Gravitational Lensing Applications

Since the revival of interest for gravitational lensing in the 1960s, many applications of lensing have been designed to infer cosmological and astrophysical information. With the development of observational techniques, esp. the replacement of photometric plates with CCD detectors, the development of radio interferometry, high resolution imaging with HST, and the increasing of size and depth of optical / near IR surveys, many of these applications have been realized. This lecture intends to give an introduction to different types of applications, and then focus on weak lensing applications. For more background knowledge, and the derivations of equations which will be missing in this notes, see

Ref 1. "Gravitational Lensing: Strong, Weak, and Micro" P. Schneider, C. Kochanek, J. Wambsganss 2006

Ref 2. "Gravitational Lenses" P. Schneider, J. Ehlers, E. E. Falco 1992

Ref 3. "Weak Gravitational Lensing" M. Bartelmann, P. Schneider 2001

Some more recent reviews include

- A. Lewis, A. Challinor 2006
  "Weak gravitational lensing of the CMB"

- D. Munshi et al. 2008
  "Cosmology with weak lensing surveys"
- H. Hoestra, B. Jain 2008
  "Weak gravitational lensing and its cosmological applications"
- T. Treu 2010
  "Strong lensing by galaxies"
- S. Mao 2012
  "Astrophysical applications of gravitational microlensing"
- M. Meneghetti et. al. 2013
  "Arc statistics"

I. Fundamentals of Gravitational Lensing

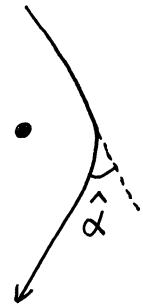
Gravitational lensing is based on 2 fundamental theories:

1) General Relativity − mass distorts spacetime
2) Fermat Principle − light follows paths with stationary light travel time, or: light follows null geodesics

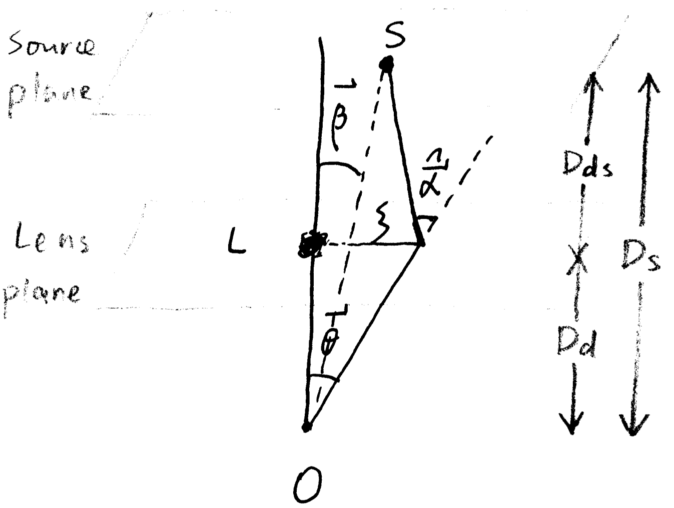
All gravitational lensing studies work in the limit of weak gravitational field. Light distortion near black holes is considered as an other field of study. The only places General Relativity appear in gravitational lensing are the cosmological distances, and the factor of "4" instead of "2" (Newtonian) in the expression of the deflection angle.

$$\hat{\alpha} = \frac{4GM}{c^2 \xi}$$

(for a point mass)



* The key to understand and describe gravitational lensing phenomena, is to see lensing as a "mapping".



mapping between the observe angular position $\vec{\theta}$ of the imag and the true angular positio $\vec{\beta}$ of the source.

Lens equation:

$$\vec{\beta} = \vec{\theta} - \frac{D_{ds}}{D_s} \hat{\vec{\alpha}}(\vec{\theta}) \equiv \vec{\theta} - \vec{\alpha}(\vec{\theta})$$

cosmological distances

position and "strength" of the lens

II. The scalar, vector, and tensor form of lens mapping

Here 3 ways of describing the mapping $\vec{\theta} \to \vec{\beta}$ will be presented, and lensing applications will be classified according to the way they are most conveniently described.

Form 1. the Fermat Principle — scalar form, $\sim \psi$ deflection potent

$$\nabla \tau(\vec{\theta}; \vec{\beta}) = 0$$

τ is the Fermat potential defined as

$$\tau \equiv \frac{1}{2} |\vec{\theta} - \vec{\beta}|^2 - \psi(\vec{\theta}),$$

with ψ being the deflection potential, defined to let $\vec{\alpha} = \nabla\psi$.

