}}r_s\delta_c/\Sigma_{\mathrm{crit}}$, with the critical density ρ_{crit}. The collapsing density contrast δ_c is related with c from numerical simulations as detailed in Navarro et al. (1997). Here we use

$$\delta_c = \frac{200}{3}\frac{c^3}{\ln(1 + c) - c/(1 + c)}. \tag{5.13}$$

For an axially symmetric mass profile, the shear can be calculated from $\gamma(x) = [\kappa(x) - \bar{\kappa}(x)]\exp[2\mathrm{i}\phi]$, where $\bar{\kappa}(x)$ is the mean surface mass density within a circle of radius x from the lens center (Bartelmann & Schneider 2001). Thus,

$$\gamma(x) = \kappa_s l(x)\mathrm{e}^{2\mathrm{i}\phi}, \tag{5.14}$$

with

$$l(x) = \begin{cases} 1 - \dfrac{8\arctan\sqrt{(x - 1)/(x + 1)}}{x^2\sqrt{x^2 - 1}} + \dfrac{4}{x^2}\ln\left(\dfrac{x}{2}\right) - \dfrac{2}{x^2 - 1} + \dfrac{4\arctan\sqrt{(x - 1)/(x + 1)}}{(x^2 - 1)^{3/2}} & (x > 1) \\[3mm] 1 - \dfrac{8\operatorname{arctanh}\sqrt{(1 - x)/(x + 1)}}{x^2\sqrt{1 - x^2}} + \dfrac{4}{x^2}\ln\left(\dfrac{x}{2}\right) - \dfrac{2}{x^2 - 1} + \dfrac{4\operatorname{arctanh}\sqrt{(1 - x)/(x + 1)}}{(x^2 - 1)(1 - x^2)^{1/2}} & (x < 1). \end{cases} \tag{5.15}$$

Defining $\mathcal{F}_s = \kappa_s D_{\mathrm{d}}/r_s$, the first flexion is calculated using $\mathcal{F} = \nabla_c\kappa = \partial\kappa/\partial x \exp[\mathrm{i}\phi]$, writes

$$\mathcal{F} = -\frac{2\mathcal{F}_s}{(x^2 - 1)^2}[2xf(x) - h(x)]\mathrm{e}^{\mathrm{i}\phi}, \tag{5.16}$$

with

$$h(x) = \begin{cases} 1 - \dfrac{2x}{\sqrt{x^2 - 1}}\arctan\sqrt{\dfrac{x - 1}{x + 1}} - 1/x & (x > 1) \\[3mm] 1 - \dfrac{2x}{\sqrt{1 - x^2}}\operatorname{arctanh}\sqrt{\dfrac{1 - x}{x + 1}} - 1/x & (x < 1), \end{cases} \tag{5.17}$$

and $f(x)$ given by Eq.(5.12). The analytic expression of the spin-3 flexion due to an NFW halo is calculated using

$$\mathcal{G} = \nabla_c\gamma = \left(\frac{\partial\gamma}{\partial x} + \frac{\mathrm{i}}{x}\frac{\partial\gamma}{\partial\phi}\right)\mathrm{e}^{\mathrm{i}\phi}. \tag{5.18}$$

Plugging in the expression of γ, we finally obtain

$$\mathcal{G} = 2\mathcal{F}_s\left[\frac{8}{x^3}\ln\frac{x}{2} + \frac{(3/x)(1 - 2x^2) + g(x)}{(x^2 - 1)^2}\right]\mathrm{e}^{3\mathrm{i}\phi}, \tag{5.19}$$

where

$$g(x) = \begin{cases} 1 - (\dfrac{8}{x^3} - \dfrac{20}{x} + 15x)\dfrac{2}{\sqrt{x^2 - 1}}\arctan\sqrt{\dfrac{x - 1}{x + 1}} & (x > 1) \\[3mm] 1 - (\dfrac{8}{x^3} - \dfrac{20}{x} + 15x)\dfrac{2}{\sqrt{1 - x^2}}\operatorname{arctanh}\sqrt{\dfrac{1 - x}{x + 1}} & (x < 1). \end{cases} \tag{5.20}$$

Fig. 5.1 shows the radial profiles of the flexion for both the NFW and SIS models. Note that the spin-3 flexion is larger in amplitude than the spin-1 flexion, as is the case for the SIS results. Moreover, the flexions due to the SIS profile are stronger than those due to NFW for small separations, since the NFW density on small scales varies as r^{-1} compared to the steeper r^{-2} for the SIS (BGRT06).

From the studies of GGL for circular symmetric halo, we see that we can use the measurements of shear and flexion to constrain the parameters of galaxy halo, such as σ_v^2 for SIS profile or r_s and c for NFW profile. In next sections, we will present more applications of GGL flexion on more complicate cases.

5.2 Radial and Tangential Flexion

We have seen that the spin-1 flexion is a vector-like quantity, and it is the gradient of the surface mass density. Therefore, the spin-1 flexion is directed towards the center of the lens in the case of a circularly symmetric lens. However, this is not a realistic case. There are several factors that make it unlikely for the spin-1 flexion to be directly towards the center of the lens. Measurement noise and systematics are two such factors, but there are two other ones of particular interest to us: elliptical halos and halo substructures. We address these issues in the following sections.

In the general case, the spin-1 flexion can be decomposed into two parts, which are named radial flexion and tangential flexion, as sketched in Fig. 5.2. They are defined as

$$
\begin{align}
\mathcal{F}_R &= -\boldsymbol{\mathcal{F}} \cdot \hat{r}; \tag{5.21}\\
\mathcal{F}_T &= \boldsymbol{\mathcal{F}} \cdot \hat{\phi} = \pm |\boldsymbol{\mathcal{F}} + \mathcal{F}_R \hat{r}|, \tag{5.22}
\end{align}
$$

where $\hat{r}$ and $\hat{\phi}$ are the unit direction vectors. Thus $\boldsymbol{\mathcal{F}} = -\mathcal{F}_R \hat{r} + \mathcal{F}_T \hat{\phi}$. This is similar to the $\mathcal{F}_R$ and $\mathcal{F}_B$ in Hawken & Bridle (2009); Leonard et al. (2007)

$$
\mathcal{F}_R = |\mathcal{F}|\cos(\theta'_F); \quad \mathcal{F}_B = |\mathcal{F}|\sin(\theta'_F), \tag{5.23}
$$

where $\theta'_F = \theta - \phi + \pi$, and θ is the angle of spin-1 flexion stretch to the coordinate axes of θ_1, θ_2. The $\mathcal{F}_B$ here does not have the same physical meaning as the cosmic shear B-mode, which is mainly due to the non-gravitational systematics, whereas the tangential flexion (or $\mathcal{F}_B$) is due to non-circular lens contribution. The decomposition in radial and tangential flexion can be easily performed in a GGL measurement. For the spin-3 flexion, there is not such a clear intuitive picture of the radial or tangential components. We mainly consider the spin-1 flexion in the following discussion.

5.3 Elliptical Halos

Most galaxies are elliptical or spiral, i.e. non circularly symmetric. Then a question may arise naturally, whether the dark halo has the same shape as the galaxy. If there exists a significant difference, that would be a strong evidence for dark matter and rule out modified gravity models. For instance, elliptical galaxy dark matter halos disfavor the MOdified Newtonian Dynamics paradigm (Hoekstra et al. 2004; Mandelbaum et al. 2006). Numerical simulations with different assumptions predict halos with different shapes, e.g. simulation with non-interacting cold dark matter predict that halos are triaxial prolate ellipsoids (Allgood et al.
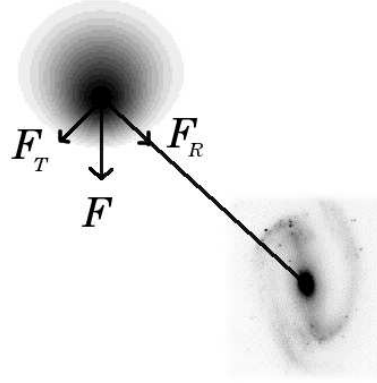
Figure 5.2: Sketch of the galaxy-galaxy lensing flexion. The spin-1 flexion $\mathcal{F}$ is decomposed into radial flexion $\mathcal{F}_R$ and tangential flexion $\mathcal{F}_T$.

2006). Therefore, the shape of dark matter halos can provide constraints on galaxy formation models and the nature of dark matter.

Hawken & Bridle (2009) discussed the use of flexion to measure halo ellipticity, and found that the constrains from flexion are comparable and tighter than those from shear. Flexion appears thus to be a potential tool for measuring the shapes of galaxy halos in the future.

We have shown the flexion profile for a circular symmetric galaxy halo. For an elliptical halo, it is difficult to obtain an analytic expression for flexion, especially for the spin-3 flexion. Here we first consider the singular isothermal elliptical (SIE hereafter) halo case. Its surface mass density profile can be written as

$$\kappa(\theta_1, \theta_2) = \frac{\theta_E}{2\sqrt{\theta_1^2/(1+\epsilon)^2 + \theta_2^2/(1-\epsilon)^2}}, \tag{5.24}$$

where the major axis a of the elliptical isodensity contours lies along the θ_1 axis, and the ellipticity ϵ is defined by

$$\epsilon = \frac{a-b}{a+b}. \tag{5.25}$$

Then the spin-1 flexion vector is

$$\mathcal{F}(\theta_1, \theta_2) = -\frac{\theta_E(1-\epsilon^2)}{2} \frac{(1-\epsilon)^2\theta_1 + \mathrm{i}(1+\epsilon)^2\theta_2}{\left[\theta_1^2(1-\epsilon)^2 + \theta_2^2(1+\epsilon)^2\right]^{3/2}}. \tag{5.26}$$

This vector field is shown in Fig. 5.3 for the case of $\epsilon = 0.3$. The radial flexion, according to the definition (Eq.5.21), is

$$\mathcal{F}_R(\theta_1, \theta_2) = \frac{\theta_E(1-\epsilon^2)}{2} \frac{1}{\left[\theta_1^2(1-\epsilon)^2 + \theta_2^2(1+\epsilon)^2\right]^{1/2}\left[\theta_1^2 + \theta_2^2\right]^{1/2}}, \tag{5.27}$$

or, in polar coordinates,

$$\mathcal{F}_R(\theta, \phi) = \frac{\theta_E(1-\epsilon^2)}{2\theta^2} \frac{1}{(1 - 2\epsilon\cos 2\phi + \epsilon^2)^{1/2}}. \tag{5.28}$$

The tangential flexion is given by

$$
\begin{aligned}
\mathcal{F}_T(\theta_1, \theta_2) &= \theta_E(1-\epsilon^2) \left[\frac{-2\epsilon\theta_1\theta_2(\theta_1 - \mathrm{i}\theta_2)}{\left[\theta_1^2(1-\epsilon)^2 + \theta_2^2(1+\epsilon)^2\right]^{3/2}(\theta_1^2 + \theta_2^2)} \right]; \\
\mathcal{F}_T(\theta, \phi) &= \frac{\theta_E(1-\epsilon^2)}{\theta^2} \frac{\epsilon \sin 2\phi}{(1 - 2\epsilon \cos 2\phi + \epsilon^2)^{3/2}}.
\end{aligned}
\tag{5.29}
$$

Here we choose the clockwise direction to be the positive one. This choice is not important to determine the halo ellipticity, which depends only on the absolute value. It is convenient to use the ratio of the tangential and radial components as an indicator of the halo ellipticity

$$
r(\theta_1, \theta_2) = \frac{\mathcal{F}_T}{\mathcal{F}_R} = \frac{4\theta_1\theta_2\epsilon}{\theta_1^2(1-\epsilon)^2 + \theta_2^2(1+\epsilon)^2},
\tag{5.30}
$$

or in polar coordinates,

$$
r(\phi) = \frac{2\epsilon \sin 2\phi}{\cos^2\phi(1-\epsilon)^2 + \sin^2\phi(1+\epsilon)^2} = \frac{2\epsilon \sin 2\phi}{1 - 2\epsilon \cos 2\phi + \epsilon^2}.
\tag{5.31}
$$

Note that the ratio is independent of θ_E and θ, i.e., does not depend on the mass of the halo, and the distance to the center of the halo. Therefore, its detection would be a clean measurement of the halo ellipticity. Essentially, the ratio of the tangential flexion to radial flexion is the tangent of the angle between the direction to the center and the actual direction of the flexion vector. Fig. 5.4 shows the flexion ratio field produced by an nonsingular isothermal elliptical (NIE hereafter) halo. One can see that there is a significant variation of the flexion ratio with orientation ϕ, and that in this case (NIE halo), it does not depend on the distance to the center of the halo. In the very inner part of the halo, where for the core model κ is flat, both components become zero. Flexion cannot be measured in this region, where the flexion becomes significantly large, and flexion concept break down. We take average of the flexion ratio over the polar angle (Eq.5.31), and obtain a simple estimator of the ellipticity of the SIE halo,

$$
\langle r \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\phi \left| \frac{2\epsilon \sin 2\phi}{1 - 2\epsilon \cos 2\phi + \epsilon^2} \right| = \frac{2}{\pi} \ln\frac{1+\epsilon}{1-\epsilon}.
\tag{5.32}
$$

Note that we take the absolute value of the ratio because the angle-average ratio is 0. Due to its simplicity, this first flexion ratio test can be easily performed in measurements of a flexion field over a given aperture.

