

# 5 *Lensing by galaxies and galaxy clusters*

---

This Chapter gives an overview on the lensing phenomena produced by extended lenses like galaxies and galaxy clusters in the universe.

The lensing events produced by such lenses fall into two broad classes. If the observer, the lens and the source happen to be well aligned along the line of sight, i.e. if the observer is looking at far sources which are projected on the sky at small angular distances from the centre of the lens, and if the combination of convergence and shear in the inner regions of the lens is appropriate, *strong lensing* events may be observed. Depending on the characteristics of the lens and of the sources, such events may consist of

- multiple images of background sources. For example, massive galaxies can split one bright QSO into several images. The displacement of such images is determined by the mass distribution of the lenses. The spectra of all these images conserve all the features contained in the spectrum of the corresponding source. Therefore, multiple images can be identified thanks to spectral analysis;
- highly distorted images. If the source is extended, the differential deflection of the light creates distortions in the images. As discussed earlier, such distortions can be in the radial or in the tangential direction. The first is particularly large around the tangential critical curves, the second close to the radial critical curves. Typical and spectacular examples of distorted images are the gravitational arcs which are observed in the core of many massive galaxy clusters;

Conversely, if the angular separation between source and lens is large, lensing shows up in the *weak* regime. Such effect is detectable only by averaging on an ensemble of extended sources (galaxies), which happen to lay behind the lens. The shape of the images of these sources may be only weakly affected by lensing. The small distortions are difficult to measure. It is impossible to do so on the images of single galaxies because the deformation is indistinguishable from the intrinsic shape of the sources. Since the distortion is coherent across a region on the sky surrounding the lens, it can be detected only averaging over several sources in a given aperture.

Both the weak and the strong lensing regimes will be discussed in the following sections.

## 5.1 Strong lensing

### 5.1.1 General considerations

Strong lensing occurs in the central regions of galaxies and galaxy clusters when the lens is “critical”. This happens when it develop extended critical lines. As seen in Chapter

2, these form where the conditions

$$\lambda_t = 1 - \kappa - \gamma = 0 \quad (5.1)$$

$$\lambda_r = 1 - \kappa + \gamma = 0. \quad (5.2)$$

The first equation defines the *tangential* critical line, the second the *radial* critical line. It is obvious that extended systems are much more complex lenses than point masses. Their strong lensing properties are determined by a much larger number of parameters than those which play a role in microlensing.

In particular

- galaxies and clusters are made of dark matter and baryons (stars, gas). These components may have different spatial distributions;
- we may think of extended lenses as composed of a **main clump of matter** and many **substructures** orbiting around it. One can imagine a galaxy or a galaxy cluster as the superposition of several surface density modes corresponding to different spatial scales. Following this idea, we can write the surface density as a multipole expansion:

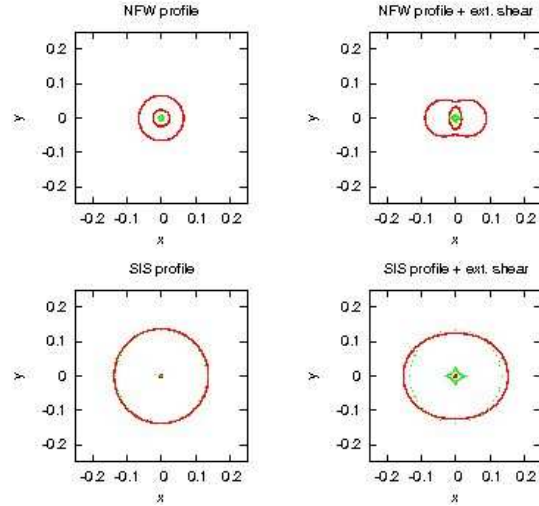
$$\kappa(\vec{x}) = \kappa_0(x) + \sum_{m=1}^{\infty} \kappa_m(x) \exp(im\phi); \quad (5.3)$$

- in this expansion the term  $\kappa_0$  is the monopole and it describes the axially symmetric part of the lens; the term with  $m = 1$  is the dipole; the term with  $m = 2$  is the quadrupole, which gives the degree of **ellipticity** of the iso-density contours. If a “power spectrum” is calculated with the coefficients  $\kappa_m$ , these result to be the most important terms;
- the **density profile**  $\kappa(x)$  [ $\rho(r)$ ] is a fundamental property of the lens. In Chapter 2, the differences between the lensing properties of the NFW and SIS profiles were highlighted: different profiles produce remarkably different lensing features, for example different multiplicities of the images, different strengths of the shear field and of the magnification pattern on the source and on the lens planes;
- depending on the shape of its density profile, a lens can be more or less sensitive to **external perturbations**. For example, an external shear  $\vec{\gamma}_e$ , which corresponds to a lensing potential

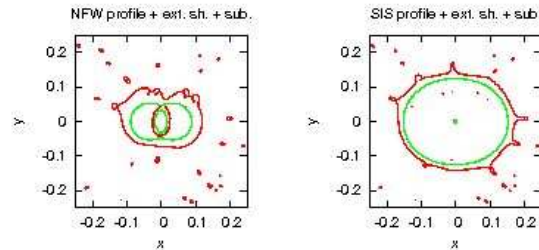
$$\Psi_\gamma(\vec{x}) = \frac{\gamma_e}{2}(x_1^2 - x_2^2), \quad (5.4)$$

perturbs much more easily an NFW than a Singular-Isothermal sphere. The reason is that the SIS density profile is significantly steeper than the NFW in the central part. The situation is illustrated in Fig. 5.1. In the left panels, the critical lines and caustics of two axially symmetric lenses with the same mass ( $M = 10^{14} M_\odot$ ) are shown. The lens and the source redshifts are assumed to be 0.3 and 1.0, respectively. An external shear, whose amplitude is  $\gamma_e = 0.1$ , is applied to both the lenses. The corresponding critical lines and caustics are shown in the two right panels. Clearly, the deformation of the critical lines (and caustics) is more significant for the NFW than for the SIS profile;

- the **impact of the substructures**, described by the higher order multipoles in Eq. 5.3, is therefore much larger in halos whose density profiles are shallower. The same two lenses which were earlier embedded into an external shear are now