For a mass distribution confined at a certain position in the light propagation direction (~ for a "thin lens"), the deflection potential is linked to the mass distribution as

$$\psi(\vec{\theta}) = \frac{4G}{c^2} \frac{D_d D_{ds}}{D_s} \int d^2\theta' \ln|\vec{\theta} - \vec{\theta}'| \cdot \int dr_3 \cdot \rho(r_3, \vec{\theta}').$$

The Fermat potential is linearly related to the light travel time,

$$T(\vec{\theta}, \vec{\beta}) = \frac{D_d D_s}{c D_{ds}} (1 + z_d) \cdot \tau + \text{const}.$$

The first part in τ, the $\frac{1}{2}|\vec{\theta} - \vec{\beta}|^2$ describes the deviation of the light ray from a straight line; the $\psi(\vec{\theta})$ term is propotional to the time delay when a light ray traverses potential ψ.

In a lensing system with multiple images of a intrinsically variable source, e.g. strong lensing of quasar, the differences of light travel time beteen 2 images is an observable ("time delay").

Time delay in quasar strong lensing is used to determine the Hubble parameter.

Form 2. the Lens equation — vector form, $\sim \psi'$

deflection angle

$$\vec{\beta}(\vec{\theta}) = \vec{\theta} - \vec{\alpha}(\vec{\theta})$$

For a thin lens,

$$\vec{\alpha}(\vec{\theta}) = \nabla\psi(\vec{\theta}) = \frac{4G}{c^2}\frac{D_d D_{ds}}{D_s}\int d^2\theta' \frac{\vec{\theta}-\vec{\theta}'}{|\vec{\theta}-\vec{\theta}'|^2}\int dr_3\, \rho(r_3, \vec{\theta}')\,.$$

- Which redshift to put a lens to achieve maximum deflection?

— Maximize $\frac{D_d D_{ds}}{D_s}$.

One (and maybe the only) example where the deflection angle $\vec{\alpha}$ is regarded as an observable*, is CMB lensing.

* note: only in a statistical sense

With the CMB as a background source, the lensing effect is that temperature at one position is mapped to another (and polarization too), $T(\vec{\beta}) \rightarrow T(\vec{\theta})$.

A simple treatment of CMB lensing Taylor expands the T-map (deflection angle is small. $\alpha \sim 2'$)

$$\tilde{T}(\vec{\theta}) = T(\vec{\beta}) + \nabla T \cdot (\vec{\theta} - \vec{\beta}) + \cdots$$

Since unlensed temperature field is well-approximated by a Gaussian random field, the deflection angle $\vec{\alpha} = \vec{\theta} - \vec{\beta}$ can be estimated with the degree of deviation of the observed temperature field from Gaussian.

Compared to other forms of lensing, CMB lensing is special for having a single source plane at high redshift, and that the signal appears on large angular scales (deflection power spectrum peaks at about $\ell \sim 40$) where linear theory works well. It can break the angular diameter distance degeneracy of CMB power spectrum studies, and is especially constraining for early dark energy models.

Form 3. the Jacobi Matrix — tensor form, $\sim \psi''$

image distortion; magnification

$$A \equiv \frac{\partial \vec{\beta}}{\partial \vec{\theta}} = \delta_{ij} - \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 \tau}{\partial \theta_i \partial \theta_j} \equiv \tau_{,ij}$$

$$= \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}$$

with $\kappa \equiv \frac{1}{2} \nabla^2 \psi$, called the "dimensionless surface mass density" or the "convergence",

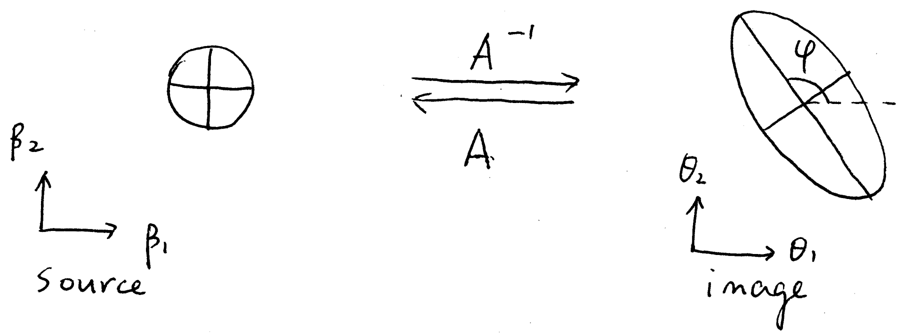
and $\gamma \equiv \gamma_1 + i\gamma_2 \equiv |\gamma| e^{2i\varphi}$

$$\equiv \frac{1}{2} (\psi_{,11} - \psi_{,22}) + i \psi_{,12} = \frac{1}{2} \partial^2 \psi$$

with $\partial \equiv \partial_1 + i \partial_2$.