5.3.1 Numerical test with NIE model

We perform the ellipticity estimator on nonsingular isothermal elliptical halos. The data is generated on a 11×11 grid convergence field

$$
\kappa(\theta_1, \theta_2) = \frac{\theta_{\mathrm{E}}}{2\sqrt{\theta_c^2 + \theta_1^2/(1+\epsilon)^2 + \theta_2^2/(1-\epsilon)^2}},
\tag{5.33}
$$

where $\theta_{\mathrm{E}} = 6$ arcsec, and we consider $\theta_c = 1, 2$ arcsec for two cases. The data points of which $|\mathcal{F}| > 0.5$ or close to the center of the halo $|\boldsymbol{\theta}| < 6$ arcsec are discarded. 100 flexion data on 1×1 arcsec^2 are used in the end.

Figure 5.3: Spin-1 flexion vector field for an elliptical isothermal density distribution with $\epsilon = 0.3$. $\mathcal{F}$ only points towards the center when the background galaxy is located on the major or minor axis of the halo.
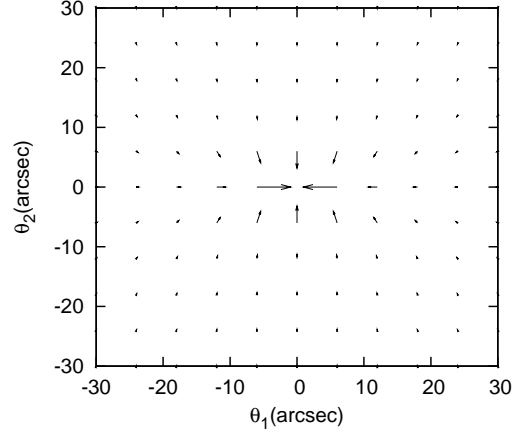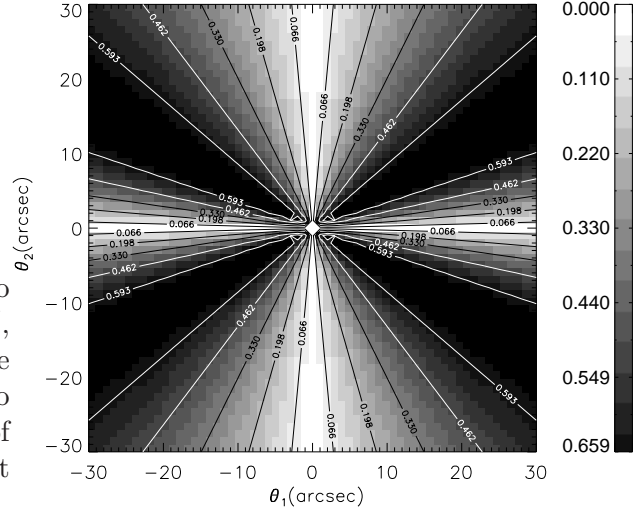

Figure 5.4: Illustration of the flexion ratio field for an elliptical core halo with $\theta_E = 6''$, $\theta_c = 2''$ and $\epsilon = 0.3$. It shows absolute values, given in linear gray-scale. The ratio is zero along the major and minor axis of the halo, where the tangential component vanishes.

We first calculate the flexion ratio for an NIE halo with $\epsilon = 0.3$. The result is shown in Fig. 5.5, the dash line is for an SIE halo (Eq.5.31), the points are result calculated from NIE model. One can see that in this ideal case, the points agree the line, and there is almost no difference between the SIE halos and NIE halo with $\theta_c = 1$ arcsec and $\theta_c = 2$ arcsec. The reason for that is that the data points that we used are outside θ_E ($> \theta_c$), and thus are not affected much by θ_c.

In addition, we test the estimator of ellipticity. 20 sets of data using Eq.(5.33) are generated with different ϵ. For each set of data, ensemble average of flexion ratio $\langle r \rangle$ is calculated, and the estimate of ellipticity is obtained through Eq.(5.32). In Fig. 5.6, the results for two kinds of halos are shown. Again, there is not difference between the halos with two θ_c. One can see that for ϵ from 0.03 to 0.6, the estimate agree with the input value.
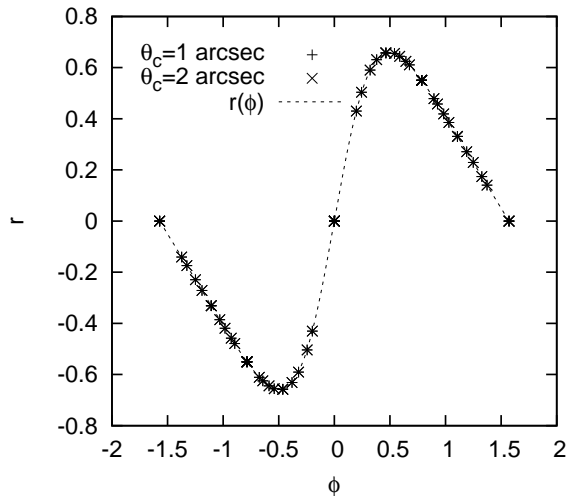
Figure 5.5: The flexion ratio r varies with ϕ. The dash line is for an SIE halo with $\epsilon = 0.3$ (Eq.5.31). The points are calculated from NIE models with $\theta_c = 1$ arcsec (plus) and $\theta_c = 2$ arcsec (cross).
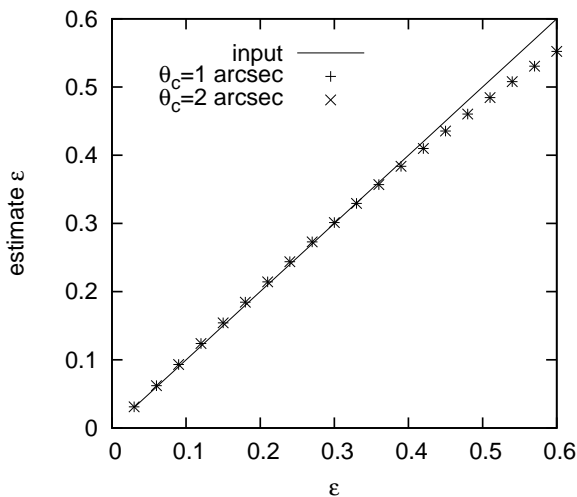

Figure 5.6: Comparison of the ellipticity estimator (Eq.5.32) and the input value. 20 ϵ (from 0.03 to 0.6) from NIE halos with $\theta_c = 1$ arcsec (plus) and $\theta_c = 2$ arcsec (cross) are shown.

5.4 Aperture flexion

Technically, this section is not really an application of GGL. But since it is tightly following the concept of radial flexion, I thus write them in the same chapter. In reality, the aperture flexion can be applied to galaxy halos and galaxy clusters.

The idea of aperture flexion is similar to the aperture mass statistic. In Schneider (1996), the formalism for the generalized aperture mass statistic for shear was shown. In Leonard et al. (2009), an approach of aperture mass for flexion was expressed. And here we present the statistic of aperture flexion, which has one integration step less.

In the previous section, we have seen that the radial flexion is caused by the gradient of mass, whereas the tangential flexion mainly responds to the asymmetry of the dark halo. We thus define the radial- and tangential- flexion aperture separately

$$Arf(\boldsymbol{x}_0, r) = \int_0^\infty \mathrm{d}^2 x \, \mathcal{F}_{\mathrm{R}}(x) \, W(r, x); \tag{5.34}$$

$$Atf(\boldsymbol{x}_0, r) = \int_0^\infty \mathrm{d}^2 x \, \mathcal{F}_{\mathrm{T}}(x) \, W(r, x), \tag{5.35}$$

where x_0 is the center of the aperture, $\mathcal{F}_\mathrm{R}$ and $\mathcal{F}_\mathrm{T}$ are with respect to the x_0. $W(r, x)$ is the weight function with characteristic radius r. Here one should notice that the lower limit of the integration should not be 0 since the flexion cannot be measured once it becomes sufficient large. This is however not a real problem when we perform the aperture flexion on the real data. We write down a discrete version of Eq.(5.35) as

$$
\begin{aligned}
Arf(\boldsymbol{x}_0, r) &= \frac{\sum_i^{N_\mathrm{g}} \mathcal{F}_\mathrm{R}(x_i)\, W(r, x_i)}{\sum_i^{N_\mathrm{g}} W(r, x_i)}; \\
Atf(\boldsymbol{x}_0, r) &= \frac{\sum_i^{N_\mathrm{g}} \mathcal{F}_\mathrm{T}(x_i)\, W(r, x_i)}{\sum_i^{N_\mathrm{g}} W(r, x_i)}.
\end{aligned}
\tag{5.36}
$$

The denominator in Eq.(5.36) is to avoid the bias from the number density of background galaxy images. Or we can define the signal to noise ratio $S(x)$ in analogy with the shear aperture mass statistic,

$$
S(\boldsymbol{x}_0, r) = \frac{1}{\sigma_\mathcal{F}} \frac{\sum_i^{N_\mathrm{g}} \mathcal{F}_\mathrm{R}(x_i)\, W(r, x_i)}{\sqrt{\sum_i^{N_\mathrm{g}} W^2(r, x_i)}},
\tag{5.37}
$$

where $\sigma_\mathcal{F}$ is the dispersion of flexion which is to be obtained from measurements. The signal to noise ratio is only defined for Arf, since Atf is only sensitive to the asymmetry of mass distribution in the aperture, therefore here it is mainly used for a check for the quality of data.

5.4.1   Weight function

Two different families of weight functions are considered for the aperture flexion. The signal of aperture flexion depends on various factors, such as the structure of the lens cluster and the size of aperture. The optimal choice of weight function is strongly dependent on the flexion profiles of the structures being studied. The weight function used here is not chosen to be optimal for a lens profile, which allows us to detect the (sub-)structures without a priori information of the cluster.

We use two simple weight functions,

$$
\begin{aligned}
w_1(x) &= \frac{1}{\pi r^2} \mathrm{Exp}\left(\frac{-x^2}{r^2}\right); \\
w_2(x) &= \frac{x^2 \mathrm{e}}{\pi r^4} \mathrm{Exp}\left(\frac{-x^2}{r^2}\right),
\end{aligned}
\tag{5.38}
$$

where r is the characteristic scale or the size of the aperture. In Fig. 5.7, one can see that $w_1(x)$ is a normal Gaussian filter, which peaks at $x = 0$, but $w_2(x)$ peaks at $x = r$. If we calculate the expected signal to noise, again we need to notice the lower limit in the integration, since flexion cannot be measured when it become significant large, and the consequence would be that the signal to noise is overestimated. For the weight function w_2, the signal to noise decreases with increasing aperture size r (???). On the other hand, small size aperture will lower the number of galaxy images. There is no unique way to choose the size of aperture, especially in clumpy clusters, with different sizes of substructures. We thus stack several apertures of different r. This is equivalent to using the weight function that sum up different r.
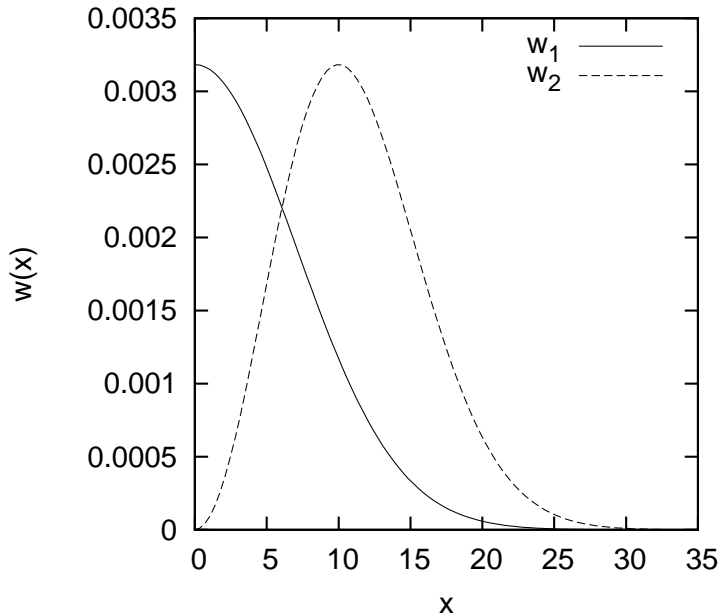
Figure 5.7: The two weight functions (Eq.5.38) with characteristic scale $r = 10$.

5.4.2 Mock data

To test the aperture flexion method, we use a simulated cluster, which is taken from the Millennium Simulation (Springel et al. 2005). The cluster is at redshift $z = 0.62$. Besides the dark matter of the cluster, stellar mass components are added (Hilbert et al. 2008). The flexion maps are created by using the multiple-plane ray-tracing method (Hilbert et al. 2009). All the sources are at redshift $z = 2.1$.

We use 64×64 background images on a grid field of 6×6 arcmin^2. The data with $|\mathcal{F}| > 0.5$ arcsec^{-1} are discarded since they cannot be measured.

