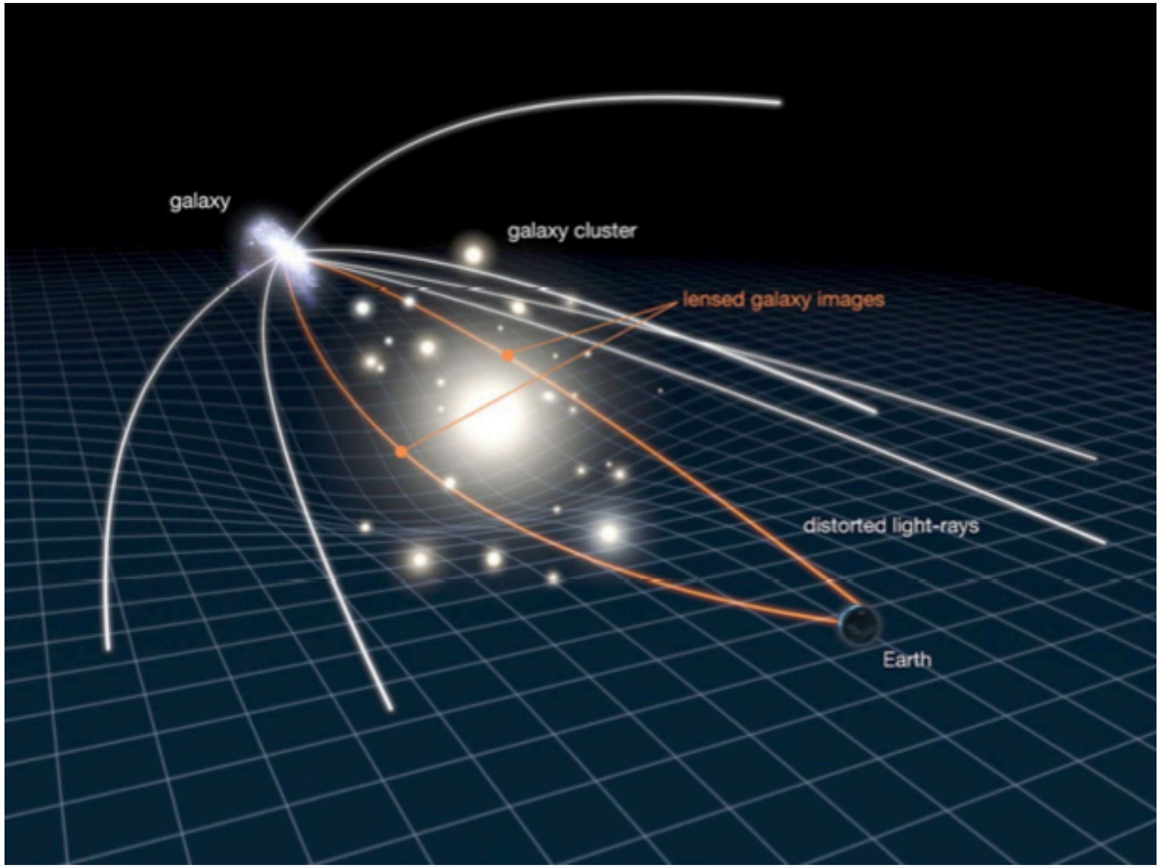
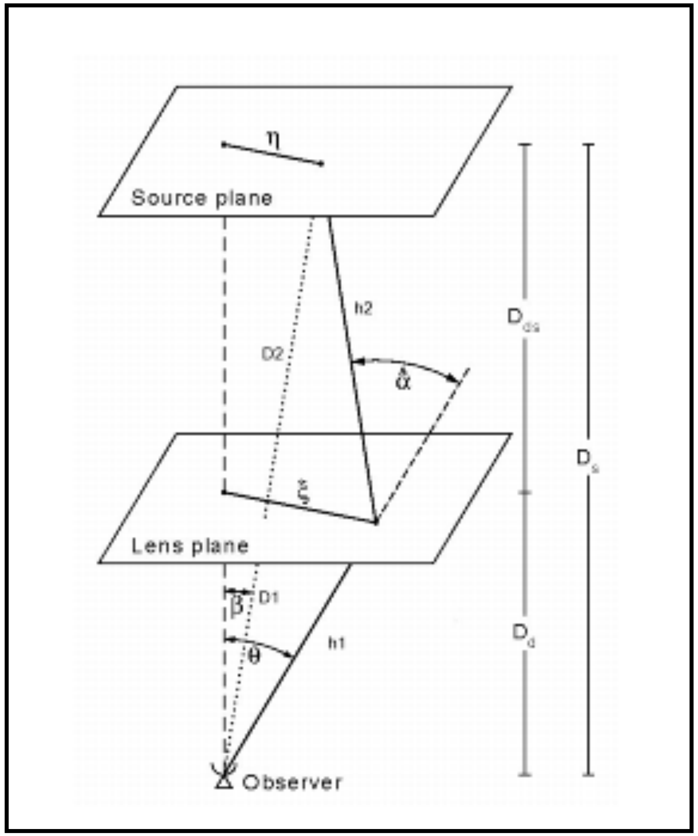
Gravitational Lensing

Deflection of light

The basis of gravitational lensing is the effect of deflection of light caused by gravity.



In general, we define a **source - lens - observer system**



source position in the source plane

deflection angle

impact parameter in the lens plane

image position in the image plane

optical axis

Light from a point emitted at an angular position β is observed at a different angular position θ.

It is deflected by a deflection vector α induced by gravity.

The lens equation, relating source and lens planes can be found from the diagram above, by using simple trigonometry (vector addition on the source plane):

$$\boxed{D_s\vec{\theta} = D_s\vec{\beta} + 2D_{ds}\frac{\hat{\vec{\alpha}}}{2}}$$

α is determined by the properties of the lens : it contains the physical (gravitational field) information we want to find out.

Measuring the change between θ and β we can find α if we know the distances (there is a degeneracy with the distance).