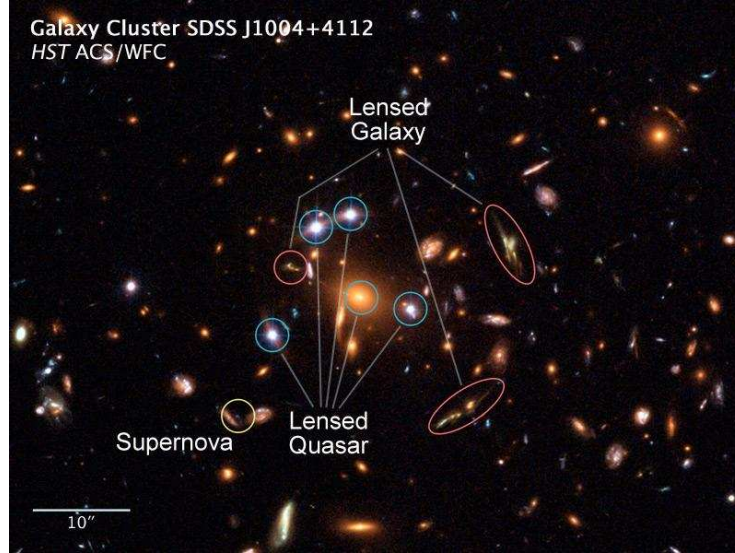


**Figure 5.1:** The figure illustrate how the shape of the critical lines (red) and of the caustics (green) change when the lens is embedded into an external shear. The upper and the bottom panels refer to an NFW and to a SIS lens model, respectively.



**Figure 5.2:** Different sensitivity to substructures of lenses with NFW (left) and SIS (right) density profiles. The red curves show the critical lines after populating the halo of the lens with 30 subhalos with masses  $10^{10} < M/M_{\odot} < 10^{11}$ . The green curves show the corresponding critical lines in absence of the subhalos.

shown in Fig. 5.2 after having been populated with 30 axially symmetric subhalos. Each subhalo has mass between  $10^{10}$  and  $10^{11} M_{\odot}$  and has been modelled as a SIS. One could think that they represent galaxies in the halo of a galaxy cluster. The same subhalos have been used to populate both the lenses. Two important features can be noticed. First, for both the NFW and SIS profiles, the critical lines of the main clump are expanded. The relative expansion of the critical line of the NFW lens is significantly larger than that of the SIS. Substructures typically enhance the strong lensing ability of the lens, because they provide additional convergence and shear. Second, the critical lines around the individual subhalos have different sizes, depending on density profile of the halos within which they are embedded. In other words: substructures are much stronger lenses when the structure they belong to has a shallow density profile;



**Figure 5.3:** The galaxy cluster SDSS J1004+4112 has 5 images of a QSO close to its centre. At the same time, several other background galaxies are strongly lensed by this cluster.

### 5.1.2 Observables

What can we learn from images like that in Fig. 5.3? In fact, we can extract a huge amount of information from this kind of observations. Indeed, lensing is a unique tool for tracing the total mass distribution of the lens and at the same time it magnifies distant sources, allowing to see objects which otherwise would be very faint. Finally, since lensing is a geometrical effect, for lensing systems at cosmological distances some important information about the geometry of the universe can be derived.

It is then crucial to use in the proper way the observational constraints given a lens system.

Galaxies and galaxy clusters can strongly lens point or extended sources. The angular separation between multiple images is of the order of the Einstein ring,

$$\begin{aligned}\theta_E &= 0.9'' \left( \frac{M}{10^{11} M_\odot} \right)^{1/2} \left( \frac{D}{\text{Gpc}} \right)^{1/2} \\ &= 1.5' \left( \frac{M}{10^{15} M_\odot} \right)^{1/2} \left( \frac{D}{\text{Gpc}} \right)^{1/2},\end{aligned}\quad (5.5)$$

which means that the multiple images can be well resolved.

Among the things that can be measured for a lens are

- (1) the relative positions of the components (lens and images; astrometric constraints);
- (2) the relative fluxes of the images;
- (3) the time delays between the images;
- (4) several properties of the lens (dynamical properties, light distribution in different bands, etc.);

(5) the microlensing of the images.

The astrometric constraints are the most important. We can usually measure the relative positions of the lensed components very accurately (5 mas or better). Substructures and other perturbers set a lower limit of order 1-5 mas with which it is safe to impose astrometric constraints. A given lens-source model provides a configuration of the images which can be compared with the observations. If the images are extended, additional constraints derive from the fact that we can measure the relative transformation between one image and the other.

Flux ratios are relatively easy to measure, but not very useful because of the systematic uncertainties (see later in the text). As we will see, flux ratios are predictable. However, if taken at a single epoch they are affected by time variability in the source (which appears with some time delay in the different images), microlensing (in galaxies) or lensing by subhalos and substructures (in galaxies and clusters), absorption by ISM (galaxies) and IGM (clusters). In fact, most applications of flux ratios have focused on probing these perturbers rather than on studying the mass distribution of the lenses.

Measurements of time delays are possible for sources which are lensed by galaxies. Indeed, the angular separations between multiple images in galaxy clusters are too large and the time delays too long. For example, from Eqs. 2.64 and 3.48, the time delay between the two images produced by a SIS is

$$\Delta t_{SIS} = \frac{1}{2} \frac{D_L D_S}{c D_{LS}} (1 + z_L) (\theta_A^2 - \theta_B^2). \quad (5.6)$$

Such delay amounts to some months in the case of galaxies ( $\theta_E \sim 1''$ ) and to some decades in the case of galaxy clusters. To date, time delays have primarily been used to estimate the Hubble constant rather than the surface density, but if we assume to know  $H_0$  or consider only time delays ratios, then time delays can be used to constrain the mass distribution.

Any independent measurement of the mass of a component will also help to constrain the structure of the lens. For galaxies, this primarily means making stellar dynamical measurements of the lens galaxy and comparing the dynamical mass estimates to those from the lens geometry. For clusters, the dynamics of cluster galaxies, the X-ray emission from the hot intra-cluster gas, or weak lensing (see later) can be used to estimate the cluster mass. Unfortunately, these independent mass estimates are frequently in disagreement with those derived from strong lensing.

Using microlensing variability to constrain the mass distribution around where images appear is actually more theory than practice due to the lack of microlensing light curves for almost all lenses.

### 5.1.3 Mass modelling

One of the most important applications of strong lensing is that of using the observables listed above to constrain the mass distribution of the lens in its inner region.

