Weeks 11-12: Gravitational Lensing

April 17, 2017

1 The Basics

According to general relativity, the trajectories of light rays can be bent by gravitational fields. Massive objects can therefore gravitationally "lens" the images of objects behind them. A cosmological examples of gravitational lensing is lensing of galaxies or quasars at redshifts $z \sim 1$ by elliptical galaxies or galaxy clusters along the line of sight. If the lensing is "strong," then there will be multiple images of the source. If the lensing is "weak, " then there will be no extra images, but the images we see will be distorted by lensing. Weak lensing is now a common and powerful tool to measure the mass distribution in galaxy clusters. There is also "gravitational microlensing" of stars in the Galactic bulge or LMC by other stars along the line of sight. In gravitational microlensing, there are multiple images, but they are so close together that they cannot be resolved. Instead, the combined light from the two unresolved images makes the source star appear brighter than it would otherwise. There is also "cosmic shear," weak gravitational lensing of galaxies at $z \sim 1$ by the large-scale distribution of mass in the Universe. Large-scale correlations in the mass induce large-scale correlations to distant-galaxy images.

The description of gravitational lensing is relatively simple.

Let us choose the origin of our coordinates on the sky to be at the position of the lens. We will call the true position of a source behind the lens $\vec{\theta}_s$. The lens is at a distance D_L from us, and the source is at a distance D_S. The distance between the source and the lens is D_{LS}. For now, consider lensing only in Minkowski space (i.e., neglect cosmological effects); we will clarify the meaning of these distances in the cosmological context below. Light rays from the source are bent by a "deflection angle,"

$$\vec{\alpha} = \frac{2}{c^2} \int \vec{a}_\perp \, dl, \tag{1}$$

where the integral is along the line to a source, and $a_\perp$ is the gravitational acceleration perpendicular to the line of sight at distance l. As a result of this bending of light rays, the source appears at a position $\vec{\theta}_I$ ("I" for image). In the astrophysical context, the size of the lens R is $R \ll D$, so all the light deflection takes place at the distance of the lens. We assume, throughout, that all deflection angles are small: $\alpha \ll 1$ and $|\vec{\theta}_I - \vec{\theta}_S| \ll 1$.

The relation between the deflection angle and the "displacement angle" $\vec{\theta}_I - \vec{\theta}_S$ is the "lens equation,"

$$\vec{\alpha}(D_L\theta_I) = \frac{D_S}{D_{LS}}(\vec{\theta}_I - \vec{\theta}_S). \tag{2}$$

If we are theorists, then we assume we know the positions $\vec{\theta}_S$ of the sources, and we then solve the lens equation to get the positions $\vec{\theta}_I$ of the images. If we are observers, we have the image positions, and (in principle) we solve the lens equation to tell us where the sources are. What usually happens, even if we are observers, is that we assume source positions and then solve the lens equation to obtain $\vec{\theta}_I(\vec{\theta}_S)$.

The solution of the lens equation provides a mapping from the "source plane," parameterized by $\vec{\theta}_S$ to the "image plane," parameterized by $\vec{\theta}_I$. The mathematics that describes this process is known as "catastrophe theory", the theory of mapping of a two-dimensional surface into another two-dimensional surface. If the mapping $\vec{\theta}_S \to \vec{\theta}_I$ is one-to-one (which usually happens if α is sufficiently small), then we have weak gravitational lensing; i.e., no multiple images. If the function $\theta_I(\vec{\theta}_S)$ is multiple-valued, then there will be multiple images of the source. This is called "strong lensing," and it usually occurs if α is sufficiently large.

With either weak or strong lensing, gravitational lensing may distort the image shapes. However, lensing always conserves surface brightness because the intensity I_ν/ν^3 is conserved. What this means is that if lensing distorts a galaxy image, then the brightness of the source will be amplified in proportion to the increase in the size of the image. Remember that the function $\vec{\theta}_I(\vec{\theta}_S)$ maps every point in the "source plane" into an "image plane." The Jacobian of this transformation,

$$A = \left| \frac{\partial(\vec{\theta}_I)}{\partial(\vec{\theta}_S)} \right|, \tag{3}$$

is therefore the increase in the size of the image and therefore the amplification of the image. According to catastrophe theory, an image may have $A > 1$ or $A < 1$, but you always have at least one image with $A > 1$. In particular, for weak lensing, $A > 1$ always.

Note also that we are assuming the "geometrical optics" limit: i.e., we are assuming that the wavelength λ of light is small compared with the Schwarzchild radius GM/c^2.

Let's now think more carefully about the distances D_L, D_{LS}, and D_S. To derive the lens equation, we simply assumed that physical sizes are an angle times a distance. It is therefore clear that in the cosmological context, the distances we should use are angular-diameter distances. Care should also be taken to note that D_{LS} is the angular-diameter distance of the source as measured by an observer at the lens.