How does the deflection angle relate to the lens gravitational potential?

Let us consider **light propagation from source to observer** in the Universe described by the Robertson-Walker metric with a small inhomogeneity representing the lensing potential:

$$ds^2 = -\left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 + \left(1 - \frac{2\Phi}{c^2}\right) \left[dx_1^2 + dx_2^2 + dx_3^2\right]$$

The deflection may be derived using the principle of Fermat: **light follows a path of extremal time.**

Light follows null geodesics, and setting $ds^2 = 0$ we can immediately write the speed of light when travelling in the gravitational field of the lens.

It is:

$$c' = \frac{|\mathrm{d}\vec{x}|}{\mathrm{d}t} = c\sqrt{\frac{1 + \frac{2\Phi}{c^2}}{1 - \frac{2\Phi}{c^2}}} \approx c\left(1 + \frac{2\Phi}{c^2}\right)$$

We can think of the gravitational field as a “change of medium” since it effectively changes the speed of light propagation.

This medium is thus associated to an effective index of refraction, given by:

$$n = c/c' = \frac{1}{1+\frac{2\Phi}{c^2}} \approx 1 - \frac{2\Phi}{c^2}$$

In terms of properties of light propagation, the perturbed metric is like a medium where the speed of light is v < c → it bends the light with respect to the homogeneous spacetime where v = c.

Now, let x(l) be a light path crossing the medium.

The light travel time is then proportional to:
(since the refraction index is basically dt/dx)

$$\int_A^B n[\vec{x}(l)]\mathrm{d}l$$

and we want to find the path of extremal (minimum) time, i.e.,

$$\delta \int_A^B n[\vec{x}(l)]\mathrm{d}l = 0$$

This is a standard **variational problem**, that as we know will lead to the Euler-Lagrange equations.

The extremal light path verifies:

$$\delta \int_A^B n(\vec{x}) dx = 0 = \delta \int_{\lambda_A}^{\lambda_B} n(\vec{x}(\lambda)) \frac{dx}{d\lambda} d\lambda = \delta \int_{\lambda_A}^{\lambda_B} n(\vec{x}(\lambda)) |\dot{\vec{x}}| d\lambda = \delta \int_{\lambda_A}^{\lambda_B} L(x, \dot{x}; \lambda) d\lambda,$$

where λ is an arbitrary affine parameter, labeling the positions along the path,

and we found out that $n[\vec{x}(\lambda)] \left| \frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda} \right|$ role of a Lagrangian.

Having found the Lagrangian we can now describe the light path using the Euler-Lagrange equations:

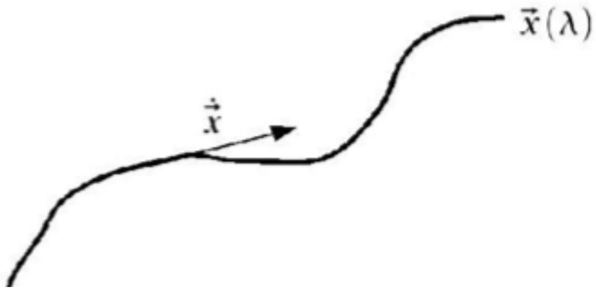
$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{\vec{x}}} - \frac{\partial L}{\partial \vec{x}} = 0$$

From our Lagrangian, we compute:

$$\frac{\partial L}{\partial \dot{\vec{x}}} = n\frac{\dot{\vec{x}}}{|\dot{\vec{x}}|}$$

$$\frac{\partial L}{\partial \vec{x}} = |\dot{\vec{x}}|\frac{\partial n}{\partial \vec{x}} = (\vec{\nabla} n)|\dot{\vec{x}}|$$

This means that the Euler-Lagrange equation is an equation for the evolution of $\dot{\vec{x}}$, which is a vector tangent to the light path.



$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = \quad \frac{d}{d\lambda}(n(\vec{x})\vec{u}_x) - \vec{\nabla} n = n\dot{\vec{u}}_x + (\vec{\nabla} n.\vec{u}_x)\vec{u}_x - \vec{\nabla} n.$$

(u is the normalised vector tangent to the path) and so the Euler-Lagrange equation is:

$$n\dot{\vec{u}}_x + (\vec{\nabla}n.\vec{u}_x)\vec{u}_x - \vec{\nabla}n = 0$$

$$\Leftrightarrow \quad \dot{\vec{u}}_x = \frac{1}{n(\vec{x})}\left(\vec{\nabla}n - (\vec{\nabla}n.\vec{u}_x)\vec{u}_x\right)$$

this is the gradient of n perpendicular to the light path

$$\Leftrightarrow \quad \dot{\vec{u}}_x = \frac{1}{n(\vec{x})}\vec{\nabla}_\perp n(\vec{x}) = (1+\frac{2\Phi}{c^2})(-\frac{2}{c^2}\vec{\nabla}_\perp\Phi) \approx -\frac{2}{c^2}\vec{\nabla}_\perp\Phi$$

and therefore, the gradient of the potential.

Now, the derivative of the tangent vector is by definition the deflection.

So **we found that the deflection is the gradient of the lens potential in the plane orthogonal to the tangent to the path (i.e. on the lens plane).**

Notice the minus sign, meaning the gradient of the potential points away from the lens centre and the deflection angle points toward the lens (light is pulled towards the lens).

The potential changes from point to point along the light path, so the **total deflection** is the integral over the ”pull” of the gravitational potential perpendicular to the light path:

$$\boxed{\vec{\alpha} = -\frac{2}{c^2}\int_{\lambda_A}^{\lambda_B} \dot{\vec{u}}_x \, d\lambda = \frac{2}{c^2}\int_{\lambda_A}^{\lambda_B} \vec{\nabla}_\perp \Phi \, d\lambda,}$$

Note that:

- The integral should be made over the actual light path
(a priori unknown before computing the deflection → so it is a recursive problem).