The mass modelization of the lens is never an easy task. In particular, the modeler has to deal with huge degeneracies in the parameter space which make several models able to fit the observed data.

There exist two different approaches to this problem. The first consists in using parametric fitting models to recover the mass distribution of the lens. The monopole term is modelled using some analytic profile like those discussed in Chapter 3 (SIS, NFW, SIE, etc.). The quadrupole is added by introducing ellipticity into the model. Higher order terms in the multipole expansion of the surface density are generated by adding additional mass components to the lens. The effect of the environment is mimicked by

adding external shear. Each of the ingredients which are used to construct the model is described by some parameters. The list  $\mathbf{p}$  of all these parameters defines the lens model.

Suppose we use only the astrometric constraints. The lens equation supplies the source position for each combination of the model parameters:

$$\vec{\beta}_i = \vec{\theta}_i - \vec{\alpha}(\vec{\theta}_i, \mathbf{p}) \quad (5.7)$$

It is easy to project the images onto the source plane and then minimize the difference between the projected source positions. This can be done with a  $\chi^2$  fit statistic of the form

$$\chi_{src}^2 = \sum_i \left( \frac{\vec{\beta} - \vec{\beta}_i}{\sigma_i} \right)^2. \quad (5.8)$$

The position of the source  $\vec{\beta}$  is unknown, therefore it is a model parameter. By minimizing the  $\chi^2$ , we look for the minimal scatter between the source points leading to the observed images.

The advantage of using  $\chi_{src}^2$  is that mapping the observed positions of the images on the source plane is straightforward and fast. In the previous equation the  $\sigma_i$  are the astrometric uncertainties. These are known on the lens plane, not on the source plane. Therefore, this method is conceptually wrong.

Alternatively, one can define a  $\chi^2$  fit statistic on the lens plane. However, this implies that the lens equation has to be solved in the opposite direction, which is numerically much more expensive. The  $\chi^2$  is then

$$\chi_{img}^2 = \sum_i \left( \frac{\vec{\theta}_i(\vec{\beta}) - \vec{\theta}_i}{\sigma_i} \right)^2. \quad (5.9)$$

Additional  $\chi^2$  variables can be constructed to fit the other observables, but, as we said in the previous section, they provide much less solid constraints than the astrometric data.

Note that even for relatively simple models the number of parameters can be very large. Some parameters define the density profile: for example, the NFW density profile has two free parameters, namely the scale radius  $r_s$  and the characteristic density  $\rho_s$ . These must be multiplied by the number of lens components. Each of them is also characterized by an ellipticity and a position angle. The external shear also brings some parameter in. For complicated models, the number of parameters grows significantly. Conversely the constraints are usually few.

Another approach is the non-parametric one. The basic idea behind non-parametric mass models is that the effective lens potential and the deflection equations are linear functions of the surface density. As we saw earlier the surface density can be decomposed in multipoles but also in pixels or any other form in which the surface density is represented as a linear combination of density functionals multiplied by unknown coefficients  $\mathbf{k}$ . (Kochanek, 2004; Diego et al., 2005). In any such model the lens equation for image  $i$  takes the form

$$\vec{\beta} = \vec{\theta}_i - A_i \mathbf{k} \quad (5.10)$$

where  $A_i$  is the matrix that gives the deflection at the position of image  $i$  in terms of the coefficients of the surface density decomposition  $\mathbf{k}$ . An example is the following. Suppose that we observe  $N_{img}$  images and that the lens plane has been divided into  $N_c$

cells. Each cell contains a mass  $m_j$  with  $1 \leq j \leq N_c$ . Then, we have  $N_c$  coefficients which are the masses  $m_j$  and the matrix  $A$  is a matrix of  $2N_{img} \times N_c$  elements casting the mass vector into a vector of displacements angles. The factor 2 in front of  $N_i$  comes from the fact that the images (and the sources) are identified by two coordinates.

For a lens with  $N_{img}$  images of the same source, such a system can be solved exactly if there are enough degrees of freedom in the description of the surface density. For example, if there are two images, we can eliminate the position of the source, obtaining the system of equations

$$\vec{\theta}_1 - \vec{\theta}_2 = (A_1 - A_2)\mathbf{k}, \quad (5.11)$$

and we can solve it by taking the inverse of the matrix  $A_1 - A_2$ :

$$\mathbf{k} = (A_1 - A_2)^{-1}(\vec{\theta}_1 - \vec{\theta}_2). \quad (5.12)$$

Unfortunately there are usually more degrees of freedom than constraints. In the example given above, if the position of the sources are unknown, there are  $2N_{img}$  linear equations for  $2N_{img} + N_c$  unknowns. In order to identify a suitable solution for such a system, we need to add extra information or impose constraints.

One possibility is that of reducing the number of unknowns by removing the source positions from the unknown category. If the images are known to be multiple images of the same source, this can be achieved by minimizing the spread of the source points. If there are images which are known to derive from different sources, one can make the assumption that these sources are as much compact as possible.

#### 5.1.4 When theory crashes against reality: lensing near cusps

We consider now a particular class of images originated by sources close to the cusps of the tangential caustic. Such an image forms at the position  $\vec{x}^{(0)}$  close to the tangential critical line. The following conditions must be satisfied:

- given the Fermat potential

$$\phi(\vec{x}, \vec{y}) = \frac{1}{2}(\vec{x} - \vec{y})^2 - \psi(\vec{x}) \quad (5.13)$$

images form where  $\vec{\nabla}\phi = 0$ . Thus,

$$\phi_1^{(0)} = 0 = \phi_2^{(0)}; \quad (5.14)$$

- $\det A|_0 = 0$  obviously, given that, as said,  $\vec{x}^{(0)}$  is close to the critical line. Given that  $A_{ij} = \phi_{ij}$ ,

$$\phi_{11}^{(0)}\phi_{22}^{(0)} - (\phi_{12}^{(0)})^2 = 0; \quad (5.15)$$

- $\text{tr}A \neq 0$ , unless the source lays at the interception between the radial and the tangential caustic;
- $(\vec{\nabla} \det A)|_0 \neq 0$ , so that  $\det A$  changes sign by crossing the critical line;
- the image on the source plane of a vector tangent to the critical line must vanish (definition of cusp). Since  $(\vec{\nabla} \det A)|_0$  is normal to the critical line, a tangent vector is obtained by applying a rotation of  $\pi/2$ ,  $R(\pi/2)$ . If the source is on the cusp, the image on the source plane of such a vector tangent to the critical line must be

$$A_0 R(\pi/2)(\vec{\nabla} \det A)|_0 = 0. \quad (5.16)$$

We choose the coordinate axes such that the image  $\vec{x}^{(0)}$  and its source  $\vec{y}^{(0)}$  are at the origin, the axes on the lens and on the source planes are parallel and the Jacobian matrix at  $\vec{x}_0$  is diagonal  $A_{11}^{(0)} \neq 0$ ,  $A_{22}^{(0)} = 0$ . This implies that

$$\begin{aligned}\phi_{12}^{(0)} &= \phi_{21}^{(0)} = 0 \\ \phi_{11}^{(0)} &\neq 0 \\ \phi_{22}^{(0)} &= 0.\end{aligned}\quad (5.17)$$