We will now see that the quantity describing the source relevant for lensing is the surface mass density, the mass density projected along the line of sight. The deflection angle is

$$\vec{\alpha} = \frac{2}{c^2}\int \vec{\nabla}_\perp \Phi \, dl = \vec{\nabla}_\perp \frac{2}{c^2}\int \Phi \, dl, \tag{4}$$

where Φ is the gravitational potential. Therefore, the lens equation can be re-written,

$$\vec{\theta}_I - \vec{\theta}_S = \frac{D_{LS}}{D_S}\vec{\alpha} \equiv \vec{\nabla}_\theta \psi(\vec{\theta}_I), \tag{5}$$

where $\psi(\vec{\theta}_I)$ is the projected potential, a function of position in the image plane. Since the gravitational potential is related to the mass density ρ through the Poisson equation, $\nabla^2\Phi = 4\pi G\rho$, it follows that

$$\nabla^2_\theta \psi = \frac{D_L D_{LS}}{D_S}\frac{8\pi G}{c^2}\Sigma \equiv 2\frac{\Sigma}{\Sigma_c}, \tag{6}$$

where $\Sigma = \int \rho\, dl$ is the projected mass density, or "surface mass density," and Σ_c is the "critical surface mass density,"

$$\Sigma_c \equiv \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}}. \tag{7}$$

Finally, the two-dimensional Poisson-like equation, $\nabla^2\psi = 2\Sigma/\Sigma_c$ can be written in integral form as

$$\psi(\vec{\theta}) = \frac{1}{\pi\Sigma_c}\int \Sigma(\vec{\theta}')\,\ln|\vec{\theta}-\vec{\theta}'|\,d^2\theta. \tag{8}$$

2 Simple lens model

Let's now consider a simple lens model. We'll take the simplest case, which is that for a circularly symmetric lens mass distribution. In that case, the deflection angle α is a function of "impact parameter" $b = D_L\theta_I$, the point of closest approach of the *unperturbed* light trajectory to the lens center of mass. In particular, for a circular lens,

$$\alpha = \frac{4G}{c^2}\frac{M(<b)}{b}, \tag{9}$$

where $M(<b)$ is the mass seen *in projection* within a distance b of the source. The simplest example is, of course, a point mass, which has a deflection angle $\alpha = 4GM/c^2 b$. Another example is the singular isothermal sphere (SIS) which has a density profile $\rho(r) = \sigma^2/(2\pi G r^2)$ as a function of radius r, where σ^2 is the velocity dispersion. Recall that the SIS has a flat rotation curve, making it a workable lowest-order description of a galaxy. The surface mass density is $\Sigma = \sigma^2/(2Gr)$, and the deflection angle is $\alpha = 4\pi\sigma^2/c^2$; it is independent of impact parameter.

It is easy to see that with the generic circular lens, the lens equation predicts that there should be three images. To do so, we look at the lens equation, which says that the displacement $|\vec{\theta}_I - \vec{\theta}_S| \propto \alpha(\vec{\theta}_I)$, the deflection angle. For a circular lens, the symmetry dictates that the images will be situated along the line on the sky connecting the source and with the lens center. Then, we can neglect the vector character of the positions and consider θ_I and θ_S to be aligned distances from the center of mass. We then plot the right and left-hand sides and see that the two curves will intersect three times. In general, the source is displaced from the lens, so $\theta_S \neq 0$. Then, $\theta_S - \theta_I$ is a positively-sloped line that has a zero at $\theta_I = \theta_S$. The function $\alpha(\theta_I)$ will usually fall to zero at the origin, as the enclosed mass $M(<b) \to 0$ as $b \to 0$ (unless the source has a density cusp $\rho \propto 1/r^2$ which would yield $\Sigma \propto 1/b$). Then, if the source has a finite extent, then α will then recede to zero (as $\alpha \propto M/\theta_I$) at large θ_I. There can then be three crossings of these two curves. There is an image A on the same side of the lens as the source, but at larger radius, one (B) on the other side, close to the lens center, and another (C) at larger distance on the other side. The C image is parity-reversed.

If we put the source right behind the lens—i.e., as $\theta_S \to 0$, the source B becomes an "Einstein ring" at a radius θ_I that satisfies $\alpha(\theta_I) = \theta_I$. For a point-mass lens, this "Einstein radius" is

$$\theta_E = \left(\frac{4GM}{c^2}\frac{D_{LS}}{D_L D_S}\right)^{1/2} = \left(\frac{M}{10^{11.09}\,M_\odot}\right)^{1/2}\left(\frac{D_L D_S/D_{LS}}{\mathrm{Gpc}}\right)^{-1/2} \mathrm{arcsec}, \tag{10}$$

and for the SIS, it is

$$\theta_E = \left(\frac{4\pi\sigma^2}{c^2}\right)\frac{D_{LS}}{D_S} = \left(\frac{\sigma}{186\,\mathrm{km/sec}}\right)^2 \frac{D_{LS}}{D_S}\,\mathrm{arcsec}. \tag{11}$$

The Einstein radius plays a crucial role in gravitational lensing, as it demarcates strong from weak lensing. If a source falls within the Einstein radius, then it is strongly lensed, and there will be multiple images. If the source falls outside the Einstein radius, then it will be weakly lensed. (Strictly speaking, there may be a second image right behind the lens, but it is extremely faint and therefore just academic.)