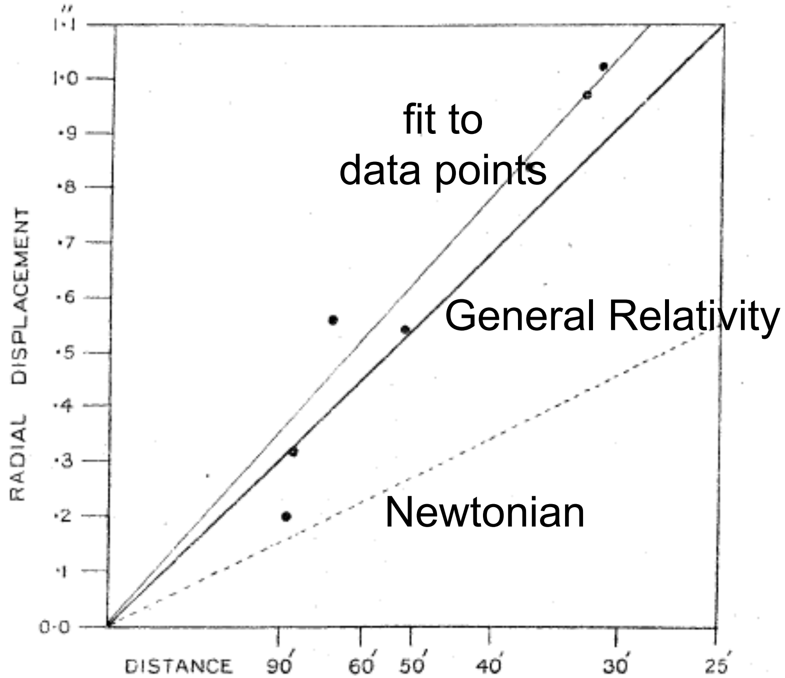
However, given the smallness of the potential $\Phi/c^2 \ll 1$, the deflection angle is usually small and in practice we can integrate over the unperturbed light path. (This is called the Born approximation, also used in scattering theory).

- Since the speed of light is effectively slowed down in the gravitational field, the travel time to cross a given length is larger than it would be in the absence of a lens. This is called the Shapiro delay.

- The value of the deflection angle computed in GR (that was we saw contains a factor of 2)

is twice the value predicted by Newtonian gravity, or by considering the equivalence principle (gravity - acceleration) in special relativity.

The well-known Eddington eclipse expedition of 1919 measured the deflection angle produced at the edge of the Sun disk with the purpose of comparing the measurement with the two predictions. **It was the first test of GR.**

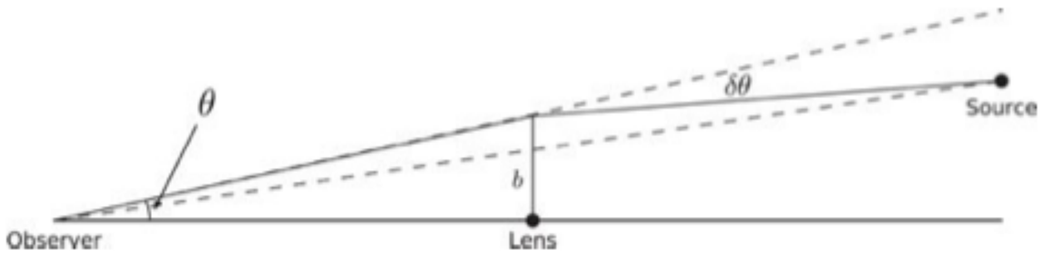


Having found the relation between deflection angle and gravitational potential, we can compute the deflection of the light emitted by a **point source** when passing by a lens.

Let us consider a point mass lens, with potential

$$\Phi = -\frac{GM}{r}$$

Light from the source travels along the z-axis towards the observer and crosses the lens plane (i.e., the plane x,y orthogonal to the z-axis), at a distance b from the point mass. b is called the impact parameter.



$$r = \sqrt{x^2 + y^2 + z^2} = \sqrt{b^2 + z^2},\ b = \sqrt{x^2 + y^2}$$

The potential on the lens plane is

$$\vec{\nabla}_\perp \phi = \begin{pmatrix} \partial_x \Phi \\ \partial_y \Phi \end{pmatrix} = \frac{GM}{r^3} \begin{pmatrix} x \\ y \end{pmatrix}$$

where $\begin{pmatrix} x \\ y \end{pmatrix} = b \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$

and the resulting deflection vector is:

$$\begin{aligned} \hat{\vec{\alpha}}(b) &= \frac{2GM}{c^2} \begin{pmatrix} x \\ y \end{pmatrix} \int_{-\infty}^{+\infty} \frac{\mathrm{d}z}{(b^2+z^2)^{3/2}} \\ &= \frac{4GM}{c^2} \begin{pmatrix} x \\ y \end{pmatrix} \left[\frac{z}{b^2(b^2+z^2)^{1/2}} \right]_0^\infty = \frac{4GM}{c^2 b} \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} \end{aligned}$$

From the x and y components of the deflection angle vector, we compute its norm, which is the well-known **result**:

$$|\hat{\vec{\alpha}}| = \frac{4GM}{c^2 b}$$

Note that the impact parameter is strongly constrained.

The source emits in all directions, and various light paths reach the lens plane.
But only one is deflected towards the observer.

From the lens equation (from the source-lens-observer diagram), we can see it is the one that passes at $b = D_d D_{ds} / D_s$

D_d = distance from observer to lens (deflector)
D_{ds} = distance from lens to source
D_s = distance from observer to source

For this reason, **all lensing systems have a fundamental degeneracy between distances and lens properties.**

We can only compute the mass of the lens if we know the distances involved in the system.

Conversely, lensing can be used as a geometric probe of the Universe (i.e., it can be used to measure cosmological distance and use them to infer the density parameters) if the mass of the lens is known.