5.4.3 Aperture flexion from mock data

To compute the Arf on the field, we use a grid of 177×177 apertures of radius of $r = 6i$ arcsec, where $i = 1, 2...10$. Two cases are shown in Fig. 5.10 for w_1 and Fig. 5.11 for w_2. In both small radius $6''$ maps, there are several peaks corresponding to the cluster and substructures. In larger radius $36''$ maps, the peaks only show the center of the cluster, and are not sensitive to the substructures which are smaller than the aperture. However, the stacked map (Fig. 5.12) takes the advantage of all apertures sizes.
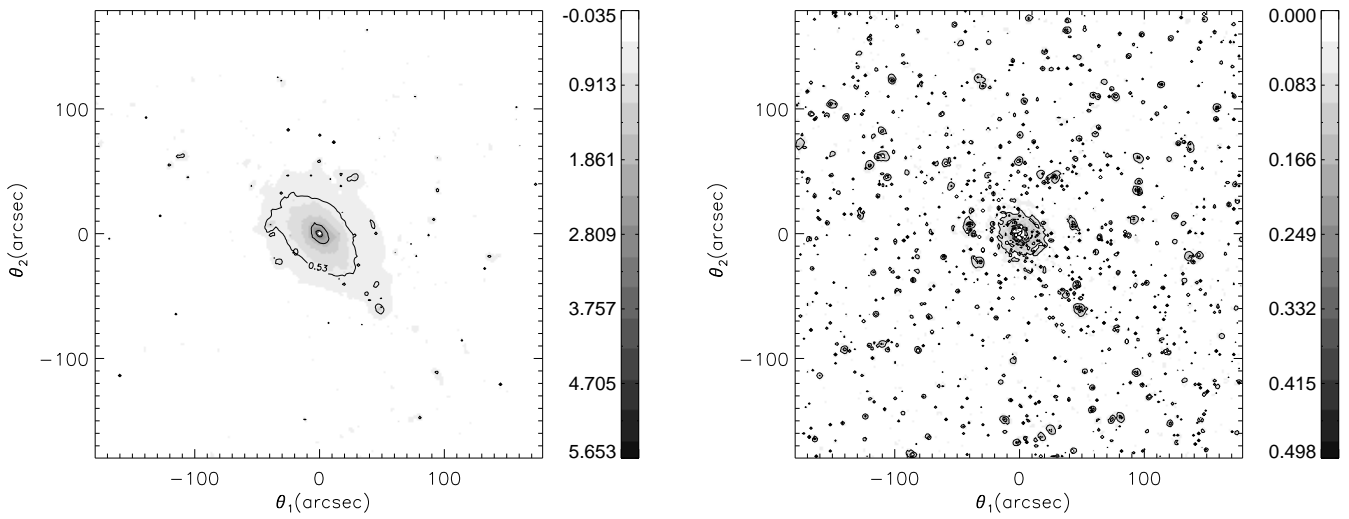
Figure 5.8: The lensing properties of a simulated cluster used for generating the mock data. Left: The surface mass density κ, right: the absolute value of the spin-1 flexion $|\mathcal{F}|$, both for sources at redshift $z = 2.1$, and are given in linear gray-scale and contours.
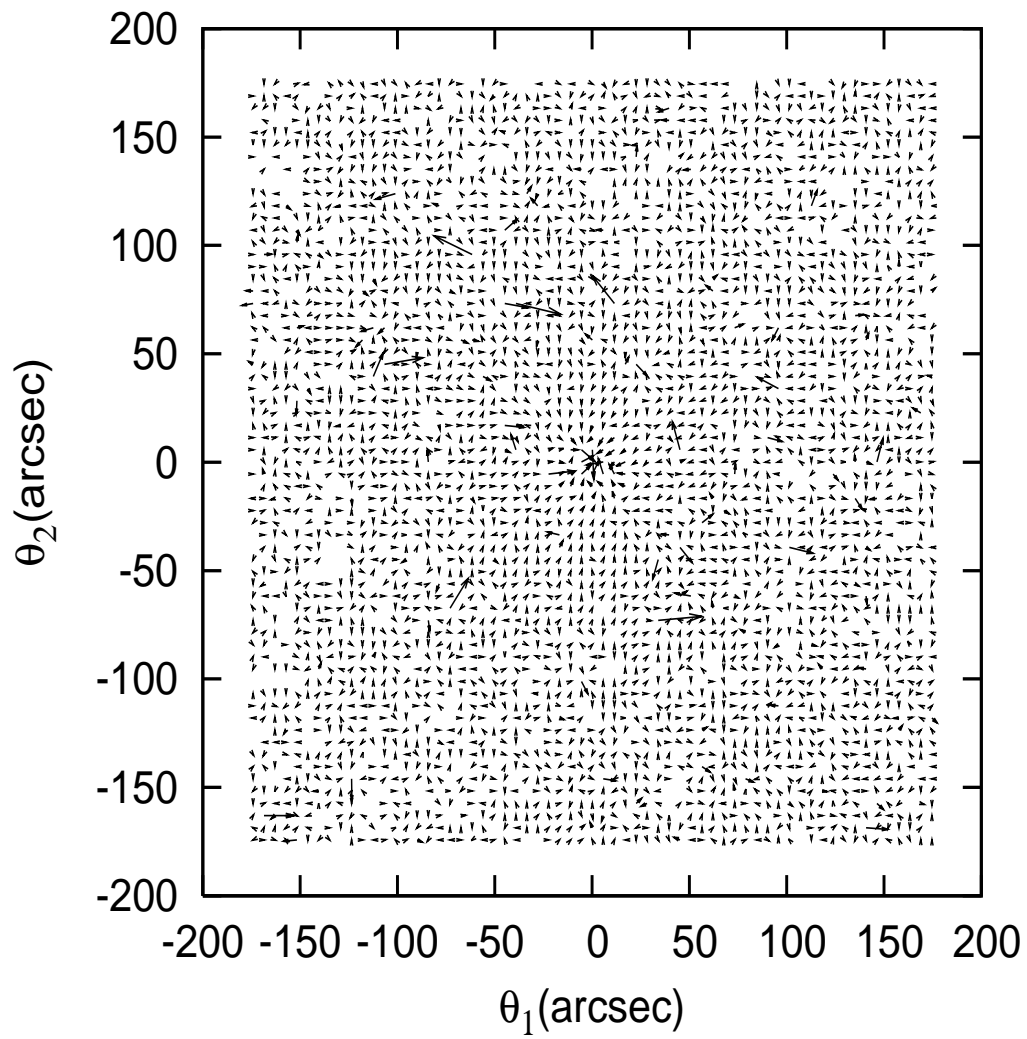
Figure 5.9: The flexion vector field of the simulated cluster in Fig. 5.8. The points for which $|\mathcal{F}| > 0.5$ arcsec^{-1} are omitted.
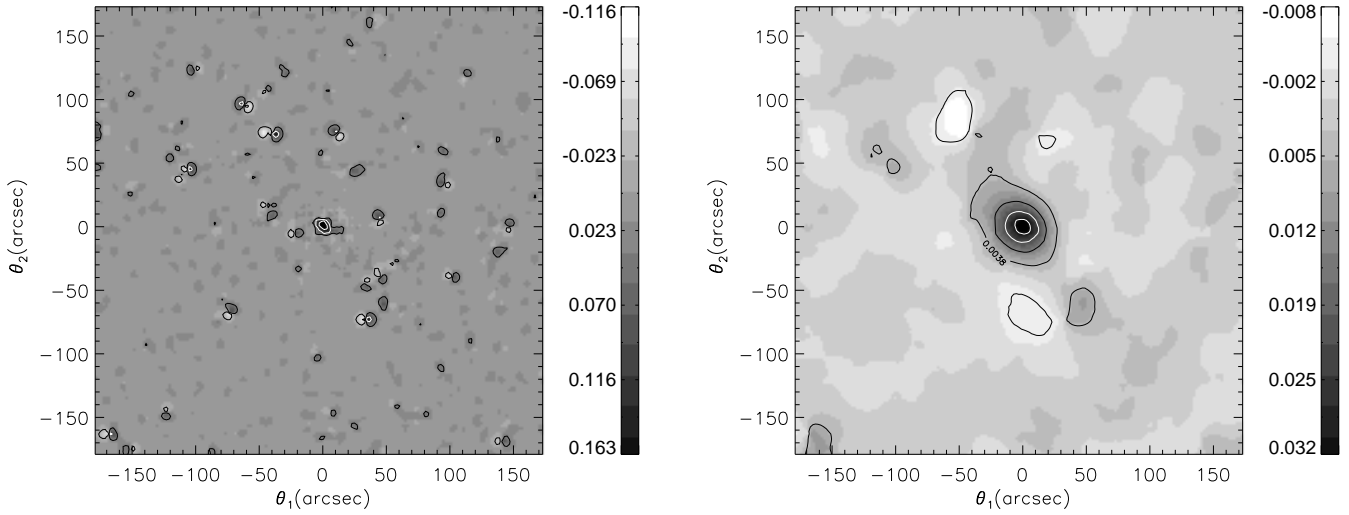
Figure 5.10: Aperture radial flexion with radius of $6''$ and $36''$ of weight function $w1(x)$.
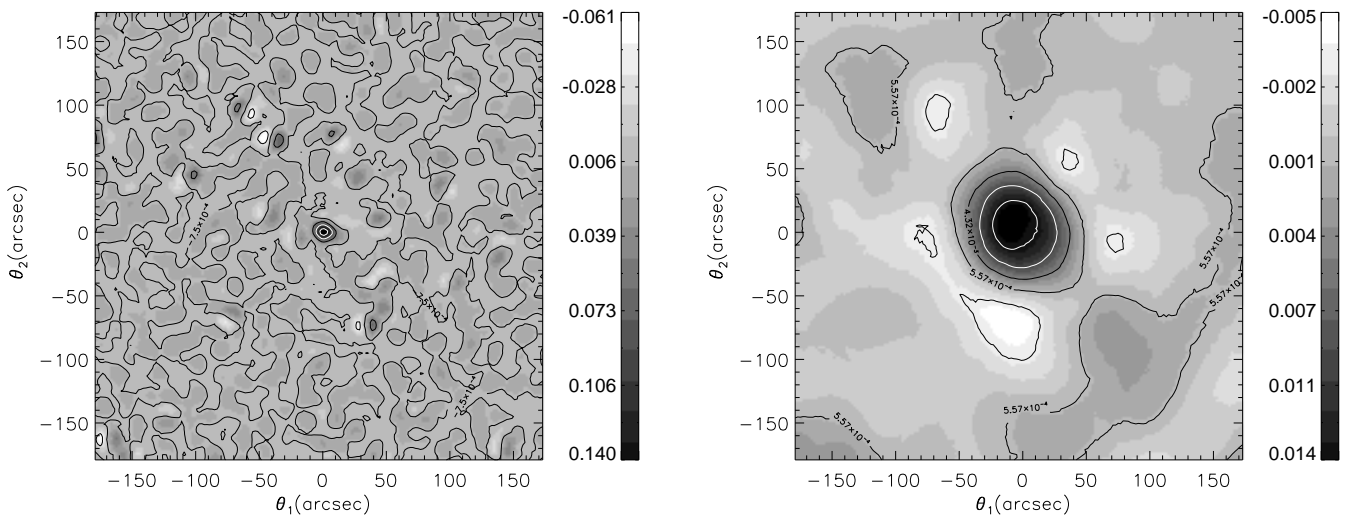

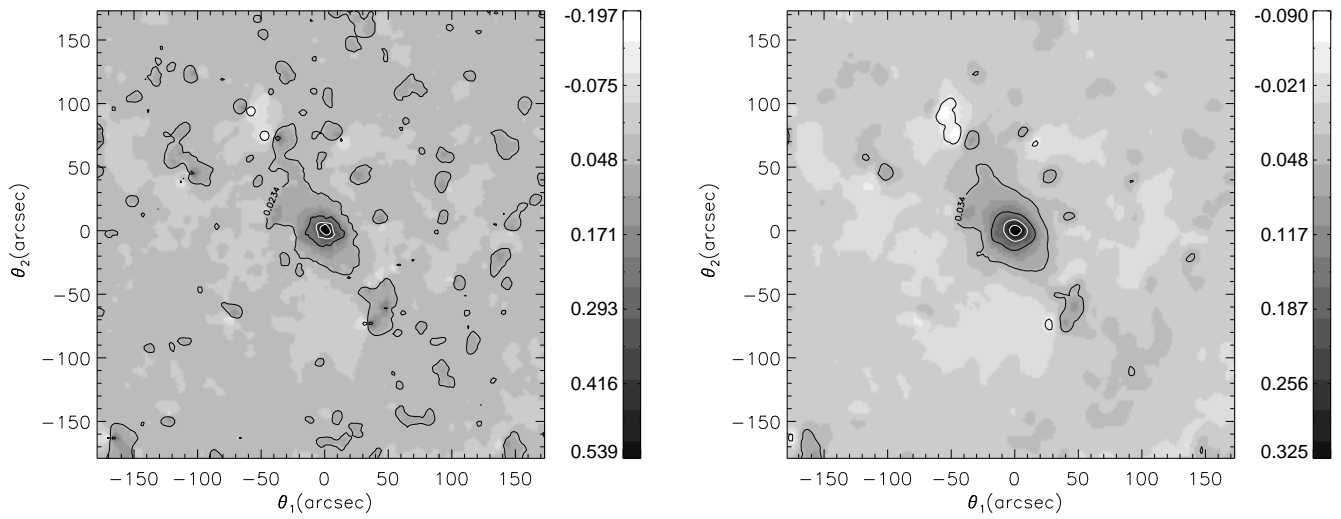Figure 5.11: Aperture radial flexion with radius of $6''$ and $36''$ of weight function $w2(x)$.