For Galactic gravitational microlensing by point sources, the characteristic separations are (using $M \sim M_\odot$ and $D \sim 10$ kpc) milli-arcseconds apart and therefore cannot be resolved with optical telescopes. We can, however, see the amplification in the brightness of the source as the lens passes in front of it. The total amplification (for both sources when the source is multiply lensed) due to a point-mass lens is

$$A = \frac{1 + x^2/2}{x\sqrt{1 + x^2/4}}, \qquad \text{where} \quad x \equiv \frac{\theta_S}{\theta_E}. \tag{12}$$

We can also define for the point-mass lens a lensing cross section,

$$\sigma(>A) = \left(\frac{4\pi\sigma^2}{c^2}\right)^2\left(\frac{D_L D_{LS}}{D_S}\right)^2 \begin{cases} \frac{4\pi}{A^2} & A > 2, \\ \frac{\pi}{(A-1)^2}, & A < 2. \end{cases} \tag{13}$$

Note that the lenses are most effective (i.e., the lensing cross section is maximized) when $D_L \simeq D_S/2$.

In the cosmological context, most lenses are extended sources (e.g., elliptical galaxies, galaxy clusters, or even larger-scale inhomogeneities). And in most cases, they do not have circular symmetry. It therefore behooves us to study the more general properties of lenses. Recalling that the displacement angle is $\vec{\theta}_I - \vec{\theta}_S = \vec{\nabla}_\theta \psi$, we can write

$$\frac{\partial \theta_{I,,i}}{\partial \theta_{S,j}} = \delta_{ij} + \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j} \equiv \left(\frac{\partial(\vec{\theta}_I)}{\partial(\vec{\theta}_S)}\right)_{ij}. \tag{14}$$

Since this is a symmetric 2×2 matrix, by rotating to align with the principal axes, the matrix can be diagonalized and written in the form

$$\frac{\partial \theta_{I,,i}}{\partial \theta_{S,j}} = \begin{pmatrix} 1 + \gamma - \kappa & 0 \\ 0 & 1 - \gamma - \kappa \end{pmatrix}, \tag{15}$$

where κ, the trace, is the *convergence*,

$$\kappa = \frac{4\pi G}{c^2}\int \frac{D_L D_{LS}}{D_S}\rho\, dl = \frac{\Sigma}{\Sigma_c}. \tag{16}$$

and γ, the *shear*, is the trace-free part. The convergence measures how much brighter the source becomes, and the shear measures the distortion to the shape (e.g., if the source is round, it can become elliptical after lensing). The curves in the image plane where $\kappa = 1$ are known as caustics; they separate regions of strong and weak lensing. They occur when the surface mass density exceeds the critical density,

$$\Sigma_c = \frac{c^2 D_S}{4\pi G D_L D_{LS}} \simeq 0.35 \frac{(D_S/\mathrm{Gpc})}{(D_L/\mathrm{Gpc})(D_{LS}/\mathrm{Gpc})} \mathrm{~g~cm}^{-2}. \tag{17}$$

This surface density is reached in galaxy clusters, and sometimes in the inner parts of galaxies.

Weak gravitational lensing. Away from the center of a galaxy cluster, the surface density will be $\Sigma < \Sigma_c$, and we will be in the weak-lensing regime. Let us define the *magnification tensor* $\psi_{ij} \equiv \partial^2\psi/\partial\theta_i\partial\theta_j$, so $\partial\theta_{I,i}/\partial\theta_{S,j} \equiv J_{ij} = \delta_{ij} + \psi_{ij}$. As we said before, the change in the image area, and therefore brightness, is the trace of J_{ij}: $A \simeq 1 + \psi_{11} + \psi_{22} = 1 + \nabla^2\psi = 1 + 2\Sigma/\Sigma_c$. Since sources are being amplified, lensing will decrease the luminosity threshold, and the sky density of sources at each point on the sky will depend on the value of A at that point, as you will explore in a homework assignment.

Weak gravitational lensing will induce a "shear" distortion, described by the trace-free part of the magnification tensor ψ_{ij},

$$\begin{aligned} S_{ij} &\equiv J_{ij} - \delta_{ij} \mathrm{Tr}\, J \\ &= \delta_{ij}\left(1 - \frac{1}{2}\nabla^2\psi\right) + \psi_{ij} \\ &= \begin{pmatrix} 1 + \gamma\cos 2\theta & \gamma \sin 2\theta \\ \gamma \sin 2\theta & 1 - \gamma \cos 2\theta \end{pmatrix}, \end{aligned} \tag{18}$$

where the entries in the matrix simply parameterize the most general symmetric 2×2 matrix. More specifically, $\gamma\cos 2\theta = (\psi_{11} - \psi_{22})/2$ and $\gamma \sin 2\theta = \psi_{12}$. Here, γ is the elongation of a circular source, and θ is the orientation angle (relative to the $+\hat{x}$ axis) of that elongation: $\theta = 0$ implies stretching along the x axis; $\theta = \pi/2$ implies elongation along the y axis; and $\theta = \pi/4$ and $\theta = 3\pi/4$ imply elongation along axes at $45°$ with respect to the x and y axes.