From

$$(\vec{\nabla} \det A)|_0 = \vec{\nabla}[\phi_{11}^{(0)}\phi_{22}^{(0)} - (\phi_{12}^{(0)})^2] \neq 0 \quad (5.18)$$

together with 5.17, we obtain

$$\phi_{111}^{(0)}\phi_{22}^{(0)} + \phi_{11}^{(0)}\phi_{221}^{(0)} - 2\phi_{12}^{(0)}\phi_{121}^{(0)} = \phi_{11}^{(0)}\phi_{221}^{(0)} \neq 0, \quad (5.19)$$

and/or

$$\phi_{112}^{(0)}\phi_{22}^{(0)} + \phi_{11}^{(0)}\phi_{222}^{(0)} - 2\phi_{12}^{(0)}\phi_{122}^{(0)} = \phi_{11}^{(0)}\phi_{222}^{(0)} \neq 0, \quad (5.20)$$

so  $\phi_{122}^{(0)}$  and  $\phi_{222}^{(0)}$  cannot be both zero.

Moreover,

$$R(\pi/2)(\vec{\nabla} \det A)|_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \phi_{11}^{(0)} \begin{pmatrix} \phi_{221}^{(0)} \\ \phi_{222}^{(0)} \end{pmatrix} = \phi_{11}^{(0)} \begin{pmatrix} -\phi_{222}^{(0)} \\ \phi_{221}^{(0)} \end{pmatrix}. \quad (5.21)$$

Mapping to the source plane, we find

$$\begin{aligned}A_0 R(\pi/2)(\vec{\nabla} \det A)|_0 &= \begin{pmatrix} \phi_{11}^{(0)} & 0 \\ 0 & 0 \end{pmatrix} \phi_{11}^{(0)} \begin{pmatrix} -\phi_{222}^{(0)} \\ \phi_{221}^{(0)} \end{pmatrix} \\ &= \phi_{11}^{(0)} \begin{pmatrix} -\phi_{11}^{(0)}\phi_{222}^{(0)} \\ 0 \end{pmatrix}.\end{aligned}\quad (5.22)$$

Given Eq. 5.16, we obtain

$$\phi_{222}^{(0)} = 0 \quad (5.23)$$

and thus

$$\phi_{122}^{(0)} \neq 0. \quad (5.24)$$

Summarizing:

$$\phi_1^{(0)} = \phi_2^{(0)} = \phi_{12}^{(0)} = \phi_{22}^{(0)} = \phi_{222}^{(0)} = 0 \quad (5.25)$$

$$\phi_{11}^{(0)} \neq 0 \neq \phi_{122}^{(0)}, \quad (5.26)$$

thus,  $\phi$  near the critical point (cusp point) is

$$\begin{aligned}\phi &= \phi^{(0)} + \frac{1}{2}\vec{y}^2 - \vec{x}\vec{y} + \frac{1}{2}\phi_{11}^{(0)}x_1^2 + \frac{1}{6}\phi_{111}^{(0)}x_1^3 + \frac{1}{2}\phi_{112}^{(0)}x_1^2x_2 \\ &+ \frac{1}{2}\phi_{122}^{(0)}x_1x_2^2 + \frac{1}{24}\phi_{2222}^{(0)}x_2^4 + \dots\end{aligned}\quad (5.27)$$

Images form where  $\vec{\nabla}\phi = 0$ . Thus,

$$\begin{aligned}\partial_1\phi &= -y_1 + \phi_{11}^{(0)}x_1 + \frac{1}{2}\phi_{111}^{(0)}x_1^2 + \phi_{112}^{(0)}x_1x_2 + \frac{1}{2}\phi_{122}^{(0)}x_2^2 = 0 \\ \partial_2\phi &= -y_2 + \frac{1}{2}\phi_{112}^{(0)}x_1^2 + \phi_{122}^{(0)}x_1x_2 + \frac{1}{6}\phi_{222}^{(0)}x_2^3 = 0\end{aligned}\quad (5.28)$$

Then, the lens mapping near the cusp point is (using only the leading terms in  $x_1$  and  $x_2$ ):

$$\begin{aligned}y_1 &= cx_1 - \frac{b}{2}x_2^2 + dx_1x_2 \\ y_2 &= \frac{d}{2}x_1^2 - bx_1x_2 - ax_2^3,\end{aligned}\quad (5.29)$$

where

$$a = -\frac{1}{2}\phi_{222}^{(0)}, b = -\phi_{122}^{(0)}, c = \phi_{11}^{(0)}, d = \phi_{112}^{(0)}.\quad (5.30)$$

The Jacobian matrix is

$$A = \frac{\partial\vec{y}}{\partial\vec{x}} = \begin{pmatrix} c + dx_2 & dx_1 - bx_2 \\ dx_1 - bx_2 & -bx_1 - 3ax_2^2 \end{pmatrix}.\quad (5.31)$$

The determinant is

$$\begin{aligned}\det A &= -(c + dx_2)(bx_1 + 3ax_2^2) - (dx_1 - bx_2)^2 \\ &= -(cbx_1 + 3acx_2^2 + bdx_1x_2 + 3adx_2^3) - (d^2x_1^2 + b^2x_2^2 - 2bdx_1x_2) \\ &= -bcx_1 - (3ac + b^2)x_2^2 - d(dx_1^2 - bx_1x_2 + 3ax_2^3).\end{aligned}\quad (5.32)$$

All term in  $\det A$  involving  $d$  turn out to be negligible near the origin although this was not obvious before.