being the "shear" which describes the traceless part of the Jacobian.

See the effect of the Jacobian on a image:



major axis $\frac{\text{lensed}}{\text{unlensed}}$

$1 - \kappa + |\gamma|$

minor axis $\frac{\text{lensed}}{\text{unlensed}}$

$1 - \kappa - |\gamma|$

The total magnification of the image equals the ratio of lensed and unlensed sizes, $\mu = \frac{1}{\det A} = \frac{1}{(1-\kappa)^2 - |\gamma|^2}$, due to the conservation of surface brightness under a lens mapping.

Almost all other lensing applications use shear or magnification as their observables.

III. Strong, weak, and micro lensing

One can also classify lensing applications by asking

- How strong is the lensing effect?

  or related questions: is the lens mapping 1-1 ?

  are there singularities in the mapping?

- Strong lensing occurs around singularities of the lens mapping where $\det A = 0$, and a single light source can be mapped to several images. The study of singularities of differential mapping is the catastrophy theory, see eg. book by Arnold 1984.
- Weak lensing by definition lies well in the 1-1 mapping regime. There the deflection potential & deflection angle are small, and the Jacobian is close to the identity matrix, meaning $|\gamma| \ll 1$, $\kappa \ll 1$.
- Micro lensing refers to situations of strong lensing for which the separation between multiple images is too small to resolve, and is usually observed as a variation in the flux of the source.

A more detailed classification requires asking "what is the soure, the lens, and the observable" in a particular lensing system. Below is a partial list of lensing applications classified this way.

	source	lens	observable	can infer
STRONG LENSING	galaxy	cluster	giant arcs (incl. Einstein rings), multiple images	cluster inner mass profile; properties of very high-z galaxies
	quasar	galaxy	time delay	H_0
			position & shape of multiple images & arcs	galaxy mass
			differential reddening of optical images	ISM of lens galaxy
			violation of cusp relation ...	substructure ?
MICRO LENSING	star	star (+ planet)	spikes in microlensing lightcurve	search for extrasolar planets
	star	MACHO	statistics of microlensing events	existence ? of MAssive Compact Halo Objects
	quasar	star (in foreground galaxy)	multifrequency light curves of multiple images	size of quasar emission region

... size of θ_E ? cross section?

WEAK LENSING

source	lens	"observable"	can infer
galaxy	cluster	$\gamma_t(\theta)$ around 1 cluster	mass profile of cluster
galaxy	galaxy	$\gamma_t(\theta)$ around stacked galaxies at small θ	mass profile of galaxy (averaged)
		* $\langle \kappa_g \gamma_t \rangle (\theta)$ for large θ	galaxy bias
galaxy	LSS	statistics of γ ; statistics of magnification	matter power spectrum ; cosmological parameters

* $$\kappa_g := \frac{N(\vec{\theta}) - \bar{N}}{\bar{N}}$$

Why is the source for weak lensing always galaxies ?

- Doesn't have to be, but weak lensing requires statistical power, and distant galaxies observed in the optical or near-IR band form the densest population of distant objects in the sky.

  The same reason explains the need for large and deep photometric (+spestrocopy) surveys in weak lensing.

Cosmic Shear

Light bundles get continuously deflected and distorted by the gravitational field of the inhomogeneous mass distribution. The corresponding shape distortion of light sources is usually very mild, typically of the order of a few percent. Thus cosmic shear has to be detected and studied in a statistical way, using images of a large number of distant galaxies.