If we had a source that we knew to be perfectly round, we could measure the shear by measuring the elongation of the image as determined, for example, by the quadrupole moments of the source:

$$Q'_{ij} = \int I(\vec{\theta})\theta_i\theta_j d^2\vec{\theta}. \tag{19}$$

If we align this with the principle axes, then the quadrupole tensor will be

$$Q'_{ij} = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}, \tag{20}$$

where we choose $a > b$. We then scale this quadrupole moment by one half its trace (a rotational invariant) so that its trace is unity (i.e., so that it represents only a shear distortion, not an amplification). After doing so, we have

$$Q_{ij} = \begin{pmatrix} 1+\epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix}, \tag{21}$$

where $\epsilon = (a^2 - b^2)/(a^2 + b^2)$ is the "ellipticity". Rotating back to the original axes, the ellipticity tensor is

$$Q_{ij} = \begin{pmatrix} 1 + \epsilon \cos 2\theta & \epsilon \sin 2\theta \\ \epsilon \sin \theta & 1 - \epsilon \cos 2\theta \end{pmatrix}, \tag{22}$$

where θ is the position angle of the source. People sometimes work with the ellipticity pseudo-vector,

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \frac{1}{Q_{11} + Q_{22}} \begin{pmatrix} Q_{11} - Q_{22} \\ 2Q_{12} \end{pmatrix}. \tag{23}$$

For a nearly circular image, $a \propto 1 + \gamma$ and $b \propto 1 - \gamma$, so

$$\epsilon = \frac{a^2 - b^2}{a^2 + b^2} = \frac{2\gamma}{1 + \gamma^2} \simeq 2\gamma. \tag{24}$$

Therefore,

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \psi_{11} - \psi_{22} \\ 2\psi_{12} \end{pmatrix}. \tag{25}$$

By measuring $e_1(\vec{\theta})$ and $e_2(\vec{\theta})$ as function of $\vec{\theta}$, we can therefore get the projected mass distribution $\Sigma(\vec{\theta})$ by inverting the Poisson equation.

The problem, however, is that in practice we have no sources that we know *a priori* to be perfectly round. However, we know that according to statistical isotropy, there should *on average* be no preferred orientation in a given region of the sky. Therefore, if we look at a large number of sources in a given region of the sky, we can use the mean ellipticity—i.e., $\langle e_1 \rangle = (1/N) \sum_{i=1}^{N} e_1^i$, where N is the number of sources used in the determination of the shear in that region and e_1^i the ellipticity of source i, and similarly for $\langle e_2 \rangle$—in a given region of the sky as an estimator for the shear at that point. If the sources have a characteristic or mean elongation e, then there will be a Poisson error $\sim e/\sqrt{N}$ in the determination of the shear in any given region.

3 Time delays

It is impossible to measure the travel time of photons from cosmological sources. However, some of the cosmological sources that are observed (e.g., quasars) exhibit variability, and if these sources are strongly lensed and multiply imaged, then there may be a difference in the time at which certain variable events (e.g., flares or outbursts of the source) are observed. The time of arrival of the ith image is

$$t^{(i)} = (1 + z_L) \left[\frac{D_L D_S}{2 D_{LS}} |\vec{\theta}_I^{(i)} - \vec{\theta}_S|^2 - \psi(\vec{\theta}^{(i)}) \right].$$

Here, the first term is due to the difference in the path length, and the second is the *Shapiro time delay*, the general-relativistic "slowing" of light as it passes through a gravitational-potential well. The factor $1 + z_L$ takes into account the redshifting of the time delay at the lens when observed at redshift $z = 0$. Let's consider some characteristic values: If the characteristic distance is a Gpc$\simeq 3 \times 10^9$ light-years and the typical separation is an arcsec$\simeq 5 \times 10^{-6}$, then the typical contribution of the first (the geometric) term in the time delay is roughly a month. Estimation of the second term is a bit more difficult, but it is usually of the same order of magnitude. So, for example, if the typical potential-well depth (for a cluster) is $(\sigma/c)^2$ with $\sigma \sim 1000\,\text{km/sec}$, and the size of the potential well is ∼Mpc, then the Shapiro time delay will be $\sim 10^{-1}$ year. The time delay is sought after in cosmology as it might, at least in principle, provide a way to determine the Hubble constant. The idea is that if we see multiple images of the same source, then the surface-mass distribution in the lens can be modeled. If so, then by measuring the time delay, we measure the angular-diameter distance in the prefactor. In practice, modeling the mass distribution in the lens to the requisite accuracy is difficult, but still worth a try. People generally try to find lens systems in which there are four images of a given source, as the multiple images provide multiple constraints and cross-checks to the mass model.