The critical curve requires  $\det A = 0$ , which implies

$$x_1 = -\frac{1}{bc}(3ac + b^2)x_2^2.\quad (5.33)$$

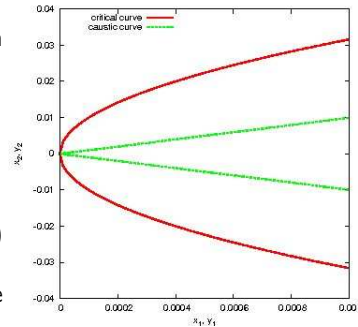
Thus the critical curve is a parabola (red curve in the Fig. on the right).

The caustic curve is obtained from Eq. 5.29:

$$\begin{aligned}y_1 &= -\frac{3ac + b^2}{b}x_2^2 - \frac{b}{2}x_2^2 = -\frac{3}{2b}(2ac + b^2)x_2^2 \\ y_2 &= \frac{3ac + b^2}{c}x_2^2 - ax_2^3 = \frac{1}{c}(2ac + b^2)x_2^3\end{aligned}\quad (5.34)$$

Then the caustic is a semicubic parabola (green curve in the Fig. on the right),

$$y_1^3 = -\frac{27c^2(2ac + b^2)}{8b^3}y_2^2.\quad (5.35)$$



Let us now study the inversion of the lens mapping 5.29 near the cusp. We have:

$$\begin{aligned}
y_1 &= cx_1 - \frac{b}{2}x_2^2 + dx_1x_2 \\
\Rightarrow x_1(c + dx_2) &= y_1 + \frac{b}{2}x_2^2 \\
\Rightarrow x_1 &= \frac{1}{c + dx_2} \left( y_1 + \frac{b}{2}x_2^2 \right) \\
&\approx \frac{1}{c} \left( y_1 + \frac{b}{2}x_2^2 \right), \tag{5.36}
\end{aligned}$$

and

$$\begin{aligned}
y_2 &= \frac{d}{2} - bx_1x_2 - ax_2^3 \\
&= \frac{d}{2c^2} \left( y_1^2 + by_1x_2^2 + \frac{b^2}{4}x_2^4 \right) - \frac{b}{c} \left( y_1x_2 + \frac{b}{2}x_2^3 \right) - ax_2^3 \\
&= \frac{d}{2c^2} \left( y_1^2 + by_1x_2^2 + \frac{b^2}{4}x_2^4 \right) - x_2^3 \left( a + \frac{b^2}{2c} \right) - \frac{b}{c}y_1x_2 \\
\Rightarrow x_2^3 \frac{2ac + b^2}{2c} + y_2 + \frac{b}{c}y_1x_2 - \frac{d}{2c^2} \left( y_1^2 + by_1x_2^2 + \frac{b^2}{4}x_2^4 \right) &= 0. \tag{5.37}
\end{aligned}$$

Neglecting the last term, we obtain

$$x_2^3 + \frac{2cy_2}{2ac + b^2} + \frac{2by_1x_2}{2ac + b^2} = 0. \tag{5.38}$$

The second of these equations is in third order in  $x_2$  and can be solved analytically. This solution can then be inserted in Eq. 5.36 to yield  $x_1$ . The number of solutions of Eq. 5.38 depends on the discriminant  $\Delta$ ,

$$\Delta = \frac{1}{(2ac + b^2)^2} \left[ c^2y_2^2 + \frac{8}{27} \frac{b^3y_1^3}{2ac + b^2} \right]. \tag{5.39}$$

If  $\Delta < 0$  three real solutions exist. This correspond to the case of a source inside the caustic. If  $\Delta > 0$  only one real solution exists, which correspond to the case of a source outside the caustic. If  $\Delta = 0$ , two solutions merge. It can be easily seen that

$$\Delta = 0 \Rightarrow y_1^3 = -\frac{27c^2(2ac + b^2)}{8b^3}y_2^2, \tag{5.40}$$

which is the equation of the caustic curve. Thus, two images disappear when the source crosses the caustic curve. The caustic separates the regions of one and three solutions on the source plane.

Since the general solution of Eqs. 5.36 and 5.38 is very complicated, we restrict to the special case  $y_2 = 0$ , when the source is located on the symmetry axis of the cusp. Then,

$$\begin{aligned}
x_2^3 + \frac{2by_1x_2}{2ac + b^2} &= 0 \\
\Rightarrow x_2 = 0 \text{ or } x_2^2 + \frac{2by_1}{2ac + b^2} &= 0 \\
\Rightarrow x_2 = 0 \text{ or } x_{2\pm} &= \frac{1}{2} \left[ \pm \sqrt{-\frac{8by_1}{2ac + b^2}} \right] \\
\Rightarrow x_2 = 0 \text{ or } x_{2\pm} &= \left[ \pm \sqrt{-\frac{2by_1}{2ac + b^2}} \right]. \tag{5.41}
\end{aligned}$$

- If  $x_2 = 0$ ,  $x_1 = \frac{y_1}{c}$  (from Eq. 5.36);
- if  $x_2 = \pm \sqrt{-\frac{2by_1}{2ac+b^2}}$ ,  $x_1 = \frac{y_1}{c} - \frac{b}{2c} \frac{2by_1}{2ac+b^2} = \frac{2ay_1}{2ac+b^2}$ ;
- thus, images form at

$$\left(\frac{y_1}{c}, 0\right); \left(\frac{2ay_1}{2ac+b^2}, \pm \sqrt{-\frac{2by_1}{2ac+b^2}}\right). \quad (5.42)$$

The magnifications are given by the inverse of the Jacobian determinant,

$$\det A = -bcx_1 - (3ac + b^2)x_2^2. \quad (5.43)$$

- For image 1:

$$\det A = -by_1; \quad (5.44)$$

- for image 2:

$$\begin{aligned} \det A &= -\frac{2abcy_1}{2ac+b^2} + (3ac+b^2)\frac{2by_1}{2ac+b^2} \\ &= \frac{(-2abc+6abc+2b^3)y_1}{2ac+b^2} = \frac{2by_1(b^2+2ac)}{2ac+b^2} \\ &= 2by_1, \end{aligned} \quad (5.45)$$

which is valid also for image 3.