The direct observable of cosmic shear is the ellipticities of galaxy images ϵ (measured from 2nd moment of galaxy surface brightness distribution). In the weak lensing limit ($\kappa \ll 1$, $|\gamma| \ll 1$),

$$\epsilon = \epsilon_i + \gamma$$

where ϵ_i being the intrinsic ellipticity of the galaxy. It is not only non-zero, but even dominates ϵ. So one has to work under the (reasonable) assumption that $\langle \epsilon_i \rangle = 0$ (no preferred direction). This is another reason that γ can be only measured in a statistical way. — ϵ is an unbiased but very noisy estimator of γ.

- Lens mapping for cosmic shear

For cosmic shear, the lens is no longer a localized object, but all the intervening matter between the source and the observer. The form of lens mapping in this case can be worked out by considering the geodesic equation in a slightly disturbed metric.

See e.g. Sect 6.1 of Ref. 1 for a derivation. Amazingly, the formalism of gravitational lensing for a single lens plane still holds for cosmic shear to the linear order (under the "Born Approximation", to be more precise). Under this approximation, lensing by the 3D matter distribution can be treated as being lensed by an equivalent lens plane, with deflection potential ψ in the form of

$$\underbrace{\psi(\vec{\theta}, \chi)}_{\substack{\text{source position} \\ \text{(comoving distance)}}} = \frac{3H_0^2\Omega_m}{2\pi c^2} \int d^2\theta' \ln|\vec{\theta}-\vec{\theta}'| \int_0^{\chi} \underbrace{d\chi'}_{\text{lens position}} \frac{f_K(\chi-\chi')f_K(\chi')}{\underbrace{f_K(\chi)}_{\text{comoving ang. dia. distan}}} \frac{\delta(f_K(\chi')\vec{\theta}, \chi')}{a(\chi')}$$

then we have still $\kappa = \frac{1}{2}\nabla^2\psi$, $\gamma = \frac{1}{2}\partial^2\psi$. The form of κ reads

$$\kappa(\vec{\theta}, \chi) = \frac{3H_0^2\Omega_m}{2c^2} \int_0^{\chi} d\chi' \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)} \frac{\delta(f_K(\chi')\vec{\theta}, \chi')}{a(\chi')}$$

- the equivalent convergence of a source at comoving distance χ ~~with~~ which is observed at angular position $\vec{\theta}$.
- projected density field.

In cosmic shear studies, one usually considers the lensing distortion of a population of source galaxies which spread along the line of sight. It's useful to define the convergence for a particular population of galaxies

$$\kappa^{(i)}(\vec{\theta}) = \int_0^{r_h} d\chi\, p_s^{(i)}(\chi)\, \kappa(\vec{\theta}, \chi)$$

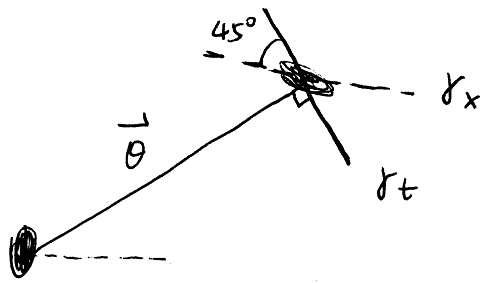
with $p_s^{(i)}(\chi)$ being the distance probability distribution of this galaxy population.

- Cosmic Shear 2-point statistics

In weak lensing studies, the shear γ is the "observable" (estimated from ϵ) and the convergence κ is the "theory" (theoretically predicted from the density field and cosmological distances). For cosmic shear, the most standard way of linking "observable" to "theory" is through 2-point statistics, i.e. the 2-pt statistics of shear is measured from galaxy shear catalog, and then used to constrain cosmology with the help of the relation between them and the 2-pt statistics of the convergence.

Since the shear field is a polarization field which we describe with a complex number at each position, the 2-pt correlator of shear at two positions separated by $\vec{\theta} = \theta e^{i\varphi}$ correspond to more than 1 real function. Define tangential and cross-component of the shear regarding to the direction of $\vec{\theta}$, as

$$\gamma_t \equiv -\mathrm{Re}(\gamma e^{-2i\varphi}), \qquad \gamma_\times \equiv -\mathrm{Im}(\gamma e^{-2i\varphi}).$$