Figure 5.12: Stacked aperture radial flexion map of w_1 (left) and w_2 (right) for 10 different characteristic radius r, which are taken from $6''$ to $60''$ increasing by $6''$ each time.

Chapter 6

Summary and Outlook

6.1 Summary

In this thesis, I studied the higher order weak gravitational lensing effect, flexion, and some of its potential applications in cosmology.

Flexion, as a new subject in lensing, has shown its advantages in various studies (at least theoretically). Flexion is more sensitive to small-scale variations in the potential than shear, and therefore has a higher signal-to-noise, in some regions than the shear.

In Chapter 1, I briefly introduced the standard model of cosmology, the ΛCDM model. Although supported by most observations, such as CMB and large-scale galaxy surveys, the ΛCDM model still has several unsolved problems. The very big unknowns are the nature of dark energy and dark matter, which comprise more than 95% of the content of the whole universe. In order to understand that, we need to know how the universe evolves and how the matter is distributed in the universe.

Gravitational lensing offers a powerful tool to study the matter distribution in the universe. I presented the basic concepts and definitions of gravitational lensing in Chapter 2, focusing on weak lensing. The central observation in weak lensing studies is the ellipticity of background galaxies. Due to mass-sheet degeneracy, the ellipticity is only an estimate of a reduced shear. Further, I present some lensing statistical quantities since a single image does not provide enough information. There are the 2-point correlation functions and aperture mass which are widely used in weak lensing studies. Various approaches are used to analyze different properties on different scales, such as cosmic shear, cluster mass reconstruction and galaxy-galaxy lensing.

In Chapter 3, I introduced the higher-order weak lensing effect – flexion – and studied the effect of flexion in weak gravitational lensing. A general flexion field can be decomposed into a pair of components which is due to a shear field and a pair of components not related to shear. The former pair can be further separated into flexion due to an E- and B-mode shear, with only the E-mode flexion expected to arise from gravitational lensing. For the second pair of components are most due to noise or intrinsic shape effects of source. Owing to the mass-sheet degeneracy, only reduced flexion can be measured. The second-order lens equation is given as well as the relations between the brightness moments of source and image in terms of the reduced shear and the reduced flexion. I present approximate estimates for the reduced shear and flexion using these moments equations. In a number of numerical tests I have studied the bias of the reduced flexion estimators. The product of flexion and source size

matters in the accuracy of estimates. I also pointed out a limit where the flexion formalism ceases to be valid, namely when the product of source size and flexion is sufficiently large that parts of the source are multiply imaged locally, i.e., where a caustic cuts through the source. I gave this limitation in some cases and also a complete classification of the critical curves of the second-order lens equation employed in flexion studies.

Chapter 4 deals with galaxy clusters, which are an important probe of the matter content of the universe. I present some methods for cluster mass reconstruction, especially detailed the method which combines strong lensing, weak lensing shear and flexion information. This method allows one to extend the weak lensing analysis into the inner part of the clusters and of substructures within the clusters. We tested the method with numerical simulations, finding an agreement between the input and reconstructed mass also on the substructures. Using flexion allows us to obtain a significant improvement on the results of inner part of clusters and resolve the substructures. We conclude that with high resolution imaging data the method can accurately reconstruct cluster masses and substructures.

In Chapter 5, I showed preliminary results on galaxy-galaxy lensing. The spin-1 flexion can be decomposed into radial and tangential components, which respond to different properties of the galaxy halo. The ratio of tangential flexion to radial flexion can be used to measure the ellipticity of dark matter halos. In an ideal numerical test, the result perfectly agrees with the input value. I also presented an aperture statistics for the radial flexion, which can be used for substructure detections.

6.2   Outlook

In the next few years, there will be several weak lensing projects, e.g. PanSTARRS[1], KIDS[2], and even space-based missions such as Euclid[3] and SNAP[4]. The number density and quality of the galaxy images will be highly improved. At that time we can really put the flexion to extensive use. But before that, there are several preparation that need to be done.

First of all how to measure flexion from galaxy images. Although some methods have been constructed, e.g. shapelets and higher-order KSB/HOLICs, none of them have been tested by simulation. Thus we have no idea if these flexion estimates would be biased by the magnitude, size, ellipticity of the galaxy images, or especially important PSF. From preliminary studies, the bias of the flexion estimator is size dependent, since flexion has a dimension of length inverse and the intrinsic flexion is inversely proportional to the image size. As mentioned before, the second-order lens equation contains five essential parameters. The bias of an estimator for reduced shear and flexion will depend on these parameters as well as on the intrinsic ellipticity (and higher-order moments) distribution of sources. Beside the bias, the covariance of flexion and shear is also interesting to calculate. Furthermore, the weak lensing community has been studying and comparing several existing different methods to measure shear [5] (Heymans et al. 2006; Massey et al. 2007a), and more new methods (even from computer science) will be studied [6]. The development of different techniques for measuring flexion will be also interesting.

---

[1]http://www.ps1sc.org

[2]http://www.astro.wise.org/projects/KIDS

[3]http://www.dune-mission.net

[4]http://snap.lbl.gov

[5]http://www.physics.ubc.ca/∼heymans/step.html

[6]http://www.great08challenge.info

The intrinsic flexion (higher-order moments) is another important issue in flexion studies. Since the intrinsic flexion is inversely proportional to the image size, the variance of intrinsic flexion is also correlated with the size of the image, and it is thus not a constant. Moreover, what can we really observe is the reduced flexion, which is combination of flexion and reduced shear. This introduces another noise into flexion. Thus what we studied here about flexion is based on very simple intrinsic flexion model. The constraints by flexion in lensing analysis need much more accuracy and careful study.

On the other hand, flexion shows several interesting applications on constraining the substructure within clusters and galaxy halos. Chapter 5 provides a good starting point for further work. On the scale of galaxies, the radial and tangential flexion and their ratio provide a measurement of the ellipticity of the dark matter halo and its orientation. I have only shown the result in ideal cases, namely no intrinsic noise, perfect knowledge of the halo center and only for the NIE halo. It is interesting to see how well can we constrain these parameters for a certain level of errors and whether the estimate is unbiased, and which precision do we need if we want to perform this method on real data. Besides the elliptical halo, using GGL shear and flexion to measure the circular symmetric halo is also interesting. In addition, I present the preliminary study of aperture flexion. In aperture statistics, flexion is separated into radial and tangential components too. Since the tangential flexion mainly responds to the asymmetry of dark halo, and our main goal is to detect the cluster and substructures, the aperture tangential flexion is designed for checking the quality of the data. But the second-order statistics of aperture flexion is unclear, and shows some interesting behavior. It is thus worthy some more studies. Besides that, an optimal weight function should be constructed for the aperture flexion and also an combination of shear and flexion in aperture statistics.

Appendix A

Higher-order KSB

In this Appendix, we come to a realistic method on how to measure flexion. The observed galaxy images are not only distorted by a gravitational field, but also affected by non-gravitational lensing various effects, e.g. the distortion by the telescope and the camera, and the atmosphere in case of ground based telescopes. The image read out from a CCD is a convolution of the unperturbed image with the point-spread function (PSF). In terms of surface brightness, $f(\theta)$, the transformations may be written as

$$I^{\text{obs}}(\theta) = p(\theta) \otimes (G\,I^{\text{source}}(\theta)), \tag{A.1}$$

where G is the gravitational lensing operator and $p(\theta)$ is the surface brightness of the PSF.

The correction of PSF is important for weak lensing studies, otherwise we get a wrong estimate of the lensing shear or flexion. To account for this unwanted distortions, the individual correction have to be taken into account using stellar images as references before being combined into a final image. It is in principle possible, if the PSF is properly sampled on the pixel grid and across the image. Kaiser, Squires & Broadhurst have developed a formalism for the PSF correction, which is called the KSB method (Kaiser et al. 1995; Luppino & Kaiser 1997; Hoekstra et al. 1998). Erben et al. (2001), Bacon et al. (2001) and Hoekstra et al. (2002b) presented tests of KSB+ shape measurement on image simulations containing an artifical shear signal. The PSF distortion can be determined by measuring the shapes of stellar images at different positions, since they are point-like or intrinsically round sources. Since for flexion measurement, higher-order brightness moments are used, the KSB method which is desiged for shear measurements is not accurate enough for that of flexion. Here I outline an upgraded version, higher-order KSB, which can be used for the future flexion measurements. A similar approach (HOLICs) by Okura et al. (2008) was performed on real data for flexion measurement. But what I present here is a bit different treatment about centroid shift (Eq.A.23) and for weighting function (Eq.A.52). The work in this Appendix is done together with Barnaby Rowe.

A.1 Notation

The notations in this chapter are a little different from others, since the higher-order KSB method needs many parameters while they might stand for different quantities in other chapters, and they only hold in this chapter. Firstly, the complex numbers are defined in a similar
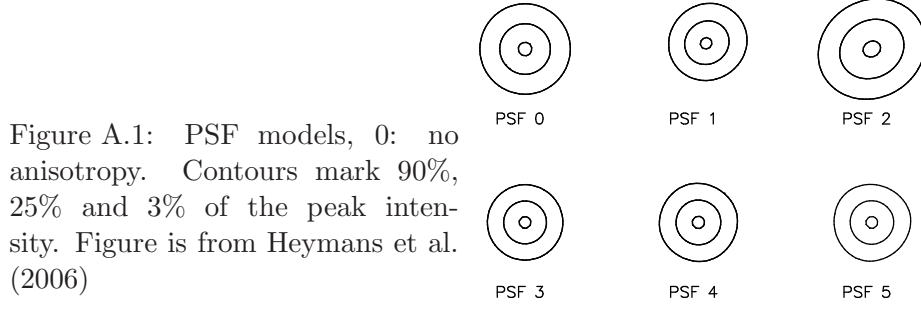
Figure A.1: PSF models, 0: no anisotropy. Contours mark 90%, 25% and 3% of the peak intensity. Figure is from Heymans et al. (2006)

fashion to equation A1 of Okura et al. (2008) (OUF08 hereafter)

$$\theta_M^N = (\theta_1 + i\theta_2)^{\frac{N+M}{2}}(\theta_1 - i\theta_2)^{\frac{N-M}{2}}. \tag{A.2}$$

In this way, notation N gives the order of the complex number, and M gives the spin, for instance $\theta_1^3 = \theta^*\theta\theta$ and $\theta_0^2 = \theta\theta^* = |\theta^2|$ etc. With the complex derivative definition $\partial = \partial_1 + i\partial_2$, the following results is useful:

$$\partial\theta_M^N = (N - M)\theta_{M+1}^{N-1}; \quad \partial^*\theta_M^N = (N + M)\theta_{M-1}^{N-1}, \tag{A.3}$$

from which we can see that the complex derivative is spin up operator, and the conjugate derivative is spin down operator. In particular for the weight function $W = W(\theta_0^2/\sigma^2)$, Eqs.A.3 become

$$\partial W = \frac{2\theta_1^1}{\sigma^2}W', \quad \partial^*W = \frac{2\theta_{-1}^1}{\sigma^2}W', \quad \text{with } W' = \frac{dW}{d(\theta^2/\sigma^2)}. \tag{A.4}$$

The stellar anisotropy moments q_i, q_{ij} and q_{ijk} defined in Bartelmann & Schneider (2001) can also be defined in complex notation

$$d = \int d^2\phi \, q(\phi) \, \phi_1^1 = q_1 + iq_2 \tag{A.5}$$

$$q_0 = \int d^2\phi \, q(\phi) \, \phi_0^2 = q_{11} + q_{22} \tag{A.6}$$

$$q_2 = \int d^2\phi \, q(\phi) \, \phi_2^2 = q_{11} - q_{22} + 2iq_{12} \tag{A.7}$$

$$t_1 = \int d^2\phi \, q(\phi) \, \phi_1^3 = q_{111} + q_{122} + i(q_{112} + q_{222}) \tag{A.8}$$

$$t_3 = \int d^2\phi \, q(\phi) \, \phi_3^3 = q_{111} - 3q_{122} + i(3q_{112} - q_{222}) \tag{A.9}$$

$$f_0 = \int d^2\phi \, q(\phi) \, \phi_0^4 = q_{1111} + 2q_{1122} + q_{2222} \tag{A.10}$$

$$f_2 = \int d^2\phi \, q(\phi) \, \phi_2^4 = q_{1111} - q_{2222} + 2i(q_{1112} + q_{1222}) \tag{A.11}$$

$$f_4 = \int d^2\phi \, q(\phi) \, \phi_4^4 = q_{1111} - 6q_{1122} + q_{2222} + 4i(q_{1112} - q_{1222}) \tag{A.12}$$