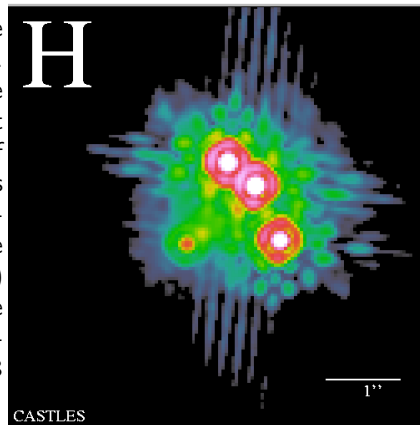
Thus:

$$\mu^{(1)} = -\frac{1}{by_1}; \quad \mu^{(2)} = \frac{1}{2by_1} = \mu^{(3)} \quad (5.46)$$

$$\mu^{(1)} + \mu^{(2)} + \mu^{(3)} = 0 \quad \text{signed magnifications} \quad (5.47)$$

$$|\mu^{(2)}| = \frac{1}{2}|\mu^{(1)}| \quad \text{unsigned magnifications} \quad (5.48)$$

Thus, for cusp lenses we expect the three brightest images to satisfy these conditions. In fold lenses, analogous relations can be found for the two brightest images. What happens in reality? Unfortunately none of the systems observed so far show flux ratios in agreement with the theoretical expectations. For example, the famous quadruple quasar B1422+231 (Patnaik et al., 1992) displayed on the right strongly violates these rules. The three brightest images are labeled in Patnaik et al. (1992) A, B and C. B is the brightest images.



In our previous discussion it corresponds to the image 1 (inside the critical line). The images A and C correspond to the images 2 and 3, which are outside the critical curve. Based on the theoretical expectations, the flux ratios  $A/B$  and  $C/B$  should be  $\sim 0.5$ ,

while the flux ratio  $A/C$  should be  $\sim 1$ . Instead, VLBA observations of this system found

$$\begin{aligned}\frac{A}{B} &\sim 0.9 \\ \frac{C}{B} &\sim 0.5 \\ \frac{C}{A} &\sim 0.5.\end{aligned}$$

Many authors have attempted at finding a solution to this problem. For example, Mao & Schneider (1998) suggested that such anomalies may be due to substructures in the lenses. However, recently Mao et al. (2004) using numerical simulations found that at typical image positions the fraction of the surface density in substructures is  $< 0.5\%$  in a  $\Lambda$ CDM cosmology. Such low fraction seems lower than required to explain the flux anomalies observed in B1422+231 and in other systems (e.g. B2045+265, B0712+472, B1933+503). The problem of the “flux anomalies” in lenses quasars is still unresolved. Generally, mass models can be found which fit very well all the observables but they fail at reproducing the flux ratios between the images.

### 5.1.5 General results from strong lensing

#### Galaxies

There are about 70 cases of strong lensing by galaxies now known. Most of them have two or four images, but a few have higher image numbers. Several important properties of these lenses have been revealed thanks to lensing. Some of them are:

- image splittings allow the projected mass to be constrained which is enclosed by the images. In particular, galaxy lensing suggests that galaxies have inner density profiles near to isothermal (Rusin et al., 2003). For example, we have seen in the previous Chapter that gravitational lenses are expected to produce odd number of images, unless the surface-mass distribution has a steep inner density profile. All but very few observed galaxy-lens systems have an even image number, most of them either two or four, as required by isothermal density profiles;
- the time delay between images of the same source, can be measured if the source is intrinsically variable. The same features then are repeated in the light curves of all the images but shifted in time. If a model for the lens mass distribution is given, the time delay can be used to measure the Hubble constant. The most recent determinations of  $H_0$  via lensing,  $H_0 = 75_{-6}^{+7} \text{ km s}^{-1} \text{ Mpc}^{-1}$  (Koopmans et al., 2003), are in agreement with the HST Key Project results, which give  $H_0 = (72 \pm 8) \text{ km s}^{-1} \text{ Mpc}^{-1}$  and with the WMAP constraints,  $H_0 = (74 \pm 3) \text{ km s}^{-1} \text{ Mpc}^{-1}$ ;
- the observed abundance of quadruples is particularly large, comparable to that of doubles. Such a large fraction of quads supports the view that galaxy lenses contain substructures, which contribute to the shear field in the lens, pushing the critical lines into regions of low density and then extending the lensing cross sections for quadruples. The large scale structure indeed seems to provide insufficient external shear for explaining the observed frequency of quadruples.

#### Clusters

Much can be learned on clusters studying the arcs that some of them produce:

- there must be more matter in clusters than what is contained in their galaxies, and that dark matter must be more smoothly distributed; otherwise, arcs would appear smaller and more strongly curved;
- prominent counter-arcs are often missing, thus clusters cannot typically be symmetric;
- arcs are thin, thus cluster density profiles must be steep; this can be seen as follows: for critical curves to appear, we need  $\kappa \sim 1$  in the core; the radial magnification of a tangential arc is

$$\mu_r = \frac{1}{1 - \kappa + \gamma}, \quad (5.49)$$

while  $1 - \kappa - \gamma = 0 \Rightarrow \gamma = (1 - \kappa)$ . Thus

$$\mu_r = \frac{1}{2(1 - \kappa)}. \quad (5.50)$$

Thin arcs require  $\mu_r \lesssim 1$ , which implies  $\kappa \lesssim 1/2$ . Thus, the density profile must be steep enough to drop from  $\kappa \sim 1$  to  $\kappa \sim 1/2$  within  $\sim \theta_E$ ;

- cluster cores, if they exist, need to be small, because otherwise no arcs would be formed, or radial arcs would be at much larger cluster centric distances;
- substructures need to be abundant in galaxy clusters, because otherwise “straight arcs” could not appear;
- the simple estimates of cluster properties using arcs, radial or tangential, are complicated in reality by cluster substructures or asymmetries; shear plays an important role.

## 5.2 Weak lensing

### 5.2.1 Weak lensing by galaxy clusters

Due to the very high number density of distant galaxies, clusters appear placed in front of a “cosmic wallpaper”, which they distort in a very characteristic way. There are about 30-40 galaxies per square arcmin of the sky (but in deep space observations the amazing density of 80-100 galaxies per square arcmin has been reached!), thus approximately 20000 of them on the face of the full moon. Can the distortions imprinted on these galaxies by clusters be used for reconstructing cluster mass distributions?