One can then define their 2-pt correlation functions $\langle \gamma_t \gamma_t \rangle(\theta)$, $\langle \gamma_t \gamma_\times \rangle(\theta)$, and $\langle \gamma_\times \gamma_\times \rangle(\theta)$. Among them, $\langle \gamma_t \gamma_\times \rangle(\theta)$ is expected to vanish due to parity symmetry, and $\langle \gamma_t \gamma_t \rangle$, $\langle \gamma_\times \gamma_\times \rangle$ are usually combined into

$$\xi_\pm(\theta) \equiv \langle \gamma_t \gamma_t \rangle(\theta) \pm \langle \gamma_\times \gamma_\times \rangle(\theta), \quad \text{with}$$

$$\xi_+ = \langle \gamma \gamma^* \rangle, \qquad \xi_- = \langle \gamma \gamma \rangle e^{-4i\varphi}.$$
$$= \langle \kappa \kappa \rangle$$

They relate to the convergence power spectrum as

$$\xi_+(\theta) = \int_0^\infty \frac{d\ell\,\ell}{2\pi} J_0(\ell\theta)\, P_\kappa(\theta),$$

$$\xi_-(\theta) = \int_0^\infty \frac{d\ell\,\ell}{2\pi} J_4(\ell\theta)\, P_\kappa(\theta).$$

A number of other 2-pt statistics have been proposed to extract cosmological information from the data, e.g. the shear dispersion, the aperture mass statistics, the ring statistics, the "COSEBIs", Most of them with the motivation of separating E- and B- mode signals to lend a control on the systematics. All these statistics can be derived from $\xi_\pm$. (see Ref. 1 & Schneider+10)

In cosmic shear studies, the angular correlation function $\xi_\pm$ are usually the statistics to be directly applied on the data (shear catalog), instead of the power spectrum of shear. Why? The main reason is that shear signal peaks at small angular scales, $\ell^2 P_\kappa(\ell)$ for a mean source redshift of 1.5 peaks at $\ell \sim 10^4$, corresponding to $\theta \sim 1'$. At these angular scales, there are many stars, satellite tracks, etc. to be masked out of an optical / near IR survey of galaxies. The survey geometry after masking is usually highly complicated and is hard to properly take account of when Fourier space statistics are used, and thus the motivation of using correlation functions which are not much affected by complicated survey geometry. In addition, the density fluctuations that cosmic shear measures at those small angular scales (and at low redshifts) are highly non-Gaussian, therefore the biggest advantage of using Fourier space analysis (independent signal at each ℓ-mode for Gaussian random fields) loses its significance.

A more fundamental question: why use shear statistics? why not construct a convergence map out of shear signal and measure 2-pt statistics of κ on that map? One does want to have a κ-map for other purposes, but for constraining cosmological parameters with cosmic shear, it is 1) not necessary, 2) can be easily biased due to the non-local $\gamma-\kappa$ relation and the complicated survey geometry.

- The size, depth, and photometric quality of galaxy surveys have significantly increased in recent years. e.g. sizes of largest "lensing quality" surveys:

VIRMOS-DESCART	~ 2000 A.D.	8.5 deg^2
CFHTLens	~ 2010 A.D.	~ 150 deg^2
DES	2010 - 2020 A.D.	~ 5000 deg^2
EUCLID	2020+ A.D.	~ 15000 deg^2

i.e. the statistical power of cosmic shear will soon be increased by an order of magnitude, making it highly competitive in constraining cosmology. In particular, it is the most promising probe of dark energy.

However, this also means that cosmic shear study will soon enter systematics dominated era. Usually, systematics to cosmic shear studies are classified into 3 catagories:

- Observational systematics
  e.g. in shape measurements (PSF) $\longrightarrow$ go to space; "lensing quality data
  photo-z measurements $\longrightarrow$ spectroscopy for calibration
  (note: z information not critical for statistical power, but crutial for removing intrinsic alignments)

- Theoretical systematics

  e.g. theoretical prediction for non-linear power spectrum and effects of baryons → simulations and understanding simulations

- Astrophysical systematics

  e.g. the intrinsic alignments

  $\langle \epsilon_i^2 \rangle \neq 0$ (intrinsic alignment, only for auto-correlations within a redshift bin)

  Hirata & Seljak 04: $\langle \epsilon_i \gamma \rangle \neq 0$ (intrinsic-shear alignment, for cross-correlations between different z-bins)

  → modeling / nulling / combine with $\langle g\gamma \rangle$, $\langle gg \rangle$

V. Cluster Weak Lensing

To use clusters for cosmology, one key step is to determine their masses. Whereas X-ray and SZ measurements constrain M_{gas}, lensing and dynamical methods give the total gravitating mass. Different probes are sensitive to mass at different radii, in particular, strong lensing at $0.01\,r_{vir} \lesssim r \lesssim 0.1\,r_{vir}$, weak lensing at $0.1\,r_{vir} \lesssim r \lesssim 1\,r_{vir}$. At even smaller radii $r < 0.01 r_{vir}$ stellar dynamics in the central LBG is needed; galactic dynamics can help constraining the mass profile in the range of weak lensing (but usually with a lower S/N), and can extend to even larger radii $r \sim$ a few $* r_{vir}$.