Let us consider that the lens is not a point mass but it is an extended object (extended lens)

Since the deflection angle depends linearly on the mass M, the effect from a finite lens in a plane is just the sum of the deflection angles created from all points in the lens. If we discretize the lens as a set of N point lenses of masses M_i at positions ξ_i on the lens plane, then the deflection angle of a light ray crossing the plane at ξ will be:

$$\hat{\vec{\alpha}}(\vec{\xi}) = \sum_i \hat{\vec{\alpha}}_i(\vec{\xi} - \vec{\xi}_i) = \frac{4G}{c^2} \sum_i M_i \frac{\vec{\xi} - \vec{\xi}_i}{|\vec{\xi} - \vec{\xi}_i|^2}$$

We can also consider a lens in 3D with mass density ρ. The z extension of the lens is always just a small segment of the full source-observer light path, and it can be considered that it is in a plane - the thin-screen approximation. In this approximation, the lensing matter distribut ribed by its **surface mass density**:

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

and the total deflection is given by:

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

Gravitational Lensing

Gravitational lensing, in a strict sense, refers to the case of **extended sources**, which give rise to differential effects.

Indeed, neighbouring points in the source suffer slightly different deflections in the lens plane: it is a differential effect that makes **the image of an extended source** (i.e. non point-like) **to become distorted**.

This is easily seen if we Taylor-expand the lens equation. Remember the lens equation is a mapping from image positions to source positions (it is usually written in that order, and not as a mapping from source to image). So a given point θ in the image plane corresponds to an original position β(θ) in the source plane, related by the deflection angle:

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

(here the vectors have absorbed the distance factors present in the original lens equation) $\vec{\alpha}(\vec{\theta}) \equiv \frac{D_{\mathrm{LS}}}{D_{\mathrm{S}}} \hat{\vec{\alpha}}(\vec{\theta})$

The Taylor expansion of β(θ) to linear order is $\beta(\theta) = \beta(\theta_0) + A(\theta_0).(\theta - \theta_0)$

where A is the amplification matrix (the Jacobian) and describes the **lensing transformation between source and image planes** to first order:

$$A_{ij}(\theta) = \frac{\partial \beta_i}{\partial \theta_j} = \left(\delta_{ij} - \frac{\partial \alpha_i}{\partial \theta_j} \right)$$

it is a 2D matrix, since β (position in the source plane and θ (position in the lens plane) are 2D vectors.

Now, remember that a general matrix can be decomposed in 3 parts:

(traceless) symmetric + (traceless) antisymmetric + diagonal

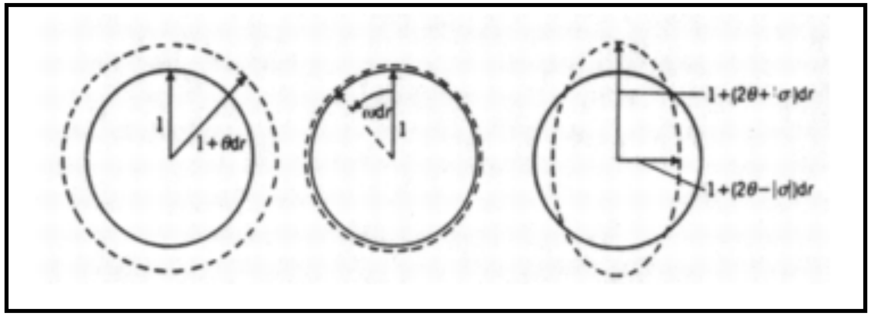
$$\begin{bmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{bmatrix} + \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} + \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

Applying a diagonal matrix to an image will expand it (or contract it) radially in an isotropic way → k is called **convergence.**

Applying an antisymmetric matrix to an image will rotate it → ω is called **rotation.**

Applying an symmetric matrix to an image will distort it in an anisotropic way, contracting in one dimension and expanding in the other → γ is called **shear.**

This means that any linear distortion of an image is a combination of convergence/expansion, rotation and shear



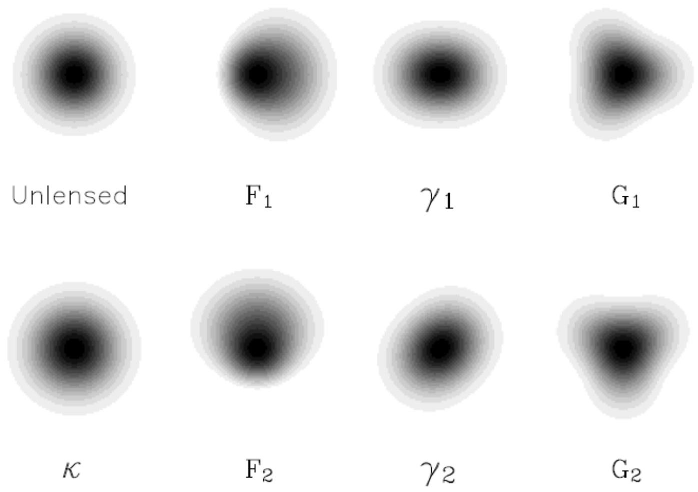
The amplification matrix is then written as

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

Note that usually actual lensing distortions do not includes rotation because the gravitational field is a gradient field (completely defined by a potential), and so its rotational is zero (it is a so-called E field) and the deflection vector field does not produces rotations.

The presence of rotations in a lensed image (due to so-called B-modes) is an indication of systematic effects, i.e., distortion effects with non-lensing origin.