4 The magnification tensor

In cosmology, most lenses (e.g., galaxies or clusters) are extended, and so the point-mass approximation breaks down. In this cases, we can define a *lensing potential*,

$$\psi(\vec{\theta}) = 2\frac{d_{LS}}{d_L d_S}\int \phi(d_L\vec{\theta}, s)\, ds,$$

and then the reduced lensing angle is

$$\vec{\alpha} = \vec{\nabla}_\theta \psi = 2\frac{d_{LS}}{d_S}\int \vec{\nabla}_\perp \phi\, ds.$$

The Laplacian of the lensing potential is the *convergence*,

$$\kappa(\vec{\theta}) \equiv \frac{1}{2}\nabla^2_\theta \psi = \frac{d_L d_{LS}}{d_S}\int \nabla^2 \phi\, ds,$$

which, via the Poisson equation ($\nabla^2\phi = 4\pi G\rho$), is proportional to the surface mass density.

Now consider what lensing does to a source. It takes rays from positions $\vec{\beta}$ on the sky and maps them to positions $\vec{\theta}$. We can therefore define a tensor,

$$A_{ij} \equiv \frac{\partial \beta^i}{\partial \theta^j},$$

and since $\beta = \theta - \alpha$, we have

$$A_{ij} = \delta_{ij} - \frac{\partial \alpha^i}{\partial \theta^j} = \delta_{ij} - \psi_{ij},$$

where

$$\psi_{ij} \equiv \frac{\partial^2 \psi}{\partial \theta^i \partial \theta^j}.$$

The inverse of A_{ij},

$$M = \frac{\partial \vec{\theta}}{\partial \vec{\beta}} = A^{-1},$$

is known as the *magnification tensor*. Since the lens distorts $\vec{\beta}$ into $\vec{\theta}$, the Jacobian of the transformation, the determinant of M, tells us the magnification (the increase in area) of the source,

$$\mu = |M| = \frac{1}{|A|}.$$

Moreover, by conservation of phase-space density, the surface brightness (flux per unit solid angle) remains constant under lensing. Therefore, if the source increases in size by a factor μ, it also increases in brightness by the same factor.

The magnification tensor, which describes the effects of lensing on the source, can be decomposed into the convergence $\kappa = (1/2)(\psi_{11} + \psi_{22})$, and the shear, which distorts the shape of the shape of the source. An initially round source will, under the effects of lensing, be distorted into an ellipse of ellipticity γ and position angle φ, so $\gamma_1 = \gamma \cos(2\varphi)$ and $\gamma_2 = \gamma \sin(2\varphi)$, with $\gamma = (\gamma_1^2 + \gamma_2^2)$. These components are

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}), \qquad \text{and} \qquad \gamma_2 = \psi_{12} = \psi_{21}.$$

The magnification tensor is thus,

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

and the magnification is $\mu = [(1-\kappa)^2 - \gamma^2]^{-1}$.

5 Cosmic shear

"Cosmic shear" refers to weak gravitational lensing by large-scale cosmic mass inhomogeneities. Since the large-scale density field is described statistically in terms of correlation functions and power spectra for a fractional density perturbation $\delta(\vec{x}, t)$, the predictions for the lensing by these mass density fields will also be statistical and described, e.g., by correlation functions for shear distortions to galaxies and/or characteristic higher-order correlations in the CMB temperature and polarization. We will begin by considering the description of the lensing field and then the observational consequences for galaxy shapes and then for CMB fluctuations.

To begin, we review what we've learned so far. Given a mass distribution $\rho(\vec{x}, t)$, we construct a fractional density perturbation $\delta(\vec{x}, t) = [\rho(\vec{x}, t) - \bar{\rho}(t)] / \bar{\rho}(t)$. The Newtonian potential $\Phi(\vec{x}, t)$ is then obtained from the Poisson equation, $\nabla^2 \Phi = 4\pi G a^2 \bar{\rho}(t) \delta(\vec{x}, t)$, where $a(t)$ is the scale factor and ∇ is a comoving gradient operator. The projected gravitational potential is then

$$\psi(\vec{\theta}) = 2 \frac{d_{LS}}{d_L d_S} \int \phi(d_L \vec{\theta}, s)\, ds, \tag{26}$$

where the distances here are angular-diameter distances, and the convergence is $\kappa = (1/2)\nabla_\theta^2 \psi$. The two components of the shear are $\gamma_1 = (\psi_{11} - \psi_{22})/2$ and $\gamma_2 = \psi_{12}$. In Fourier space, the shear and projected potential are related by,

$$\tilde{\gamma}_1(\vec{\ell}) = \frac{-\ell_x^2 + \ell_y^2}{2} \tilde{\psi}(\vec{\ell}) \qquad \tilde{\gamma}_2(\vec{\ell}) = -\ell_x \ell_y \tilde{\psi}(\vec{\ell}), \tag{27}$$

where $\vec{\ell}$ is the wavenumber on a surface of sky sufficiently small to be considered flat. Moreover, $\tilde{\kappa}(\vec{\ell}) = -\ell^2 \tilde{\psi}(\vec{\ell})/2$. In terms of the density field.