Cluster weak lensing uses weakly distorted background galaxy images to derive the mass of clusters. The central problem is still $\gamma \to \kappa$.

Naively, this could be done by using the mathematical relation between γ and κ,

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

$$\mathcal{D}(\vec{\theta}) \equiv \frac{1}{(\theta_1 - i\theta_2)^2}, \quad \text{which is the real space}$$

correspondence of the simpler Fourier space relation

$$\tilde{\gamma}(\vec{\ell}) = \pi^{-1} \tilde{\mathcal{D}}(\vec{\ell}) \, \tilde{\kappa}(\vec{\ell})$$
$$= e^{2i\varphi_\ell} \, \tilde{\kappa}(\vec{\ell}) \qquad (\text{for } \ell \neq 0).$$

The corresponding $\gamma \to \kappa$ relation is

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

(Kaiser & Squires 1993)

the undetermined $\ell = 0$ mode

This direct inversion method suffers from many difficulties:

1). infinite noise due to θ^{-2} behavior of the kernel $\mathcal{D}$

$\rightarrow$ smoothing required

2). data availabe only on a finite field v.s. non-local relation of $\gamma \rightarrow \kappa$

$\rightarrow$ can use local relation of derivatives of γ and κ

3) additive constant κ_0 (~ mass sheet degeneracy)

$\rightarrow$ use compensated filters to derive mass proxy?

4) true "observable" is reduced shear $g \equiv \gamma/(1-\kappa)$ instead of γ.

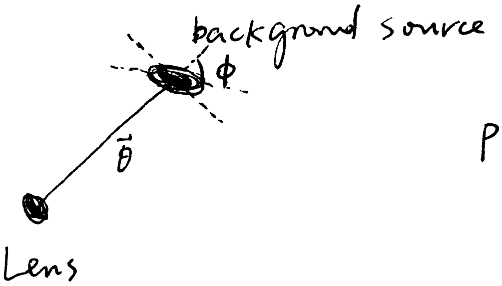
$\rightarrow$ no simple solution

Alternative methods are discussed in Ref. 1. For most current applications on real data, parametrized mass models have been used to obtain a maximum likelihood fit to the data (here the "raw data" i.e. galaxy catalog is used instead of statistics!). There exists a typical scatter of ~ 20% for 3D cluster mass determine with weak lensing signals only (and for a single cluster) even if only 1 parameter is fitted.

More precise measurements and more detailed modeling are possible with combining different probes and/or stacking clusters.

VI. Galaxy - Galaxy Lensing

The weak lensing signal around a galaxy is much weaker than that around a galaxy cluster ($S/N \propto \theta_E = \left(\frac{4GM}{c^2}\frac{D_{ds}}{D_d D_s}\right)^{\frac{1}{2}}$).
The signal of many galaxies has to be superposed statistically. On average, images of background galaxies tend to orient in the direction tangent to the line connecting the lens-soure galaxy pair,



$$p(\phi) = \frac{2}{\pi}\left[1 - \gamma_t \left\langle \frac{1}{|\epsilon_i|} \right\rangle \cos(2\phi)\right].$$

(Brainerd et al. 1996)

The amplitude of the cos-wave gives γ_t ; ~ the average strength of the shear.

Galaxy-galaxy lensing is divided into 3 regiemes according to the contributing mass to the signal:

- small θ — γ_t dominated by the mass of the galaxy
- intemediate θ — mass distribution around the lensing galaxy starts to contribute significantly
- large θ — contribution of the lens galaxy is negligible, signal generated by the LSS where the lens galaxy is embedded.

Galaxy-galaxy lensing in the small θ regieme is used to confine the density profile of stacked galaxies, that in the large θ regieme is used to infer the linear galaxy bias.

Galaxy-galaxy lensing is less sensitive to large scale PSF ~~problems~~, e.g. a uniform shear across the observation field does not change galaxy-galaxy lensing results.