The distortions applied to a circular image result in:



isotropic distortion (k, convergence) → a circle expands/contracts (full rotational symmetry)

anisotropic distortion (γ, shear) → a circle transforms into a π-rotational symmetric shape (an ellipse)

second-order distortions (by continuing the Taylor expansion) (F, G, flexion) → a circle transforms into a 120º-rotational symmetric shape (a banana-shape F or a "Mercedes logo" G)

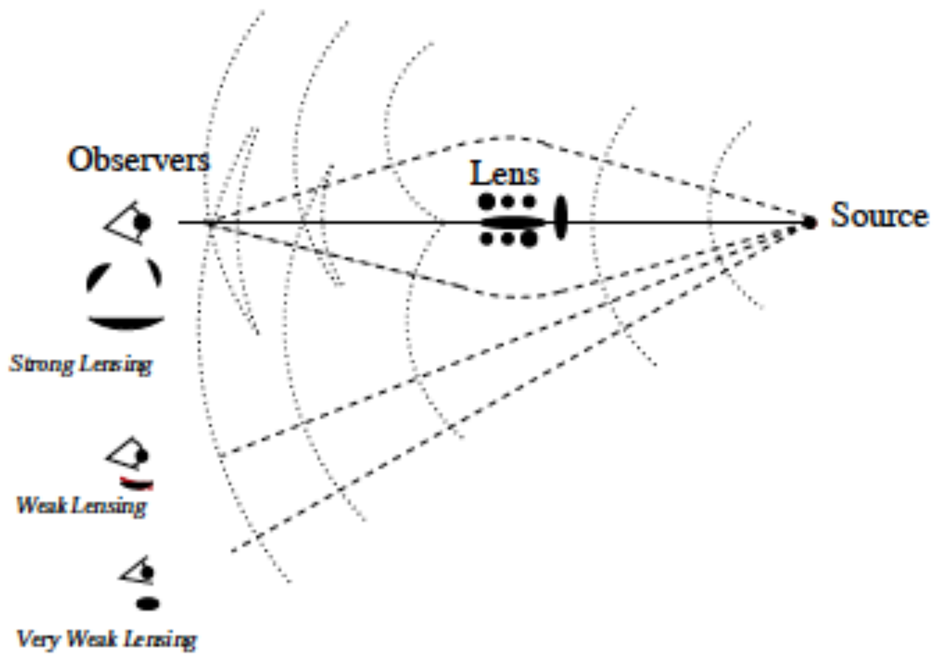
These are the fundamental distortions (also called the optical scalars) and contain the dependence on α → which contains the information on gravity

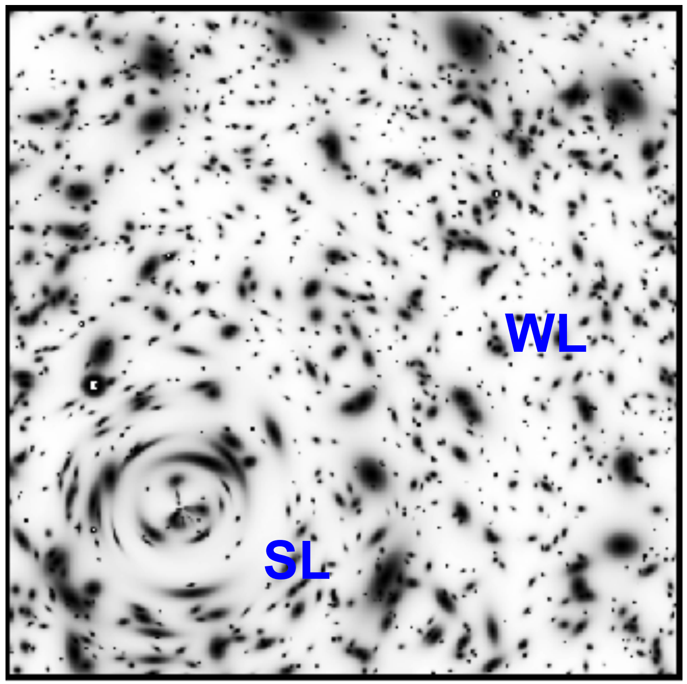
The determinant of the amplification matrix defines the magnification:

$$\mu = \frac{1}{det\ A} = \frac{1}{(1-\kappa)^2 - \gamma^2}$$

The magnification, and the amplitude of the optical scalars - which are fields in the 2D sky - define the **gravitational lensing regime** that occur in the positions of the sky.

There are two general regimes - **weak lensing** and **strong lensing** - that occur in regions of the image plane where the values of the $k(\theta)$ and γ (θ) fields are small ($<<1$) (weak lensing) or large (strong lensing).



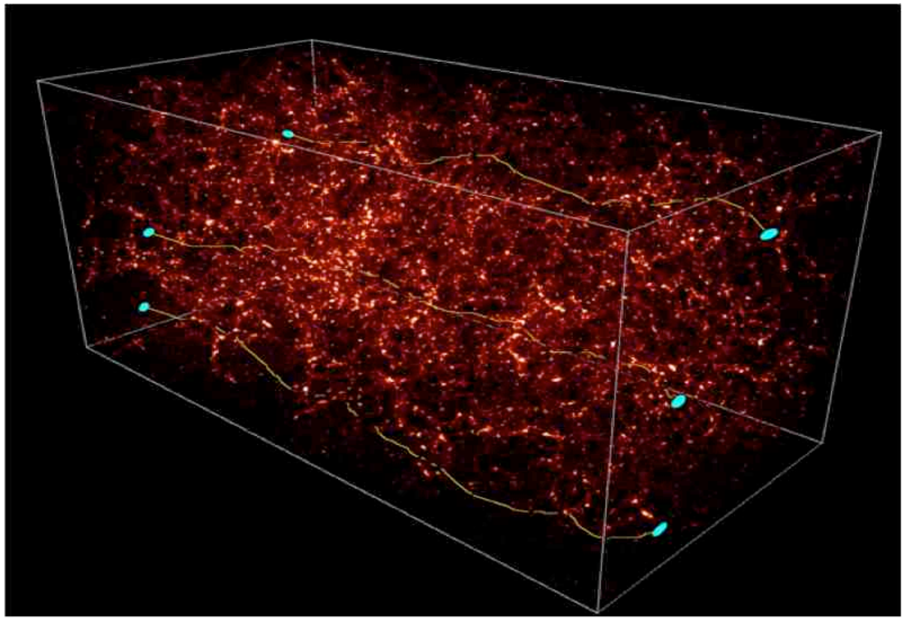


The observable effects are very different in the two regimes.