Note that for the PSF anisotropy kernel $q(\phi)$, will be defined in Eq.(A.15), is normalised and has zero mean. Thus, we have $q_0 = f_0 = 0$ by definition. The definitions above can be combined with Eq.(A.3) to give results which are necessary when performing Taylor expressions
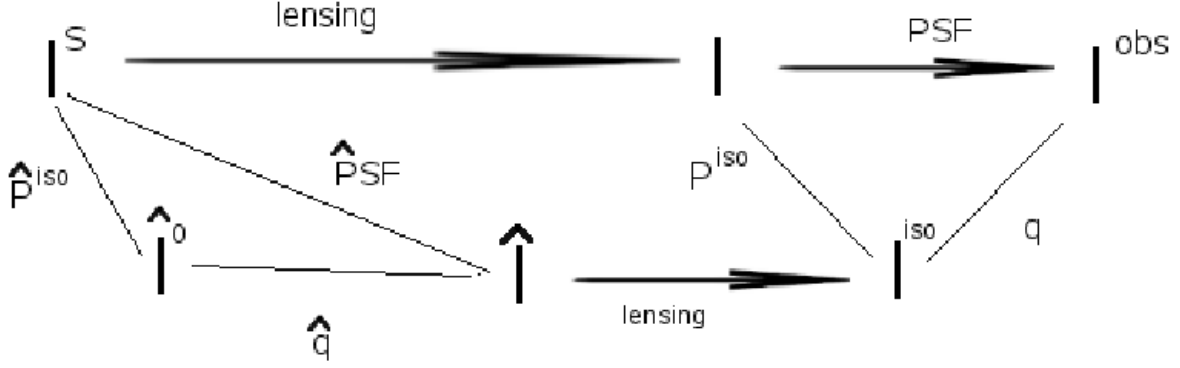
Figure A.2: The logical clue of the KSB method, with the three brightness distribution in the top line are the source brightness I^S, brightness after lensing distortion I, and the brightness after PSF smearing I^{obs}, which is the observed one. The bottom three are hypothetical ones in the KSB method. I^{iso} is the lensed image only after isotropic PSF smearing. $\hat{I}$ is the unlensed source image smeared with $\hat{\text{PSF}}$. $\hat{I}^0$ is the unlensed source image smeared with $\hat{P}^{\text{iso}}$.

using the complex notation.

$$
\begin{aligned}
q_i \frac{\partial}{\partial \theta_i} &= \frac{1}{2} \left(d \partial^* + d^* \partial \right), & \text{(A.13)} \\
q_{ij} \frac{\partial^2}{\partial \theta_i \partial \theta_j} &= \frac{1}{4} \left(q_2 \partial^* \partial^* + 2 q_0 \partial^* \partial + q_2^* \partial \partial \right) = \frac{1}{4} \left(q_2 \partial^* \partial^* + q_2^* \partial \partial \right), \\
q_{ijk} \frac{\partial^3}{\partial \theta_i \partial \theta_j \partial \theta_k} &= \frac{1}{8} \left(t_3 \partial^* \partial^* \partial^* + 3 t_1 \partial^* \partial^* \partial + 3 t_1^* \partial \partial \partial^* + t_3^* \partial \partial \partial \right), \\
q_{ijkl} \frac{\partial^4}{\partial \theta_i \partial \theta_j \partial \theta_k \partial \theta_l} &= \frac{1}{16} \left(f_4 \partial^* \partial^* \partial^* \partial^* + 4 f_2 \partial^* \partial^* \partial^* \partial + 6 f_0 \partial^* \partial^* \partial \partial + 4 f_2^* \partial^* \partial \partial \partial + f_4^* \partial \partial \partial \partial \right) \\
&= \frac{1}{16} \left(f_4 \partial^* \partial^* \partial^* \partial^* + 4 f_2 \partial^* \partial^* \partial^* \partial + 4 f_2^* \partial^* \partial \partial \partial + f_4^* \partial \partial \partial \partial \right).
\end{aligned}
$$

The form of these expressions is a direct consequence of the result for the binomial expansion. Note that each expression can be seen to explicitly maintain the spin properties of the left-hand side. The spin property of complex lensing algebra is useful to check many of the results below in advance by 'spin conservation'.

A.2 Higher-order KSB

We closely follow the logical steps of Section 4.6.2 in Bartelmann & Schneider (2001) (BS01 hereafter). The logical steps of the higher-order KSB method is show in Fig. A.2. Different I stand for 'different step' surface brightness, the mathematic definition will be given later. The 'real' process is that the source brightness is distorted by the gravitational field and

then smeared by the PSF. The KSB method begins with the observed I^{obs} (of course, that is what we have). The first step is the anisotropic correction, then the isotropic smear is done together with the lensing effect, in the end another anisotropic part is needed.

The observed brightness profile is the true surface brightness having been smeared by an anisotropic PSF

$$I^{\mathrm{obs}}(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi \, I(\boldsymbol{\varphi}) \, P(\boldsymbol{\theta} - \boldsymbol{\varphi}). \tag{A.14}$$

The PSF function $P(\boldsymbol{\varphi})$ can be decomposed into two parts

$$P(\boldsymbol{\varphi}) = \int \mathrm{d}^2\vartheta \, q(\boldsymbol{\vartheta}) \, P^{\mathrm{iso}}(\boldsymbol{\varphi} - \boldsymbol{\vartheta}) \tag{A.15}$$

where P^{iso} is the azimuthal average of P, and this is based on the assumption that the anisotropic part of $P(\boldsymbol{\varphi})$ is small compared with P. It is convenient to define the hypothetical brightness profile

$$I^{\mathrm{iso}}(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi \, I(\boldsymbol{\varphi}) \, P^{\mathrm{iso}}(\boldsymbol{\theta} - \boldsymbol{\varphi}), \tag{A.16}$$

which is the true surface brightness smeared only by the isotropic part of the PSF. For later, another profile which is the source surface brightness smeared by the same isotropic part of the PSF, is given by

$$\hat{I}^0(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi \, I^{\mathrm{s}}(\boldsymbol{\varphi}) \, P^{\mathrm{iso}}(\boldsymbol{\theta} - \boldsymbol{\varphi}). \tag{A.17}$$

A.2.1 From I^{obs} quantities to I^{iso} quantities - the "smear" correction including centroid shift

From Eqs.(A.15) and (A.16), one has

$$I^{\mathrm{obs}}(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi q(\boldsymbol{\theta} - \boldsymbol{\varphi}) I^{\mathrm{iso}}(\boldsymbol{\varphi}). \tag{A.18}$$

Using this result, the integral of an arbitrary function $f(\theta)$ can be Taylor expanded to obtain

$$
\begin{aligned}
\int \mathrm{d}^2\theta \, f(\boldsymbol{\theta}) \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \;=\;& \int \mathrm{d}^2\varphi \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \int \mathrm{d}^2\theta \, f(\boldsymbol{\theta} + \boldsymbol{\varphi}) \, q(\boldsymbol{\theta}) \\
\simeq\;& \int \mathrm{d}^2\varphi \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \, f(\boldsymbol{\varphi}) + q_i \int \mathrm{d}^2\varphi \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \, \frac{\partial f}{\partial \varphi_i} \\
& + q_{ij} \frac{1}{2} \int \mathrm{d}^2\varphi \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \, \frac{\partial^2 f}{\partial \varphi_i \partial \varphi_j} + q_{ijk} \frac{1}{6} \int \mathrm{d}^2\varphi \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \, \frac{\partial^3 f}{\partial \varphi_i \partial \varphi_j \partial \varphi_k},
\end{aligned} \tag{A.19}
$$

up to third order in derivatives of f, where q_i etc. are described in Eqs. (A.5-A.9). We then rearrange to give

$$
\begin{aligned}
\int \mathrm{d}^2\varphi \, f(\boldsymbol{\varphi}) \, I^{\mathrm{iso}}(\boldsymbol{\varphi}) \;\simeq\;& \int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \, f(\boldsymbol{\theta}) - q_i \int \mathrm{d}^2\varphi \, I^{\mathrm{obs}}(\boldsymbol{\varphi}) \, \frac{\partial f}{\partial \varphi_i} \\
& - q_{ij} \frac{1}{2} \int \mathrm{d}^2\varphi \, I^{\mathrm{obs}}(\boldsymbol{\varphi}) \, \frac{\partial^2 f}{\partial \varphi_i \partial \varphi_j} - q_{ijk} \frac{1}{6} \int \mathrm{d}^2\varphi \, I^{\mathrm{obs}}(\boldsymbol{\varphi}) \, \frac{\partial^3 f}{\partial \varphi_i \partial \varphi_j \partial \varphi_k},
\end{aligned} \tag{A.20}
$$

Intending only to calculate the PSF anisotropy corrections to linear order, here we have replaced I^{iso} by I^{obs} in the Taylor expansion terms in the above expression. The discrepancy caused by using the latter will yield terms of order q^2 and higher.

One important assumption that has gone into Eq.(A.20) is that the centroid of the object is the same for both the observed and 'iso' profiles: this is not strictly true and needs to be properly accounted for in higher order KSB. In the observed frame, the origin $\bar{\theta}^{\mathrm{obs}} = 0$ with the coordinated θ gives

$$\int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \left(\theta - \bar{\theta}^{\mathrm{obs}}\right)^1_1 W\left[(\theta - \bar{\theta}^{\mathrm{obs}})^2_0/\sigma^2\right] = 0, \tag{A.21}$$

where $W(\theta^2_0/\sigma^2)$ is a weighting function of scale length σ. One effect of the PSF anisotropy will have been to cause a small shift in the image centroid, and so a pre-smear centroid $\bar{\theta}^{\mathrm{iso}}$ is needed, which would be also calculated from the image surface brightness convolved only with the isotropic part of the PSF. This centroid would then be defined by

$$\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta}) \left(\theta - \bar{\theta}^{\mathrm{iso}}\right)^1_1 W\left[(\theta - \bar{\theta}^{\mathrm{iso}})^2_0/\sigma^2\right] = 0. \tag{A.22}$$

If one defines the smear centroid shift $\Delta\bar{\theta}^{\mathrm{sm}} = \bar{\theta}^{\mathrm{obs}} - \bar{\theta}^{\mathrm{iso}} = -\bar{\theta}^{\mathrm{iso}}$, this then gives

$$\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta}) \left(\theta + \Delta\bar{\theta}^{\mathrm{sm}}\right)^1_1 W\left[(\theta + \Delta\bar{\theta}^{\mathrm{sm}})^2_0/\sigma^2\right] = 0. \tag{A.23}$$

Assuming that $\Delta\bar{\theta}^{\mathrm{sm}}$ is small, the Taylor expansion of W to first order around $(\theta - \bar{\theta}^{\mathrm{obs}})$ is

$$W\left[(\theta + \Delta\bar{\theta}^{\mathrm{sm}})^2_0/\sigma^2\right] \simeq W\left[\theta^2_0/\sigma^2\right] + \frac{1}{2}\left(\Delta\bar{\theta}^{\mathrm{sm}}\partial^* + (\Delta\bar{\theta}^{\mathrm{sm}})^*\partial\right) W\left[\theta^2_0/\sigma^2\right].$$

Substituting this into Eq.(A.23) and ignoring terms of order $(\Delta\bar{\theta}^{\mathrm{sm}})^2$ gives the result

$$\begin{aligned}
\Delta\bar{\theta}^{\mathrm{sm}} \simeq &-\frac{1}{S^{\mathrm{iso}}}\left\{\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta})\theta^1_1 W + \frac{\Delta\bar{\theta}^{\mathrm{sm}}}{2}\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta})\theta^1_1 \partial^* W\right. \tag{A.24}\\
&\left.+ \frac{(\Delta\bar{\theta}^{\mathrm{sm}})^*}{2}\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta})\theta^1_1 \partial W\right\},
\end{aligned}$$

where the total flux is defined as

$$S^{\mathrm{iso}} = \int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta}) \, W\left[\theta^2_0/\sigma^2\right]. \tag{A.25}$$

Then, using Eqs. (A.13) and (A.20), and staying to linear order in d, q_2, t_1 and t_3 (Eqs.A.5-A.9), the centroid shift becomes

$$\begin{aligned}
\Delta\bar{\theta}^{\mathrm{sm}} \simeq &\frac{1}{2S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[d\partial^* + d^*\partial\right]\left(\theta^1_1 W\right) \tag{A.26}\\
&+ \frac{1}{8S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[q_2\partial^*\partial^* + q_2^*\partial\partial\right]\left(\theta^1_1 W\right)\\
&+ \frac{3}{48S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_1\partial^*\partial^*\partial + t_1^*\partial\partial\partial^*\right]\left(\theta^1_1 W\right)\\
&+ \frac{1}{48S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_3\partial^*\partial^*\partial^* + t_3^*\partial\partial\partial\right]\left(\theta^1_1 W\right)\\
&- \frac{\Delta\bar{\theta}^{\mathrm{sm}}}{2S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta})\theta^1_1 \partial^* W - \frac{(\Delta\bar{\theta}^{\mathrm{sm}})^*}{2S^{\mathrm{iso}}}\int \mathrm{d}^2\theta \, I^{\mathrm{iso}}(\boldsymbol{\theta})\theta^1_1 \partial W.
\end{aligned}$$