While source sizes are unknown individually, source shapes can more easily be measured. Although sources are intrinsically irregular, their large number allows averaging over several of them. Assuming that they are intrinsically randomly oriented, the average over their shapes should become circular.

If lensing is weak, the image of a circular source appears elliptical, with axes given by

$$\frac{r}{1 - \kappa - \gamma}, \quad \frac{r}{1 - \kappa + \gamma}, \quad (5.51)$$

assuming  $\kappa, \gamma \ll 1$ . Defining the ellipticity as

$$\epsilon = \frac{a - b}{a + b} = \frac{2\gamma}{2(1 - \kappa)} = \frac{\gamma}{1 - \kappa} \approx \gamma \quad (5.52)$$

in the limit of weak lensing, i.e. the ellipticity directly measures the shear.

Like the shear, the ellipticity has two components, an orientation and a magnitude, which we can write as

$$\epsilon_1 = \epsilon \cos 2\phi \quad (5.53)$$

$$\epsilon_2 = \epsilon \sin 2\phi. \quad (5.54)$$

Note the dependence on  $2\phi$  rather than  $\phi$ , which illustrates that an ellipse is symmetric under rotations by  $\pi$  rather than  $2\pi$ .

Let the intrinsic source ellipticity be  $\epsilon^{(s)}$ , then, in the weak-lensing limit,

$$\epsilon_i = \epsilon_i^{(s)} + \gamma_i = \epsilon^{(s)} \begin{pmatrix} \cos 2\phi \\ \sin 2\phi \end{pmatrix} + \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}; \quad (5.55)$$

averaging over sufficiently many sources, the first term disappears and

$$\langle \epsilon \rangle = \langle \gamma \rangle. \quad (5.56)$$

Suppose now that we know  $\gamma$  at many positions  $\vec{\theta}$  across a cluster, how can we find the mass distribution? The key is the relation between both  $\gamma$  and  $\kappa$  and the lensing potential  $\psi$ . If we transform  $\psi$ ,  $\gamma$  and  $\kappa$  to Fourier space, we find

$$\kappa = \frac{1}{2}(\psi_{11} + \psi_{22}) \Rightarrow \hat{\kappa} = -\frac{1}{2}(k_1^2 + k_2^2)\hat{\psi} \quad (5.57)$$

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}) \Rightarrow \hat{\gamma}_1 = -\frac{1}{2}(k_1^2 - k_2^2)\hat{\psi} \quad (5.58)$$

$$\gamma_2 = \psi_{12} \Rightarrow \hat{\gamma}_2 = -k_1 k_2 \hat{\psi}, \quad (5.59)$$

where  $\vec{k}$  is a wave vector conjugate to the angular position vector  $\vec{\theta}$ . From these relations, we can eliminate  $\hat{\psi}$ . Noting that

$$\left[ k^{-2} \begin{pmatrix} k_1^2 - k_2^2 \\ 2k_1 k_2 \end{pmatrix} \right] [k^{-2} (k_1^2 - k_2^2 \quad 2k_1 k_2)] = 1 \quad (5.60)$$

and writing

$$\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} = k^{-2} \begin{pmatrix} k_1^2 - k_2^2 \\ 2k_1 k_2 \end{pmatrix} \hat{\kappa}, \quad (5.61)$$

we find immediately

$$\hat{\kappa} = k^{-2} (k_1^2 - k_2^2 \quad 2k_1 k_2) \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix} = k^{-2} [(k_1^2 - k_2^2)\hat{\gamma}_1 + 2k_1 k_2 \hat{\gamma}_2] \quad (5.62)$$

As a product in Fourier space, we can equally well write the result as a convolution in real space, using the Fourier convolution theorem,

$$(f \hat{*} g) = \hat{f} \hat{g}. \quad (5.63)$$

This yields, in real space,

$$\kappa(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' [D_1(\vec{\theta} - \vec{\theta}')\gamma_1 + D_2(\vec{\theta} - \vec{\theta}')\gamma_2], \quad (5.64)$$

with

$$D_1(\vec{\theta}) = \frac{\theta_2^2 - \theta_1^2}{\theta^4} \quad (5.65)$$

$$D_2(\vec{\theta}) = \frac{2\theta_1\theta_2}{\theta^4}. \quad (5.66)$$

In that way, mass maps of galaxy clusters can be constructed from shear measurements. However, there are numerous problems in detail, which can all be overcome to some degree. First of all, ellipticities measure

$$\left\langle \frac{\gamma}{1 - \kappa} \right\rangle = \langle g \rangle, \quad (5.67)$$

the so-called *reduced* shear, rather than the shear alone. This can be overcome by writing

$$\gamma = g(1 - \kappa) \quad (5.68)$$

and using that in the convolution equation, thus

$$\kappa(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' [D_1(\vec{\theta} - \vec{\theta}')g_1(1 - \kappa) + D_2(\vec{\theta} - \vec{\theta}')g_2(1 - \kappa)], \quad (5.69)$$

which can be solved analytically starting from  $\kappa = 0$ .

Formally, the convolution for  $\kappa$  extends over all of two-dimensional space, while real data fields are finite. In practice, this leads to a bias if the field is small, but modern data fields are large enough for the method to work straightforwardly.

A problem of principle arises because the Jacobian can be transformed by multiplying it with a factor  $\lambda \neq 0$ ,

$$A \rightarrow \lambda A \equiv A', \quad (5.70)$$

without the ellipticity measurements noticing it. Scaling the Jacobian in that way simply enlarges the sources, but does not change their shape. Thus, such transformations leave the shear signal invariant. Thus, transformations like

$$1 - \kappa' = \lambda(1 - \kappa) \Rightarrow \kappa' = 1 - \lambda + \lambda\kappa \quad (5.71)$$

cannot be detected, and  $\kappa$  is only determined up to such transformations. For  $\lambda = 1 - \delta \approx 1$ ,

$$\kappa' \approx \delta + \kappa, \quad (5.72)$$

which corresponds to adding a sheet of constant surface mass density to the lens. Therefore, this invariance has been called the *mass sheet degeneracy*.