If $\phi_{\vec{\ell}}$ is the angle that $\vec{\ell}$ makes with the $\hat{\theta}_x$ direction on the surface of the sky, then the Fourier-space shears can be written,

$$\tilde{\gamma}_1(\vec{\ell}) = -\frac{\ell^2\tilde{\psi}(\vec{\ell})}{2}\left[\cos^2\phi_{\vec{\ell}} - \sin^2\phi_{\vec{\ell}}\right] = -\frac{\ell^2\tilde{\psi}(\vec{\ell})}{2}\cos\left(2\phi_{\vec{\ell}}\right) \tag{28}$$

$$\tilde{\gamma}_2(\vec{\ell}) = -\frac{\ell^2\tilde{\psi}(\vec{\ell})}{2}\left[2\cos\phi_{\vec{\ell}}\sin\phi_{\vec{\ell}}\right] = -\frac{\ell^2\tilde{\psi}(\vec{\ell})}{2}\sin\left(2\phi_{\vec{\ell}}\right). \tag{29}$$

From these relations (and the discussion of the magnification tensor above, in which we saw that γ_1 and γ_2 are components of a symmetric trace-free rank-2 tensor), we infer that γ_1 and γ_2 are components of a spin-2 field, just like the CMB-polarization Stokes parameters Q and U. We can thus define a cosmic-shear E mode,

$$\tilde{E}(\vec{\ell}) = \tilde{\gamma}_1(\vec{\ell})\cos\left(2\phi_{\vec{\ell}}\right) - \tilde{\gamma}_2(\vec{\ell})\sin\left(2\phi_{\vec{\ell}}\right) = -\ell^2\tilde{\psi}(\vec{\ell})/2 = \tilde{\kappa}(\vec{\ell}), \tag{30}$$

which, as the last equality indicates, is simply related to the projected potential and turns out to be the convergence.

Combining earlier relations, we can write the convergence as an integral over the line of sight,

$$\kappa(\vec{\theta}) = \frac{1}{D_s}\nabla_\theta^2\int_0^{d_s} dD_s\frac{D_{sl}}{D_L}\phi(D_l\vec{\theta}, D_l) \tag{31}$$

$$= \int_0^{D_s} dD_l\frac{D_{sl}D_l}{D_s}\nabla^2\phi(\vec{\theta}D_l, D_l), \tag{32}$$

where we used $\nabla_\theta = D_l\nabla$ (and we replace distances d with D to avoid confusion with the differential operator d. We then write, using Poisson's equation,

$$\kappa(\vec{\theta}) = \int_0^{D_s} dD_l W(D_l, D_s)\delta(D_l\vec{\theta}, D_l), \tag{33}$$

in terms of a cosmic-shear window function,

$$W(D_l, D_s) \equiv \frac{D_{sl}D_l}{D_s}4\pi G\bar{\rho}(a)\left[a(D_l)\right]^2. \tag{34}$$

We thus see (again) that the convergence is a weighted surface mass-density along the line of sight.

The statistical properties of the convergence field are described in terms of a convergence power spectrum $C_\ell^{\kappa\kappa}$ defined by

$$\left\langle\kappa(\vec{\ell})\kappa^*(\vec{\ell}')\right\rangle = (2\pi)^2 C_\ell^{\kappa\kappa}. \tag{35}$$

In Week 4, we derived a relation between the power spectrum $P(k)$ for a three-dimensional field and a two-dimensional projection of that 3d field. Plugging int that Limber approximation allows us to derive a relation,

$$C_\ell^{\kappa\kappa} = \int_0^{D_s} dD_l\frac{W^2(D_l, D_s)}{D_L^2}P(\ell/D_l), \tag{36}$$

between the convergence power spectrum and the density-perturbation power spectrum $P(k)$.

So far, we have assumed in our definition of the window function that all of the galaxies being lensed (the source galaxies) are at a single distance D_s. More realistically, though, they are going to be distributed over some range of distances, or redshifts, with a redshift distribution dn/dz (defined so that $(dn/dz)dz$ is the fraction of the galaxies observed in the redshift interval $z \to z + dz$). This can be accommodated by replacing the window function by,

$$W(D_l) = \frac{3\Omega_m H_0^2}{2a(D_l)} \int_0^\infty dz \frac{dn}{dz} \frac{D_{sl}(z) D_l}{D_s(z)} \Theta\left(D_s(z) - D_l\right), \tag{37}$$

where the Heaviside step function Θ is included since only sources at distances $D_s > D_l$ are lensed.