Now is a good time to define a series of weighted image brightness moments, using similarly defined moments in Chapter 3

$$S \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \; W, \tag{A.27}$$

$$D \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_1^1 \; W, \tag{A.28}$$

$$Q_0 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_0^2 \; W, \; Q_2 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_2^2 \; W, \tag{A.29}$$

$$T_1 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_1^3 \; W, \; T_3 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_3^3 \; W, \tag{A.30}$$

$$F_0 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_0^4 \; W, \; F_2 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_2^4 \; W, \; F_4 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_4^4 \; W, \tag{A.31}$$

$$V_1 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_1^5 \; W, \; V_3 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_3^5 \; W, \; V_5 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_5^5 \; W, \tag{A.32}$$

$$H_0 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_0^6 \; W, \; H_2 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_2^6 \; W, \tag{A.33}$$

$$H_4 \;=\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_4^6 \; W, \; H_6 = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_6^6 \; W. \tag{A.34}$$

When integrals are required to be taken over W' (defined in equation A.4), rather than over W, there are prime on the moments also, e.g.

$$Q_2' = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_2^2 \; W', \quad F_4''' = \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta})\theta_4^4 \; W'''. \tag{A.35}$$

Combining these definitions with the differentiation results of Sect. (A.1), one obtains the following result for the centroid shift $\Delta\bar{\theta}^{\mathrm{sm}}$ to first order in d, q_2, t_1 and t_3

$$
\begin{aligned}
\left(1 + \frac{Q_0'}{\sigma^2 S}\right)\Delta\bar{\theta}^{\mathrm{sm}} + \frac{Q_2'}{\sigma^2 S}(\Delta\bar{\theta}^{\mathrm{sm}})^* \;\simeq\;& d\left(1 + \frac{Q_0'}{\sigma^2 S}\right) + d^* \frac{Q_2'}{\sigma^2 S} \\
+\;& q_2\left(\frac{D'}{\sigma^2 S} + \frac{T_1''}{2\sigma^4 S}\right)^* + q_2^* \frac{T_3''}{2\sigma^4 S} \\
+\;& t_1\left(\frac{S'}{\sigma^2 S} + \frac{2Q_0''}{\sigma^4 S} + \frac{F_0'''}{2\sigma^6 S}\right) + t_1^*\left(\frac{3Q_2''}{2\sigma^4 S} + \frac{F_2'''}{2\sigma^6 S}\right) \\
+\;& t_3\left(\frac{Q_2''}{2\sigma^4 S} + \frac{F_2'''}{6\sigma^6 S}\right)^* + t_3^* \frac{F_4'''}{6\sigma^6 S},
\end{aligned}
\tag{A.36}
$$

where we use approximation that $S = S^{\mathrm{iso}}$, S is the flux in the observational frame.

All of this allows us to formulate a new version of Eq.(A.20), taking into account the centroid shift $\Delta\bar{\theta}^{\mathrm{sm}}$ to first order. This is to estimate the integral of some function $f(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}^{\mathrm{iso}})$ over $I^{\mathrm{iso}}(\boldsymbol{\theta})$, or to describe the hypothetical moments around the centroid of I^{iso}. Using a similar Taylor expansion to that performed in Eq.(A.24), the anisotropic q correction of

moments, Eq.(A.20) becomes

$$
\int \mathrm{d}^2\theta \; f(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}^{\mathrm{iso}}) \; I^{\mathrm{iso}}(\boldsymbol{\theta}) \;\; \simeq \;\; \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \; f(\boldsymbol{\theta}) \tag{A.37}
$$
$$
+ \;\; \frac{\Delta\bar{\theta}^{\mathrm{sm}}}{2} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \; \partial^* f(\boldsymbol{\theta}) + \frac{(\Delta\bar{\theta}^{\mathrm{sm}})^*}{2} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \; \partial f(\boldsymbol{\theta})
$$
$$
- \;\; \frac{1}{2} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[d\partial^* + d^*\partial \right] f(\boldsymbol{\theta})
$$
$$
- \;\; \frac{1}{8} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[q_2\partial^*\partial^* + q_2^*\partial\partial \right] f(\boldsymbol{\theta})
$$
$$
- \;\; \frac{3}{48} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_1\partial^*\partial^*\partial + t_1^*\partial\partial\partial^* \right] f(\boldsymbol{\theta})
$$
$$
- \;\; \frac{1}{48} \int \mathrm{d}^2\theta \; I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_3\partial^*\partial^*\partial^* + t_3^*\partial\partial\partial \right] f(\boldsymbol{\theta}),
$$

where $\Delta\bar{\theta}^{\mathrm{sm}}$ is calculated in advance using Eq.(A.36). This now gives us the relation we need to correctly go from any observed moment or function to isotropic smeared ones. As an example, for T_1 (for which $f(\boldsymbol{\theta}) = \theta_1^3 W(\theta_0^2/\sigma^2)$), we have

$$
T_1^{\mathrm{iso}} \;\; \simeq \;\; T_1^{\mathrm{obs}} + \frac{\Delta\bar{\theta}^{\mathrm{sm}}}{2} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta})\partial^* \left[\theta_1^3 W \right]
$$
$$
+ \frac{(\Delta\bar{\theta}^{\mathrm{sm}})^*}{2} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta})\partial \left[\theta_1^3 W \right]
$$
$$
- \frac{1}{2} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[d\partial^* + d^*\partial \right] \left[\theta_1^3 W \right]
$$
$$
- \frac{1}{8} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[q_2\partial^*\partial^* + q_2^*\partial\partial \right] \left[\theta_1^3 W \right]
$$
$$
- \frac{3}{48} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_1\partial^*\partial^*\partial + t_1^*\partial\partial\partial^* \right] \left[\theta_1^3 W \right]
$$
$$
- \frac{1}{48} \int \mathrm{d}^2\theta I^{\mathrm{obs}}(\boldsymbol{\theta}) \left[t_3\partial^*\partial^*\partial^* + t_3^*\partial\partial\partial \right] \left[\theta_1^3 W \right].
$$

Working all this through is similar as HOLICs calculated in OUF08. By taking the origin of the moment as $D = 0$ and $D^* = 0$, and also for $d, d^* = 0$, we get the hypothetical brightness moments that smeared by the isotropic PSF,

$$
Q_0^{\mathrm{iso}} \;\; = \;\; \int \mathrm{d}^2\theta \; I^{\mathrm{iso}}(\boldsymbol{\theta})\theta_0^2 \; W(\theta_0^2/\sigma^2)
$$
$$
\simeq \;\; Q_0^{\mathrm{obs}} + (\Delta\bar{\theta}^{\mathrm{sm}})\frac{T_1'^*}{\sigma^2} + (\Delta\bar{\theta}^{\mathrm{sm}})^*\frac{T_1'}{\sigma^2}
$$
$$
- q_2 \left(\frac{Q_2'}{\sigma^2} + \frac{F_2''}{2\sigma^4} \right)^* - q_2^* \left(\frac{Q_2'}{\sigma^2} + \frac{F_2''}{2\sigma^4} \right)
$$
$$
- t_1 \left(\frac{2D'}{\sigma^2} + \frac{5T_1''}{2\sigma^4} + \frac{V_1'''}{2\sigma^6} \right)^* - t_1^* \left(\frac{2D'}{\sigma^2} + \frac{5T_1''}{2\sigma^4} + \frac{V_1'''}{2\sigma^6} \right)
$$
$$
- t_3 \left(\frac{T_3''}{2\sigma^4} + \frac{V_3'''}{6\sigma^6} \right)^* - t_3^* \left(\frac{T_3''}{2\sigma^4} + \frac{V_3'''}{6\sigma^6} \right), \tag{A.38}
$$

$$
Q_2^{\text{iso}} = \int \mathrm{d}^2\theta \, I^{\text{iso}}(\boldsymbol{\theta})\theta_2^2 \, W(\theta_0^2/\sigma^2)
$$

$$
\simeq Q_2^{\text{obs}} + (\Delta\bar{\theta}^{\text{sm}})\frac{T_1'}{\sigma^2} + (\Delta\bar{\theta}^{\text{sm}})^*\frac{T_3'}{\sigma^2}
$$

$$
-q_2\left(S + \frac{2Q_0'}{\sigma^2} + \frac{F_0''}{2\sigma^4}\right) - q_2^*\frac{F_4''}{2\sigma^4}
$$

$$
-t_1\left(\frac{3D'}{\sigma^2} + \frac{3T_1''}{\sigma^4} + \frac{V_1'''}{2\sigma^6}\right) - t_1^*\left(\frac{2T_3''}{\sigma^4} + \frac{V_3'''}{2\sigma^6}\right)
$$

$$
-t_3\left(\frac{D'}{\sigma^2} + \frac{T_1''}{\sigma^4} + \frac{V_1'''}{6\sigma^6}\right)^* - t_3^*\frac{V_5'''}{6\sigma^6}, \tag{A.39}
$$

$$
T_1^{\text{iso}} = \int \mathrm{d}^2\theta \, I^{\text{iso}}(\boldsymbol{\theta})\theta_1^3 \, W(\theta_0^2/\sigma^2) \equiv T_1^{\text{obs}} + P_{t1}(q, Q^{\text{obs}}) \tag{A.40}
$$

$$
\simeq T_1^{\text{obs}} + (\Delta\bar{\theta}^{\text{sm}})\left(2Q_0 + \frac{F_0'}{\sigma^2}\right) + (\Delta\bar{\theta}^{\text{sm}})^*\left(Q_2 + \frac{F_2'}{\sigma^2}\right)
$$

$$
-q_2\left(\frac{2T_1'}{\sigma^2} + \frac{V_1''}{2\sigma^4}\right)^* - q_2^*\left(\frac{T_3'}{\sigma^2} + \frac{V_3''}{2\sigma^4}\right)
$$

$$
-t_1\left(S + \frac{5Q_0'}{\sigma^2} + \frac{7F_0''}{2\sigma^4} + \frac{H_0'''}{2\sigma^6}\right) - t_1^*\left(\frac{3Q_2'}{\sigma^2} + \frac{3F_2''}{\sigma^4} + \frac{H_2'''}{2\sigma^6}\right)
$$

$$
-t_3\left(\frac{Q_2'}{\sigma^2} + \frac{F_2''}{\sigma^4} + \frac{H_2'''}{6\sigma^6}\right)^* - t_3^*\left(\frac{F_4''}{2\sigma^4} + \frac{H_4'''}{6\sigma^6}\right),
$$

$$
T_3^{\text{iso}} = \int \mathrm{d}^2\theta \, I^{\text{iso}}(\theta)\theta_3^3 \, W(\theta_0^2/\sigma^2) \equiv T_3^{\text{obs}} + P_{t3}(q, Q^{\text{obs}}) \tag{A.41}
$$

$$
\simeq T_3^{\text{obs}} + (\Delta\bar{\theta}^{\text{sm}})\left(3Q_2 + \frac{F_2'}{\sigma^2}\right) + (\Delta\bar{\theta}^{\text{sm}})^*\frac{F_4'}{\sigma^2}
$$

$$
-q_2\left(\frac{3T_1'}{\sigma^2} + \frac{V_1''}{2\sigma^4}\right) - q_2^*\frac{V_5''}{2\sigma^4}
$$

$$
-t_1\left(\frac{6Q_2'}{\sigma^2} + \frac{4F_2''}{\sigma^4} + \frac{H_2'''}{2\sigma^6}\right) - t_1^*\left(\frac{5F_4''}{2\sigma^4} + \frac{H_4'''}{2\sigma^6}\right)
$$

$$
-t_3\left(S + \frac{3Q_0'}{\sigma^2} + \frac{3F_0''}{2\sigma^4} + \frac{H_0'''}{6\sigma^6}\right) - t_3^*\frac{H_6'''}{6\sigma^6},
$$

$$
F_0^{\text{iso}} = \int \mathrm{d}^2\theta \, I^{\text{iso}}(\theta)\theta_0^4 \, W(\theta_0^2/\sigma^2)
$$

$$
\simeq F_0^{\text{obs}} + (\Delta\bar{\theta}^{\text{sm}})\left(2T_1 + \frac{V_1'}{\sigma^2}\right)^* + (\Delta\bar{\theta}^{\text{sm}})^*\left(2T_1 + \frac{V_1'}{\sigma^2}\right)
$$

$$
-q_2\left(Q_2 + \frac{2F_2'}{\sigma^2} + \frac{H_2''}{2\sigma^4}\right)^* - q_2^*\left(Q_2 + \frac{2F_2'}{\sigma^2} + \frac{H_2''}{2\sigma^4}\right)
$$

$$
-t_1\left(\frac{7T_1'}{\sigma^2} + \frac{4V_1''}{\sigma^4} + \frac{Z_1'''}{2\sigma^6}\right)^* - t_1^*\left(\frac{7T_1'}{\sigma^2} + \frac{4V_1''}{\sigma^4} + \frac{Z_1'''}{2\sigma^6}\right)
$$

$$
-t_3\left(\frac{T_3'}{\sigma^2} + \frac{V_3''}{\sigma^4} + \frac{Z_3'''}{6\sigma^6}\right)^* - t_3^*\left(\frac{T_3'}{\sigma^2} + \frac{V_3''}{\sigma^4} + \frac{Z_3'''}{6\sigma^6}\right). \tag{A.42}
$$

In Eqs.(A.40) and (A.41), we write the corrections as $P_{t1}(q, Q^{\text{obs}})$ and $P_{t3}(q, Q^{\text{obs}})$.