There are other techniques for cluster reconstruction which can in principle avoid the mass sheet degeneracy. Among them are the maximum-likelihood techniques. Suppose shear measurements are given on a cluster field, then the goal is to search a lensing potential  $\psi$  such as to reproduce the shear. This can be done by minimizing

$$\chi^2 = \sum_{pixels} \frac{[\gamma_1 - \gamma_1(\psi)]^2 + [\gamma_2 - \gamma_2(\psi)]^2}{2\sigma_\gamma^2}, \quad (5.73)$$

where  $\gamma_1(\psi)$  and  $\gamma_2(\psi)$  are the usual relations between shear and potential. This minimization leads to an estimate for  $\psi$  at each point where a shear value has been measured.

Additional information can now be added. Suppose we have a means to estimate galaxy sizes, then the magnification is given by

$$\mu = \frac{d^2\Omega_{lensed}}{d^2\Omega_{unlensed}}. \quad (5.74)$$

For weak lensing, the inverse magnification is

$$R = \frac{1}{\mu} = (1 - \kappa)^2 - \gamma^2 \approx 1 - 2\kappa. \quad (5.75)$$

This can then be incorporated into  $\chi^2$ :

$$\chi^2 = \sum_{pixels} \left\{ \frac{[\gamma_1 - \gamma_1(\psi)]^2 + [\gamma_2 - \gamma_2(\psi)]^2}{2\sigma_\gamma^2} + \frac{[R - R(\psi)]^2}{2\sigma_R^2} \right\}, \quad (5.76)$$

which is now no longer invariant under the mass sheet transformation.

In practice smoothing is required to suppress the noise in the data (shot noise, intrinsic ellipticity) to a desired level.

## 5.2.2 Galaxy-galaxy lensing

Faint background galaxies are also lensed by brighter foreground galaxies. Although the effect is certainly weak, galaxies can be superposed to enhance the signal.

Let  $p^{(s)}(\epsilon)$  be the intrinsic distribution of source ellipticities. The shear caused by a lens changes the ellipticity to

$$\epsilon \rightarrow \epsilon + \gamma, \quad (5.77)$$

thus, the observed distribution of image ellipticities is

$$p(\epsilon) = p^{(s)}(\epsilon - \gamma) \approx p^{(s)}(\epsilon) - \gamma_i \left( \frac{\partial p^{(s)}}{\partial \epsilon_i} \right) (\epsilon). \quad (5.78)$$

We introduce now two angles: a position angle of the lensed galaxy with respect to the lens,  $\alpha$ , and the angle between the major axis of the lensed galaxy and the line connecting it with the lens,  $\phi$ .

Without loss of generality, we can rotate the coordinate system such that  $\alpha = 0$ . Then, we are in the principal axis system of the shear, such that

$$\gamma'_1 = -\gamma_t = \gamma_1 \cos 2\alpha + \gamma_2 \sin 2\alpha, \quad (5.79)$$

$$\gamma'_2 = 0 \quad (5.80)$$

We have defined the tangential component of the shear here as

$$\gamma_t = -(\gamma_1 \cos 2\alpha + \gamma_2 \sin 2\alpha). \quad (5.81)$$

The example of a singular isothermal sphere shows that this definition is indeed useful. There, we had

$$\gamma_1 = -\gamma \cos 2\alpha \quad (5.82)$$

$$\gamma_2 = -\gamma \sin 2\alpha, \quad (5.83)$$

such that

$$\gamma_t = -[-\gamma \cos^2 2\alpha - \gamma \sin^2 2\alpha] = \gamma. \quad (5.84)$$

Note again that the shear components are elements of a  $2 \times 2$  tensor, so that they transform under rotations as

$$\begin{aligned} \begin{pmatrix} \gamma'_1 & \gamma'_2 \\ \gamma'_2 & -\gamma'_1 \end{pmatrix} &= \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \gamma_1 \cos 2\alpha + \gamma_2 \sin 2\alpha & -\gamma_1 \sin 2\alpha + \gamma_2 \cos 2\alpha \\ -\gamma_1 \sin 2\alpha + \gamma_2 \cos 2\alpha & -\gamma_1 \cos 2\alpha - \gamma_2 \sin 2\alpha \end{pmatrix} \end{aligned} \quad (5.85)$$

We now return to the probability distribution of the image ellipticity,  $p(\epsilon)$ . The derivatives of the intrinsic ellipticity distribution can be written as

$$\left( \frac{\partial p}{\partial \epsilon_1}, \frac{\partial p}{\partial \epsilon_2} \right) \rightarrow \frac{\partial p}{\partial \epsilon} \cos 2\phi \quad (5.86)$$

in the principal axis frame of the shear, assuming that ellipticities are intrinsically isotropic, such that  $\partial p / \partial \phi = 0$ . Then

$$p(\epsilon) = p^{(s)}(\epsilon) + \gamma_t \cos 2\phi \frac{\partial p^{(s)}}{\partial \epsilon}. \quad (5.87)$$

Integrating now over  $\int_0^\infty \epsilon d\epsilon$  in order to find the angle distribution of the images, we find

$$\begin{aligned} p(\phi) &= \frac{2}{\pi} + \gamma_t \cos 2\phi \int_0^\infty \epsilon d\epsilon \frac{\partial p^{(s)}}{\partial \epsilon} \\ &= \frac{2}{\pi} - \gamma_t \cos 2\phi \int_0^\infty \epsilon d\epsilon \frac{1}{\epsilon} p^{(s)}(\epsilon) \\ &= \frac{2}{\pi} \left[ 1 - \gamma_t \cos 2\phi \left\langle \frac{1}{\epsilon^{(s)}} \right\rangle \right]. \end{aligned} \quad (5.88)$$

here, the factors  $2/\pi$  arise from the fact that  $p(\phi)$  must be normalised when integrating  $\phi$  from 0 to  $\pi/2$ , because it suffices to restrict the angle  $\phi$  to that interval.

Thus, galaxy-galaxy lensing modifies the distribution of the position angles of the lensed galaxies to

$$p(\phi) = \frac{2}{\pi} \left[ 1 - \langle \gamma_t \rangle \cos 2\phi \left\langle \frac{1}{\epsilon^{(s)}} \right\rangle \right], \quad (5.89)$$

where  $\langle \gamma_t \rangle$  is the mean tangential shear of an ensemble of lensing galaxies.  $p(\phi)$  can be observed once an estimate for  $\langle 1/\epsilon^{(s)} \rangle$  is known, which can be directly inferred from observed galaxy images. Then,  $\langle \gamma_t \rangle$  can be estimated and lens galaxy properties can be derived from it.