Let's now think a bit about measurements and the measurement errors with which the observables can be obtained. Our structure-formation theory makes a prediction for $P(k)$ from which follows a prediction for $C_\ell^{\kappa\kappa}$. Let's think a bit about how well this quantity can be measured. There will be two sources of noise: the first is the statistical noise that arises from the randomly oriented intrinsic ellipticities, as discussed above. The second is cosmic variance, which arises from the finite number of Fourier modes we have to measure. Although we did not discuss this aspect in our discussion of the CMB, what we learn here about cosmic variance applies also to measurement of CMB power spectra (and, with suitable modification, to measurements of the matter power spectrum).

As discussed above in connection with weak lensing, we construct estimators $\hat{\gamma}_{\alpha,i}$ for the shear components ($\alpha = 1, 2$) in some pixel i that contains N galaxies. The error with which we measure each shear in this pixel is then $e/\sqrt{N}$, where e is the mean intrinsic ellipticity. Since the Poisson fluctuations in each pixel are expected to be statistically independent, we have,

$$\langle \hat{\gamma}_{\alpha,i} \hat{\gamma}_{\beta,j} \rangle = \delta_{\alpha\beta} \delta_{ij} \frac{e^2}{N}. \tag{38}$$

The relation between the convergence and the shear is,

$$\tilde{\kappa}(\vec{\ell}) = \int d^2\theta \, e^{-i\vec{\ell}\cdot\vec{\theta}} \left[\gamma_1(\vec{\theta}) \cos\left(2\phi_{\vec{\ell}}\right) + \gamma_2(\vec{\theta}) \sin\left(2\phi_{\vec{\ell}}\right) \right]. \tag{39}$$

In an actual measurement, this integral over the area on the sky is replaced by a sum over the pixels in the sky,

$$\tilde{\kappa}(\vec{\ell}) = (\Delta\theta)^2 \sum_i e^{-i\vec{\ell}\cdot\vec{\theta}_i} \left[\gamma_{1,i}(\vec{\theta}) \cos\left(2\phi_{\vec{\ell}}\right) + \gamma_{2,i}(\vec{\theta}) \sin\left(2\phi_{\vec{\ell}}\right) \right], \tag{40}$$

where $(\Delta\theta)^2$ is the area of each pixel, and the sum is over all pixels.

Suppose now that there was no cosmological signal. We would still most generally infer some nonzero value $\tilde{\kappa}(\vec{\ell})$ due to random fluctuations in the intrinsic galaxy shapes. We can evaluate this to be

$$\left\langle \tilde{\kappa}(\vec{\ell}) \tilde{\kappa}^*(\vec{\ell}') \right\rangle = (\Delta\theta)^4 \frac{e^2}{N} \sum_i e^{-i\vec{\theta}_i\cdot(\vec{\ell}+\vec{\ell}')} \left[\cos\left(2\phi_{\vec{\ell}}\right) \cos\left(2\phi_{\vec{\ell}'}\right) + \sin\left(2\phi_{\vec{\ell}}\right) \sin\left(2\phi_{\vec{\ell}'}\right) \right]. \tag{41}$$

The sum is evaluated by transforming back to an integral, using $(\Delta\theta)^2 \sum_i \to \int d^2\theta$, and then the integral over $d^2\theta$ becomes $(2\pi)^2 \delta^{(2)}(\vec{\ell} - \vec{\ell}')$. We then find that even in the absence of a cosmological signal, random shape fluctuations give rise to a noise contribution,

$$C_\ell^{\kappa\kappa,\mathrm{n}} = \frac{(\Delta\theta)^2 e^2}{N} = \frac{e^2}{n}, \tag{42}$$

where $n = N/(\Delta\theta)^2$ is the density of sources on the sky. When we do a measurement of the convergence power spectrum, we will find an observed (o) value,

$$C_\ell^{\kappa\kappa,\mathrm{o}} = C_\ell^{\kappa\kappa} + C_\ell^{\kappa\kappa,\mathrm{n}}, \tag{43}$$

that is the sum of the cosmological signal $C_\ell^{\kappa\kappa}$ and the noise $C_\ell^{\kappa\kappa,\mathrm{n}}$. The cosmic signal can then be inferred by subtracting the expected noise.