A.2.2 From I^{iso} quantities to $\hat{I}^0$ quantities - the isotropic correction

First of all, we use the reduced lens equation (3.6)

$$l(\theta) \equiv \theta - g\theta^* - \Psi_3\theta^*\theta^* - 2\Psi_1\theta^*\theta - \Psi_1^*\theta\theta, \tag{A.43}$$

where for later convenience I have defined the "lens function" $l(\boldsymbol{\theta}) = \boldsymbol{\beta}$, and Ψ_1, Ψ_3 are define in Eqs.(3.7) and (3.8). We also need the Jacobian determinant which encompasses the higher-order lens equation (A.43), which is already defined in Chapter 3,

$$\det\mathcal{A} = 1 - g^*g - \eta_1\theta^* - \eta_1^*\theta, \tag{A.44}$$

where $\eta_1 = G_1 + (g^*G_3 - gG_1^*)/2$. We then follow the logical steps of BS01 closely as before, but expand our treatment as necessary.

Since $I(\boldsymbol{\theta}) = I^{\mathrm{s}}(l(\boldsymbol{\theta}))$, we may write equation (A.16) as

$$I^{\mathrm{iso}}(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi I^{\mathrm{s}}(l(\boldsymbol{\varphi}))P^{\mathrm{iso}}(\boldsymbol{\theta} - \boldsymbol{\varphi}). \tag{A.45}$$

As in BS01, we may then transform the integration variable to $\boldsymbol{\beta} = l(\boldsymbol{\varphi})$. Now that we are considering flexion, however, we need to be especially careful as equation A.43 does not give single, unique solutions for $\boldsymbol{\theta}$ and therefore cannot be uniquely inverted. As discussed in Chapter 3, we must limit ourselves to the regime in which the flexion and source size are sufficiently small to produce a single, primary image close to the origin. We may then define an inverted lensing function $l^{-1}(\boldsymbol{\beta})$ corresponding to this observable image, such that $\boldsymbol{\theta} = l(l^{-1}(\boldsymbol{\theta}))$. This can be found by solving equation (A.43) for θ and taking only the solution lying closest to the origin. We then allow ourselves to write $\boldsymbol{\beta} = l(\boldsymbol{\varphi})$ and $\boldsymbol{\varphi} = l^{-1}(\boldsymbol{\beta})$, giving

$$I^{\mathrm{iso}}(\boldsymbol{\theta}) = \int \frac{\mathrm{d}^2\beta}{\det\mathcal{A}} I^{\mathrm{s}}(\boldsymbol{\beta})P^{\mathrm{iso}}(\boldsymbol{\theta} - l^{-1}(\boldsymbol{\beta})). \tag{A.46}$$

We now define the surface brightness $\hat{I}$ as

$$\hat{I}(\boldsymbol{\theta}) = \int \mathrm{d}^2\varphi \ I^{\mathrm{s}}(\boldsymbol{\varphi})\hat{P}(\boldsymbol{\theta} - \boldsymbol{\varphi}), \tag{A.47}$$

with

$$\hat{P}(\boldsymbol{\theta}) \equiv \frac{1}{\det\mathcal{A}}P^{\mathrm{iso}}(\boldsymbol{l}^{-1}(\boldsymbol{\theta})), \tag{A.48}$$

which may be thought of as a fictional PSF relating $\hat{I}$ to I^{s}, and which due to shear and flexion is not purely isotropic in general. Putting all this together gives

$$I^{\mathrm{iso}}(\boldsymbol{\theta}) = \int \mathrm{d}^2\beta I^{\mathrm{s}}(\beta)\frac{1}{\det\mathcal{A}}P^{\mathrm{iso}}(l^{-1}(l(\theta) - \beta)) = \hat{I}(l(\boldsymbol{\theta})). \tag{A.49}$$

This is completely equivalent to equation (4.73) in BS01, except that now we have to define the inverse of the lens equation more carefully and ensure we are operating in the "non-breakdown" zone of flexion. This will then allow us to calculate moments on the $\hat{I}$ image and relate them to those we already have for the I^{iso} image. The final step will then be to relate these to moments from the $\hat{I}^0$ image and that is the same work in the Sect.A2.1.

All this will be similar to BS01 but we need to clearly understand the effect of the centroid shift. This will begin to come in for our version of equation (4.75) of BS01 (and its equivalents for the flexion moments).

Having all these preparation, we can make the isotropic correction. Using Eq.(A.49) in the source plane brightness moments integration,

$$\int \mathrm{d}^2\beta \; \beta_M^N \; \hat{I}(\beta) W(\beta_0^2/\sigma_s^2) = \int \mathrm{d}^2\theta \; \det\!\mathcal{A} \, l(\theta)_M^N \, I^{\mathrm{iso}}(\theta) W(l(\theta)_0^2/\sigma_s^2) \tag{A.50}$$

with the length-scale of the weighting

$$\sigma_s^2 = \det\!\mathcal{A} \, \sigma^2 \simeq (1-\kappa)^2(1-gg^*)\sigma^2. \tag{A.51}$$

Using the Taylor expansion to the weighting function as

$$W(\beta_0^2 + \delta) = W(\beta_0^2) + W'\delta, \tag{A.52}$$

and the linear approximation of $\det\!\mathcal{A}$, we can find out the relation between the brightness moments of I^{iso} and $\hat{I}$. As an example, for the 0-order moment, which is the total flux of the image

$$\hat{S}_0 = \int \mathrm{d}^2\beta \; \hat{I}(\beta) W(\beta_0^2) = \int \mathrm{d}^2\theta \; \det\!\mathcal{A} \, I^{\mathrm{iso}}(\theta) W(l(\theta)_0^2) = (1-gg^*)\left(S^{\mathrm{iso}} - \frac{gQ_2'^* g^* Q_2}{\sigma_s^2} \right). \tag{A.53}$$

However, here we still need to care about the centroid shift. We define the centroid shift $\hat{\bar{\beta}}$ in the source plane by the following, which is due to the isotropic PSF and lensing,

$$\int \mathrm{d}^2\beta (\beta - \hat{\bar{\beta}}) \hat{I}(\beta) \, W[(\beta - \hat{\bar{\beta}})_0^2/\sigma_s^2] = 0. \tag{A.54}$$

Assuming that $\hat{\bar{\beta}}$ is a small constant and using Eq.(A.52), we can rewrite Eq.(A.54) as

$$M[\beta] - \hat{\bar{\beta}} M'[\beta_0^2] - \hat{\bar{\beta}}^* M'[\beta_2^2] - \hat{\bar{\beta}}\hat{S}_0 + \hat{\bar{\beta}}_2^2 M'[\beta^*] + \hat{\bar{\beta}}_0^2 M'[\beta] = 0, \tag{A.55}$$

with which we can solve for $\hat{\bar{\beta}}$ numerically, and here $M[\beta]$ is stand for the brightness moment

$$\int \mathrm{d}^2\beta \; \beta \hat{I}(\beta) \, W[\beta] \simeq -\frac{5}{4}G_1^* Q_2 - \frac{3}{2}G_1 Q_0 - \frac{1}{4}G_3 Q_2^* + \frac{5}{4}g G_1^* F_0' + \frac{1}{4}g^* G_3 F_0', \tag{A.56}$$

where there are prime on M, the integrals are required to be taken over W'. After that we can define the source brightness moment with respect to the correct center

$$\int \mathrm{d}^2\beta \, f(\beta - \hat{\bar{\beta}}) \hat{I}(\beta) W[(\beta - \hat{\bar{\beta}})_0^2/\sigma_s^2] \tag{A.57}$$

$$\simeq \int \mathrm{d}^2\beta \, f(\beta - \hat{\bar{\beta}}) \hat{I}(\beta) \left[W(\beta_0^2/\sigma_s^2) - \frac{\beta^* \hat{\bar{\beta}} + \beta \hat{\bar{\beta}}^*}{\sigma_s^2} W' \right]$$

$$\simeq \int \mathrm{d}^2\theta \, \det\!\mathcal{A} \, f(l(\theta) - \hat{\bar{\beta}}) I^{\mathrm{iso}}(\theta) \left[W(\theta_0^2/\sigma_s^2) - \frac{g\theta_{-2}^2 + g^*\theta_2^2}{\sigma_s^2} W' - \frac{l(\theta)^* \hat{\bar{\beta}} + l(\theta)\hat{\bar{\beta}}^*}{\sigma_s^2} W' \right].$$

This result can be applied to the second-order moments by setting $f(\beta - \hat{\bar{\beta}}) = (\beta - \hat{\bar{\beta}})_0^2$ and $f(\beta - \hat{\bar{\beta}}) = (\beta - \hat{\bar{\beta}})_2^2$, then we obtain the correction for the isotropic PSF

$$
\begin{aligned}
\hat{Q}_0 &\simeq Q_0^{\text{iso}} \\
\hat{Q}_2 &\simeq Q_2^{\text{iso}} - 2gQ_0^{\text{iso}} + g^2 Q_2^{*\text{iso}} \\
&+ (-\hat{\bar{\beta}}T_1' - \hat{\bar{\beta}}^* T_3' - 2F_0' g - 2F_4' g^* + 4F_2'^* g^2 + 4gg^* F_2' - 2F_4'^* g^3 - 2F_0' g^2 g^*)/\sigma_s^2
\end{aligned}
\tag{A.58}
$$

The shear estimator with higher-order KSB can be obtain from these equations. But here we are more interested in the flexion, thus we set that $f(\beta - \hat{\bar{\beta}}) = (\beta - \hat{\bar{\beta}})_1^3$ and $f(\beta - \hat{\bar{\beta}}) = (\beta - \hat{\bar{\beta}})_3^3$ in Eq.(A.57). These equations give the relation for the third-order moments

$$
\begin{aligned}
\hat{T}_1 &\simeq T_1^{\text{iso}} - \frac{9F_0 - 12Q_0^2}{4}G_1 - 2\hat{\bar{\beta}}Q_0 - \hat{\bar{\beta}}^* Q_2 - \frac{\hat{\bar{\beta}}F_0' + \hat{\bar{\beta}}^* F_2'}{\sigma_s^2} \\
&+ 9\frac{(gH_2'^* + g^* H_2')}{2\sigma_s^2}G_1 + 4\frac{(gH_0' + g^* H_4')}{\sigma_s^2}G_1^* + \frac{(gH_2' + g^* H_6')}{2\sigma_s^2}G_3^* + \frac{(g^* H_0' + gH_4'^*)}{\sigma_s^2}G_3; \\
\hat{T}_3 &\simeq T_3^{\text{iso}} - \frac{3F_0}{4}G_3 - 3\hat{\bar{\beta}}Q_2 - \frac{\hat{\bar{\beta}}F_2' + \hat{\bar{\beta}}^* F_4'}{\sigma_s^2} \\
&+ 3\frac{(gH_2'^* + g^* H_2')}{\sigma_s^2}G_3 + 5\frac{(gH_0' + g^* H_4')}{\sigma_s^2}G_1,
\end{aligned}
\tag{A.59}
$$

where we only keep the linear terms of $\hat{\bar{\beta}}$. In these relations we can see that the first two terms on the right sides of the equations are the same as we have in Chap. 3 (Eq.3.64). The rest terms are correction for the centroid shift due to isotropic PSF and weight function. If we use the Gaussian weighting function, the derivative has the similar form as the original one, we omit the prime on the moment. We make some further simplification, only the terms which directly relate to G_1/G_3 are kept,

$$
\begin{aligned}
\hat{T}_1 &\simeq T_1^{\text{iso}} - \left(\frac{9F_0 - 12Q_0^2}{4} - \frac{9}{2}(gH_2^* + g^* H_2)/\sigma_s^2\right)G_1 \\
\hat{T}_3 &\simeq T_3^{\text{iso}} - \left(\frac{3F_0}{4} - 3(gH_2^* + g^* H_2)/\sigma_s^2\right)G_3.
\end{aligned}
\tag{A.60}
$$

In the end, another pre-lensing isotropic smear correction is needed. In analogy to Sect. A.2.1, we decompose $\hat{P}$ into an isotropic and an anisotropic part, the latter one assumed to be small

$$
\hat{P}(\theta) = \int \mathrm{d}^2\varphi \hat{P}^{\text{iso}}(\varphi)\hat{q}(\theta - \varphi).
\tag{A.61}
$$

Following the same way as that between I^{obs} and I^{iso}, we relate the brightness profile which would be obtained from smearing the source with the isotropic PSF P^{iso} and $\hat{I}$ through

$$
\hat{I}(\theta) = \int \mathrm{d}^2\varphi \hat{I}^0(\varphi)\hat{q}(\theta - \varphi).
\tag{A.62}
$$

Then we can write the last step of higher-KSB

$$
\begin{aligned}
\hat{T}_1^0 &= \hat{T}_1 + \hat{P}_{t1}(\hat{q}, Q'); \\
\hat{T}_3^0 &= \hat{T}_3 + \hat{P}_{t3}(\hat{q}, Q'),
\end{aligned}
\tag{A.63}
$$

where $\hat{P}_{ti}(\hat{q}, Q')$, $i = 1, 3$ stand for anisotropic correction which should in principle be calculated using $\hat{I}$ instead of I^{obs} in Eqs.(A.40) and (A.41). Here we however assume the differences between $\hat{I}$, I^{iso} and I^{obs}, and use q as $\hat{q}$ in Eq.(A.63).