Weak Lensing occurs at larger separations from the source-lens-observer line (the **line-of-sight**), or with lenses of low density contrast.

The effects are: small increase of ellipticity of the source galaxy (shear), alignment of images.

Weak lensing is a very useful probe in a cosmological system where the lens is the large-scale structure of dark matter distribution. In this case the shear is so small that it cannot be detected in individual galaxies. What can be detected is a correlation of those ellipticities because their orientations get some degree of alignment and cease being randomly oriented → this effect is used to probe the structure formation of the Universe.



Increased ellipticities: weak lensing of galaxies by the large scale structure of the Universe

Strong Lensing occurs near the line of sight, with lenses of high density contrast.

The effects are: very strong distortions (giant arcs), multiple images, flux magnification. They occur near lines where det A = 0 (infinite magnification), which are called critical lines of the image plane (the observed sky), and map back to the source plane to lines known as caustic lines.

Example:
Spherical lens

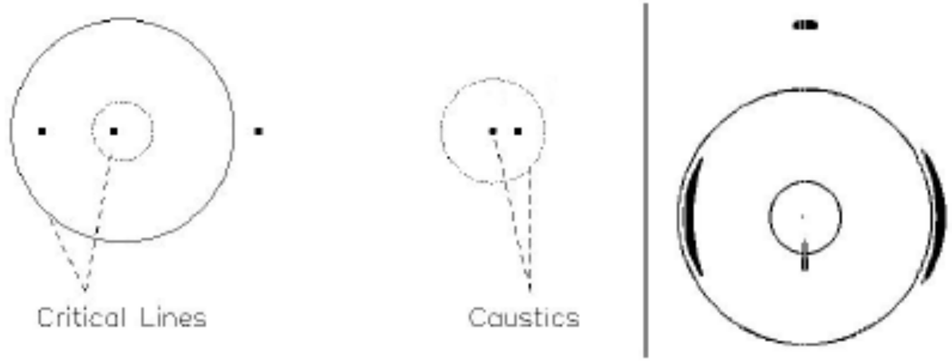


image | point source | image | extended source

Example:
Elliptical lens

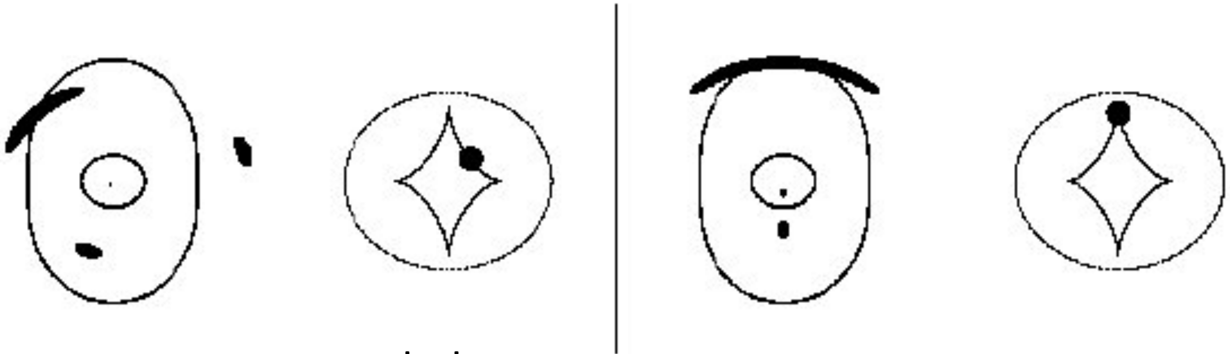
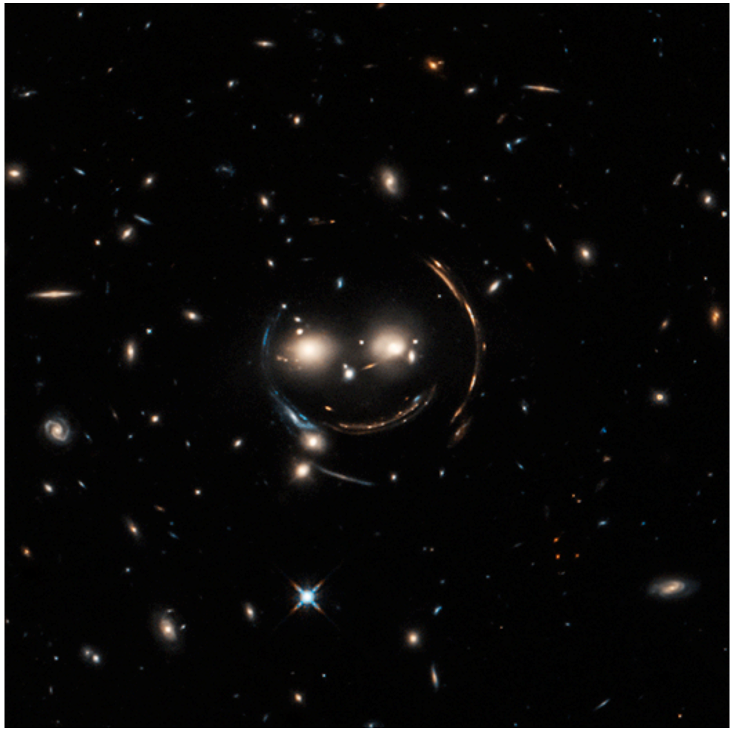


image | extended source | image | extended source

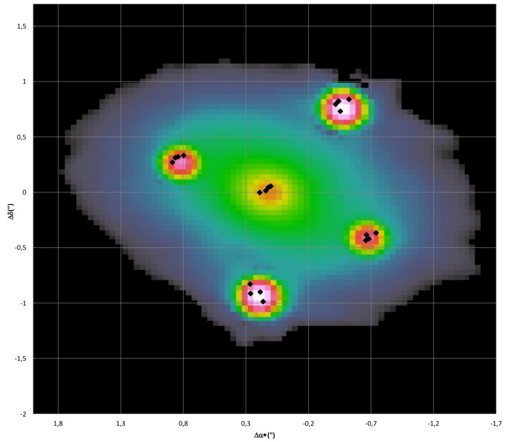
Actual observations of strong lensing:

Giant Arcs: Strong lensing of galaxies by a cluster



Giant Arcs: Strong lensing and Einstein ring of galaxies by a group that includes two massive ellipticals (The Cheshire Cat)

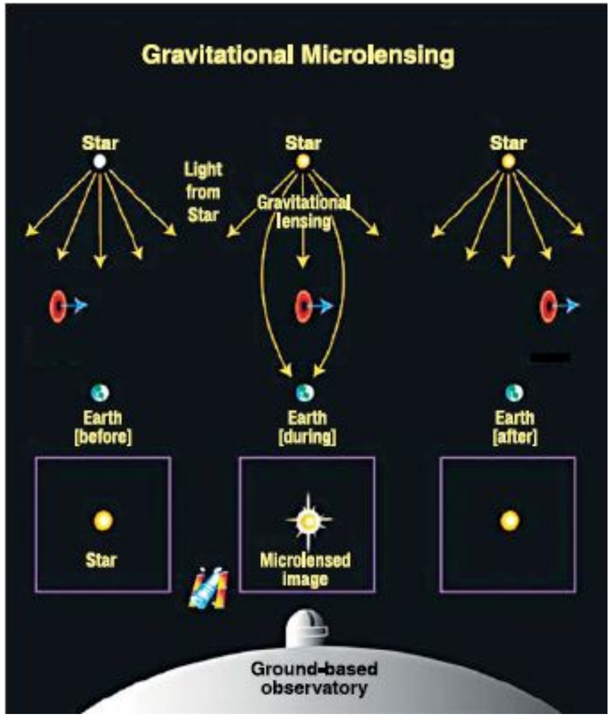
Einstein ring: Strong lensing of a galaxy by a galaxy, an infinite number of multiple images forms on a circle

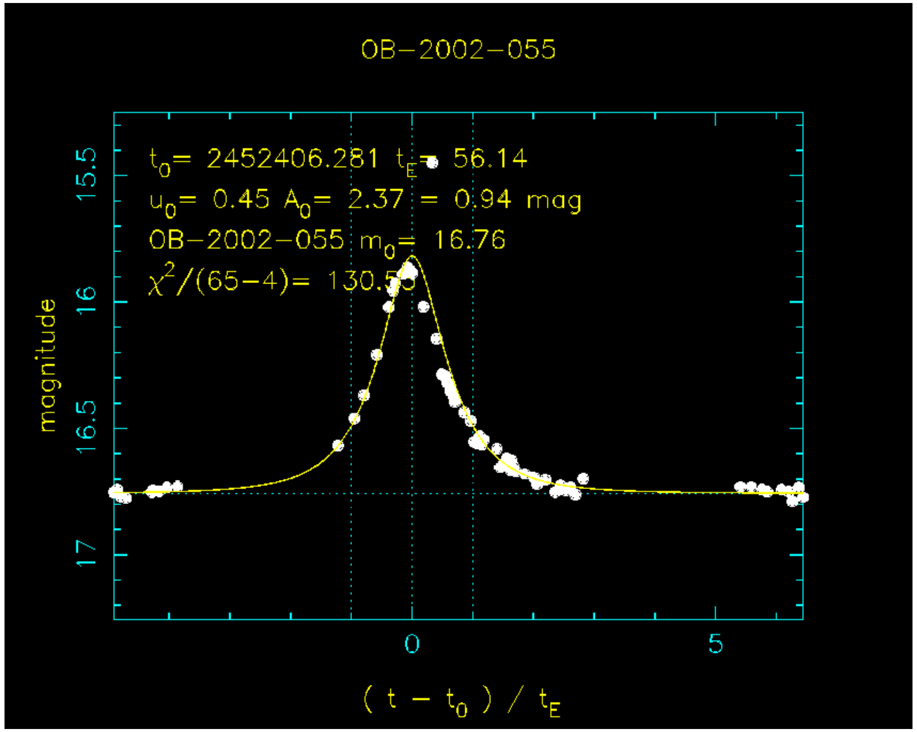


Einstein cross: Strong lensing of a quasar by a galaxy, forming a quadruple image of the quasar

When the angular scale of the strong lensing effects is small (ex: multiple images have small angular separation and are not resolved):

the strong lensing is called **microlensing.**





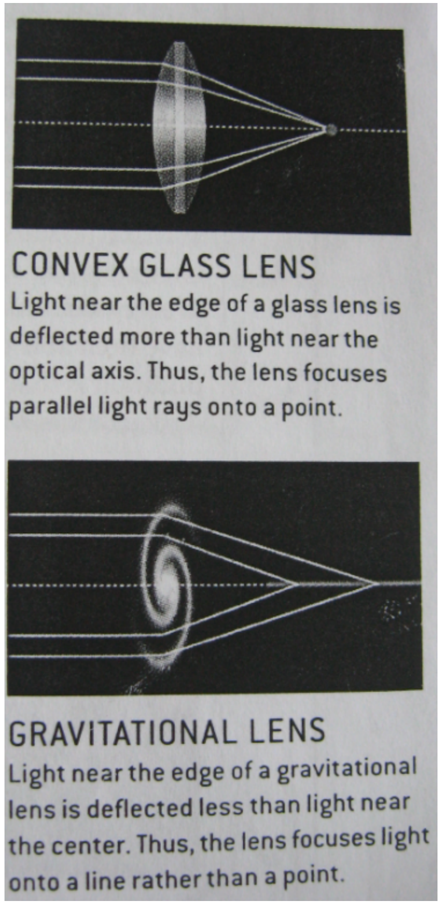
Increase of flux: → Microlensing of a star by a planet (used to detect exoplanets).