Combining all the 3 step results together, we obtain the final correction of higher-order KSB

$$
\begin{aligned}
\hat{T}_1^0 &= T_1^{\text{obs}} + P_{t1}(q, Q) + \hat{P}_{t1}(\hat{q}, Q') - G_1 \left(\frac{9F_0 - 12Q_0^2}{4} - \frac{9}{2}(gH_2^* + g^*H_2)/\sigma_s^2 \right) \\
\hat{T}_3^0 &= T_3^{\text{obs}} + P_{t3}(q, Q) + \hat{P}_{t3}(\hat{q}, Q') - G_3 \left(\frac{3F_0}{4} - 3(gH_2^* + g^*H_2)/\sigma_s^2 \right).
\end{aligned}
\tag{A.64}
$$

This equations relate the observed brightness moments to that of the source smeared by an isotropic PSF, using the anisotropy PSF and the reduced shear and reduced flexion. Since the expectation value of T_1^0 and T_3^0 are zero, these equations yield the estimator for the reduced flexion. The result shown here are preliminary. The accuracy and behavior of these estimators need further test.

Appendix B

The matrix C

In this Appendix, we list the coefficients of the matrix $\boldsymbol{C}$ which occurs in (3.56):

$$
\begin{aligned}
4(1-gg^*)C_{11} &= -2gF_2^* + (9gg^* - 3)F_0 + 6g^*(1 - 2gg^*)F_2 + g^{*2}(5gg^* - 3)F_4 \\
&+ 6gQ_2^*Q_0 - 12gg^*Q_0^2 + (3 - 9gg^*)Q_2^*Q_2 + 6g^*(4gg^* - 1)Q_0Q_2 + 3g^{*2}(1 - 3gg^*)Q_2^2 \\
4(1-gg^*)C_{12} &= 5gF_4^* - 2(5 + 6gg^*)F_2^* + 9g^*(3 + gg^*)F_0 - 2g^{*2}(12 + gg^*)F_2 + 7g^{*3}F_4 \\
&- 9gQ_2^{*2} + 6(3 + 4gg^*)Q_2^*Q_0 - 12g^*(3 + gg^*)Q_0^2 - 3g^*(5 + 3gg^*)Q_2^*Q_2 \\
&+ 6g^{*2}(8 + gg^*)Q_0Q_2 - 15g^{*3}Q_2^2 \\
4(1-gg^*)C_{13} &= -7F_4^* + 26g^*F_2^* - 36g^{*2}F_0 + 22g^{*3}F_2 - 5g^{*4}F_4 \\
&+ 15Q_2^{*2} - 54g^*Q_2^*Q_0 + 48g^{*2}Q_0^2 + 24g^{*2}Q_2^*Q_2 - 42g^{*3}Q_0Q_2 + 9g^{*4}Q_2^2 \\
4(1-gg^*)C_{14} &= -2g^*F_4^* + 6g^{*2}F_2^* - 6g^{*3}F_0 + 2g^{*4}F_2 \\
&+ 6g^*Q_2^{*2} - 18g^{*2}Q_2^*Q_0 + 12g^{*3}Q_0^2 + 6g^{*3}Q_2^*Q_2 - 6g^{*4}Q_0Q_2 \\
4(1-gg^*)C_{21} &= 2g^2F_2^* - 6g^2g^*F_0 + [4gg^*(1 + gg^*) - 2]F_2 + 2g^*(1 - 2gg^*)F_4 \\
&- 6g^2Q_2^*Q_0 + 4g(1 + 2gg^*)Q_0^2 + 6g^2g^*Q_2^*Q_2 + [2 - 4gg^*(3 + 2gg^*)]Q_0Q_2 \\
&+ 2g^*(4gg^* - 1)Q_2^2 \\
4(1-gg^*)C_{22} &= -5g^2F_4^* + 2g(7 + 4gg^*)F_2^* - 3[3 + gg^*(8 + gg^*)]F_0 + 2g^*(8 + 5gg^*)F_2 - 7g^{*2}F_4 \\
&+ 9g^2Q_2^{*2} - 2g(13 + 8gg^*)Q_2^*Q_0 + 4[3 + gg^*(8 + gg^*)]Q_0^2 \\
&+ [5 + gg^*(16 + 3gg^*)]Q_2^*Q_2 - 2g^*(16 + 11gg^*)Q_0Q_2 + 15g^{*2}Q_2^2 \\
4(1-gg^*)C_{23} &= 7gF_4^* - 2(4 + 9gg^*)F_2^* + 3g^*(7 + 5gg^*)F_0 - 2g^{*2}(9 + 2gg^*)F_2 + 5g^{*3}F_4 \\
&- 15gQ_2^{*2} + (16 + 38gg^*)Q_2^*Q_0 - 4g^*(7 + 5gg^*)Q_0^2 - g^*(13 + 11gg^*)Q_2^*Q_2 \\
&+ 2g^{*2}(17 + 4gg^*)Q_0Q_2 - 9g^{*3}Q_2^2 \\
4(1-gg^*)C_{24} &= (3gg^* - 1)F_4^* - 6gg^{*2}F_2^* + 3g^{*2}(1 + gg^*)F_0 - 2g^{*3}F_2 \\
&+ (1 - 7gg^*)Q_2^{*2} + 2g^*(2 + 7gg^*)Q_2^*Q_0 - 4g^{*2}(2 + gg^*)Q_0^2 \\
&- 3g^{*2}(1 + gg^*)Q_2^*Q_2 + 6g^{*3}Q_0Q_2
\end{aligned}
$$

The other eight elements follow trivially from the foregoing ones, since the second half of the matrix is just the complex conjugate one of the first half, i.e., $C_{44} = C_{11}^*$, $C_{34} = C_{21}^*$ etc., or in general, $C_{ij} = C_{5-i,5-j}^*$.

Bibliography

Abramowitz, M. & Stegun, I. A. 1972, Handbook of Mathematical Functions, ed. M. Abramowitz & I. A. Stegun

Albrecht, A. & Steinhardt, P. J. 1982, Physical Review Letters, 48, 1220

Allgood, B., Flores, R. A., Primack, J. R., et al. 2006, MNRAS, 367, 1781

Bacon, D. J., Amara, A., & Read, J. I. 2009, ArXiv e-prints

Bacon, D. J., Goldberg, D. M., Rowe, B. T. P., & Taylor, A. N. 2006, MNRAS, 365, 414

Bacon, D. J., Refregier, A., Clowe, D., & Ellis, R. S. 2001, MNRAS, 325, 1065

Bacon, D. J. & Schäfer, B. M. 2009, MNRAS, 396, 2167

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Bartelmann, M., Huss, A., Colberg, J. M., Jenkins, A., & Pearce, F. R. 1998, A&A, 330, 1

Bartelmann, M. & Schneider, P. 2001, Phys. Rep., 340, 291

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, Phys. Rep., 367, 1

Bernstein, G. M. & Jarvis, M. 2002, AJ, 123, 583

Bradač, M., Erben, T., Schneider, P., et al. 2005a, A&A, 437, 49

Bradač, M., Lombardi, M., & Schneider, P. 2004, A&A, 424, 13

Bradač, M., Schneider, P., Lombardi, M., & Erben, T. 2005b, A&A, 437, 39

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Bridle, S., Balan, S. T., Bethge, M., et al. 2009, ArXiv e-prints

Bryan, G. L. & Norman, M. L. 1998, ApJ, 495, 80

Chwolson, O. 1924, Astronomische Nachrichten, 221, 329

Ciardi, B. & Ferrara, A. 2005, Space Science Reviews, 116, 625

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2002, ApJ, 568, 20

Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 667, 859

Dodelson, S. 2003, Modern cosmology, ed. S. Dodelson

Eifler, T., Schneider, P., & Krause, E. 2010, A&A, 510, A260000+

Einstein, A. 1916, Annalen der Physik, 354, 769

Eisenstein, D. J. & Hu, W. 1998, ApJ, 496, 605

Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., & Schneider, P. 2001, A&A, 366, 717

Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415

Fort, B. & Mellier, Y. 1994, A&A Rev., 5, 239

Fort, B., Prieur, J. L., Mathez, G., Mellier, Y., & Soucail, G. 1988, A&A, 200, L17

Fu, L. & Kilbinger, M. 2010, MNRAS, 401, 1264

Fu, L., Semboloni, E., Hoekstra, H., et al. 2008, A&A, 479, 9

Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, MNRAS, 355, 819

Goldberg, D. M. & Bacon, D. J. 2005, ApJ, 619, 741

Goldberg, D. M. & Leonard, A. 2007, ApJ, 660, 1003

Goldberg, D. M. & Natarajan, P. 2002, ApJ, 564, 65

Gorenstein, M. V., Shapiro, I. I., & Falco, E. E. 1988, ApJ, 327, 693

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hawken, A. J. & Bridle, S. L. 2009, MNRAS, 400, 1132

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31

Hilbert, S., White, S. D. M., Hartlap, J., & Schneider, P. 2008, MNRAS, 386, 1845

Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636

Hoekstra, H., van Waerbeke, L., Gladders, M. D., Mellier, Y., & Yee, H. K. C. 2002a, ApJ, 577, 604

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, ApJ, 606, 67

Hoekstra, H., Yee, H. K. C., Gladders, M. D., et al. 2002b, ApJ, 572, 55

Irwin, J. & Shmakova, M. 2005, New Astronomy Review, 49, 83

Irwin, J. & Shmakova, M. 2006, ApJ, 645, 17

Jing, Y. P. 2002, MNRAS, 335, L89

Jing, Y. P. & Suto, Y. 2002, ApJ, 574, 538

Kaiser, N. 1994, in Clusters of Galaxies, ed. F. Durret, A. Mazure, & J. Tran Thanh van, 269–+

Kaiser, N. 1995, ApJ, 439, L1

Kaiser, N. & Squires, G. 1993, ApJ, 404, 441

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201

Kitayama, T. & Suto, Y. 1996, MNRAS, 280, 638

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330

Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, ApJ, 609, 482

Kuijken, K. 2006, A&A, 456, 827

Leonard, A., Goldberg, D. M., Haaga, J. L., & Massey, R. 2007, ApJ, 666, 51

Leonard, A., King, L. J., & Wilkins, S. M. 2009, MNRAS, 395, 1438

Liddle, A. R. & Lyth, D. H. 2000, Cosmological Inflation and Large-Scale Structure, ed. A. R. Liddle & D. H. Lyth

Limber, D. N. 1953, ApJ, 117, 134

Linde, A. D. 1982, Physics Letters B, 108, 389

Luppino, G. A. & Kaiser, N. 1997, ApJ, 475, 20

Ma, C. & Fry, J. N. 2000, ApJ, 543, 503

Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U., & Brinkmann, J. 2006, MNRAS, 370, 1008

Massey, R., Heymans, C., Bergé, J., et al. 2007a, MNRAS, 376, 13

Massey, R. & Refregier, A. 2005, MNRAS, 363, 197

Massey, R., Rowe, B., Refregier, A., Bacon, D. J., & Bergé, J. 2007b, MNRAS, 380, 229

Monaghan, J. J. 1992, ARA&A, 30, 543

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Okura, Y., Umetsu, K., & Futamase, T. 2007, ApJ, 660, 995

Okura, Y., Umetsu, K., & Futamase, T. 2008, ApJ, 680, 1

Refregier, A. & Bacon, D. 2003, MNRAS, 338, 48

Rubakov, V. A., Sazhin, M. V., & Veryaskin, A. V. 1982, Physics Letters B, 115, 189

Sawala, T., Guo, Q., Scannapieco, C., Jenkins, A., & White, S. D. M. 2010, ArXiv e-prints

Schneider, P. 1996, MNRAS, 283, 837

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses, ed. P. Schneider, J. Ehlers, & E. E. Falco

Schneider, P. & Er, X. 2008, A&A, 485, 363

Schneider, P. & Kilbinger, M. 2007, A&A, 462, 841

Schneider, P., Kochanek, C. S., & Wambsganss, J. 2006, Gravitational Lensing: Strong, Weak and Micro, ed. P. Schneider, C. S. Kochanek, & J. Wambsganss

Schneider, P. & Seitz, C. 1995, A&A, 294, 411

Schneider, P., van Waerbeke, L., Jain, B., & Kruse, G. 1998, MNRAS, 296, 873

Schneider, P., van Waerbeke, L., Kilbinger, M., & Mellier, Y. 2002a, A&A, 396, 1

Schneider, P., van Waerbeke, L., & Mellier, Y. 2002b, A&A, 389, 729

Schwarzschild, K. 1916, Abh. Konigl. Preuss. Akad. Wissenschaften Jahre 1906,92, Berlin,1907, 189

Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20

Seitz, C. & Schneider, P. 1995, A&A, 297, 287

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Seitz, S. & Schneider, P. 2001, A&A, 374, 740

Seljak, U. 2000, MNRAS, 318, 203

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Sugiyama, N. 1995, ApJS, 100, 281

Tegmark, M., Strauss, M. A., Blanton, M. R., et al. 2004, Phys. Rev. D, 69, 103501

Zaroubi, S. & Silk, J. 2005, MNRAS, 360, L64

Zentner, A. R., Berlind, A. A., Bullock, J. S., Kravtsov, A. V., & Wechsler, R. H. 2005, ApJ, 624, 505