In summary, gravitational Lensing has a number of fundamental properties:

- it depends on the projected 2d mass density distribution of the lens
- it is independent of the luminosity of the lens
- it does not have a focal point
- it is achromatic, there is no frequency shift from source to image
- it involves no emission or absorption of photons
- it conserves the surface brightness

that lead to a number of observable features:

- change of apparent **positions**
- magnification (increase of size), which combined to the conservation of brightness implies an increase of **flux** → natural telescope
- **distortion** of extended sources (ellipticities, tangential giant arcs, radial arclets)
- **multiple images**
- **time-delay** between multiple images



These observables (positions, fluxes, distortions) can be used **to estimate the total mass and mass distribution of the lens**. For example:

- in (strong or weak) cluster lensing → mass distribution of the cluster
- in LSS weak lensing (cosmic shear) → dark matter power spectrum

In all systems, the general recipe to estimate the physical properties (or cosmological parameters) is:

i) (theoretical) ***define a lens model and derive its gravitational potential.***

For example the potential of a mass distribution, or the potential of a cosmological model

ii) (theoretical) ***derive the deflection and optical scalar fields from the gravitational potential***

From the definitions in the amplification matrix, it is clear that shear and convergence are derivatives of the deflection field, and second-order derivatives of the potential:

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

convergence $\kappa = \frac{1}{2}(\psi_{,11} + \psi_{,22}).$

where ψ is the gravitational potential projected on the lens plane (i.e. integrated along z) and dimensionless (with the distance factors included), i.e.,

$$\Psi = \frac{D_{\mathsf{L}}^2}{\xi_0^2}\frac{D_{\mathsf{LS}}}{D_{\mathsf{L}}D_{\mathsf{S}}}\frac{2}{c^2}\int \Phi(D_{\mathsf{L}}\vec{\theta}, z)\mathrm{d}z$$

this is called the **lensing potential**.

Note that indeed:

$$\begin{aligned}
\vec{\nabla}_x \Psi(\vec{x}) &= \xi_0 \vec{\nabla}_\perp \left(\frac{D_{\mathsf{LS}} D_{\mathsf{L}}}{\xi_0^2 D_{\mathsf{S}}} \frac{2}{c^2} \int \Phi(\vec{x}, z)\mathrm{d}z \right) \\
&= \frac{D_{\mathsf{LS}} D_{\mathsf{L}}}{\xi_0 D_{\mathsf{S}}} \frac{2}{c^2} \int \vec{\nabla}_\perp \Phi(\vec{x}, z)\mathrm{d}z \\
&= \vec{\alpha}(\vec{x})
\end{aligned}$$

Note also that the convergence is the Laplacian of the lensing potential.
This means, from Poisson equation, that the convergence is a (projected) mass.
In particular, it is the (dimensionless) surface density:

$$\kappa(\vec{x}) \equiv \frac{\Sigma(\vec{x})}{\Sigma_{\rm cr}} \quad \text{with} \quad \Sigma_{\rm cr} = \frac{c^2}{4\pi G} \frac{D_{\rm S}}{D_{\rm L} D_{\rm LS}}$$

iii) (theoretical) ***predict the observables from the optical scalars fields***
(shear, image positions, fluxes)

iv) (observational) ***measure the observables in astrophysical images***

v) (statistical***) estimate the lens model parameters by fitting the theoretical predictions to the data***

Example: estimate the mass of a galaxy cluster (lens)

We need to build a complex model that takes into account different components of mass distribution: dark matter halo, gas, galaxy distribution,

and need to define a spatial distribution of background galaxies (sources)

and then predict the distortions, positions and fluxes on the image plane of source background galaxies.

Let us consider that the cluster only has one matter component: the dark matter halo (a NFW density profile):

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}$$ (with 2 free parameters)

The 2D **surface mass density** can be computed from the 3D density profile, and it is:

$$\Sigma(x) = \frac{2\rho_s r_s}{x^2 - 1} f(x) \ ,$$

with

$$f(x) = \begin{cases} 1 - \frac{2}{\sqrt{x^2-1}} \operatorname{arctan}\sqrt{\frac{x-1}{x+1}} & (x > 1) \\ 1 - \frac{2}{\sqrt{1-x^2}} \operatorname{arctanh}\sqrt{\frac{1-x}{1+x}} & (x < 1) \\ 0 & (x = 1) \end{cases}$$

and so the **convergence** is

$$\kappa(x) = \frac{\Sigma(\xi_0 x)}{\Sigma_{cr}} = 2\kappa_s \frac{f(x)}{x^2 - 1} \quad \text{with} \quad \kappa_s \equiv \rho_s r_s \Sigma_{\rm cr}^{-1}$$

from which we can obtain the **mass**,

$$m(x) = 2\int_0^x \kappa(x')x'dx' = 4k_s h(x)$$

with

$$h(x) = \ln\frac{x}{2} + \begin{cases} \frac{2}{\sqrt{x^2-1}}\,\mathrm{arctan}\sqrt{\frac{x-1}{x+1}} & (x > 1) \\ \frac{2}{\sqrt{1-x^2}}\,\mathrm{arctanh}\sqrt{\frac{1-x}{1+x}} & (x < 1) \\ 1 & (x = 1) \end{cases}$$

We can also compute the **lensing potential**, which is,

$$\Psi(x) = 4\kappa_s g(x)\,,$$

where

$$g(x) = \frac{1}{2}\ln^2\frac{x}{2} + \begin{cases} 2\,\mathrm{arctan}^2\sqrt{\frac{x-1}{x+1}} & (x > 1) \\ -2\,\mathrm{arctanh}^2\sqrt{\frac{1-x}{1+x}} & (x < 1) \\ 0 & (x = 1) \end{cases}$$

and the **deflection angle**, which is $\alpha(x) = \frac{4\kappa_s}{x}h(x)$

From this, we can for example predict the image positions of source galaxies, fit to the observed positions and constrain the two parameters r_s and ρ_s needed to determine the value of the cluster mass.