There is then an additional noise that arises from cosmic variance, the sample variance that arises from the finite number of ℓ modes. Suppose we want to estimate $C_\ell^{\kappa\kappa}$ from a given survey. To do this we will do is take all the Fourier mode $\vec{\ell'}$ in the survey with wavevector magnitudes ℓ' that fall within a window $\ell + \delta\ell/2 < \ell' < \ell + \delta\ell/2$ of some small width $\delta\ell' \ll \ell$ around ℓ. There will be some number $\mathcal{N}_\ell$ of such modes. The amplitude $\tilde{\kappa}(\vec{\ell'})$, of each such mode is selected from a Gaussian distribution of variance C_ℓ^{κ}. If we then infer the variance from $\mathcal{N}_\ell$ such measurements, the variance with which we can measure this variance is then

$$\Delta C_\ell^{\kappa\kappa} = \left(\frac{2}{\mathcal{N}_\ell}\right)^{1/2} C_\ell^{\kappa,\kappa,\mathrm{o}} = \left(\frac{2}{\mathcal{N}_\ell}\right)^{1/2} \left[C_\ell^{\kappa\kappa} + C_\ell^{\kappa\kappa,\mathrm{n}}\right]. \tag{44}$$

Let's now go a little further. Suppose we try to estimate each C_ℓ individually; i.e., so that we take $\delta\ell = 1$. Then the number of Fourier modes is the density of states times the density of states. This density of states is the area of an annulus of radius ℓ and width $\delta\ell = 1$ is $\mathcal{N}_\ell = 2\pi\ell\delta\ell = 2\pi\delta\ell$. Now consider the density of states. First consider a one-dimensional "volume" (a length interval) of distance d. There Fourier modes are quantized with wavenumbers that are integer multiples of $2\pi/d$ yielding a 1D density of states $(2\pi/d)^{-1}$. The density of states in two dimensions is then the square of this, or $4\pi^2/A$, where A is the area of the survey. Putting it together, we have $\mathcal{N}_\ell = A\ell/(2\pi)$, where A is the area (in radian^2) of the survey. The error with which $C_\ell^{\kappa\kappa}$ can be measured is thus,

$$\Delta C_\ell^{\kappa\kappa} = \left(\frac{1}{f_{\mathrm{sky}}\ell}\right)^{1/2} \left[C_\ell^{\kappa\kappa} + \frac{e^2}{n}\right], \tag{45}$$

where we used the area, 4π steradians, of the full sky and f_{sky} is the fraction of the full sky surveyed. Given that $C_\ell^{\kappa\kappa}$ generally decreases with ℓ (we did not show that, but you can verify), the noise generally becomes more important at smaller scale (higher ℓ). There is thus usually a critical multipole moment,

$$\ell_m \equiv \left[\frac{\ell^2 C_\ell^{\kappa\kappa}}{2\pi}\frac{2\pi n}{e^2}\right]^{1/2} \simeq 1400 \left(\frac{n}{10\,\mathrm{arcmin}^{-2}}\right)^{1/2} \left(\frac{\ell^2 C_\ell^{\kappa\kappa}/(2\pi)}{10^{-4}}\right)^{1/2} \left(\frac{e}{0.2}\right)^{-1}, \tag{46}$$

above which noise dominates the signal.

(In CMB measurements there is noise in each individual pixel that comes from instrument noise (the finite precision with which you can measure a temperature). The analogous expression for the error with which C_ℓ^{TT} can be measured (on the full sky) is

$$\Delta C_\ell^{\mathrm{TT}} = \left(\frac{2}{2\ell+1}\right)^{1/2} \left[C_\ell^{\mathrm{TT}} + \frac{4\pi\sigma_T^2}{N_{\mathrm{pix}}}\right], \tag{47}$$

where σ_T is the instrumental noise in the measurement of the temperature in each pixel, and N_{pix} the number of pixels.)

6 Weak lensing (or cosmic shear) of the CMB

Lensing by mass inhomogeneities along the line of sight to the surface of last scatter also distort the CMB temperature and polarization pattern that we see.

Let's begin with some estimates. From the discussion in the last Section, which relates the deflection angle α to the convergence and potential, we can infer that the rms deflection angle due to lensing by large-scale structure is

$$\alpha_{\rm rms}^2 \int \frac{d^2\ell}{(2\pi)^2} C_\ell^{\alpha\alpha} = 4\int_0^\infty \frac{d\ell}{\ell}\frac{C_\ell^{\kappa\kappa}}{2\pi}, \tag{48}$$

or equivalently, the deflection-angle power spectrum is $C_\ell^{\alpha\alpha} = (4/\ell^2)C_\ell^{\kappa\kappa}$. If we use the ΛCDM power spectrum in the Limber-approximation expression for the convergence power spectrum, we find that $\ell^2 C_\ell^{\alpha\alpha} \propto \ell$ for $\ell \ll 50$, and $\ell^2 C_\ell^{\alpha\alpha} \propto \ell^{-3}$ for $\ell \gg 50$. The power spectrum $\ell^2 C_\ell^{\alpha\alpha}$ peaks at $\ell \simeq 50$ at a value $\ell^2 C_\ell^{\alpha\alpha}/(2\pi) \simeq 10^{-7}$. From this, we infer that the typical deflection angle for a CMB photon is $10^{-7/2} \simeq 1$ arcmin, and this deflection is coherent on distance scales that correspond to $\ell \sim 100$, or $\sim$ hundred Mpc.

The effects of lensing of the CMB are observable. As we have seen above, the CMB is seen to be very closely Gaussian, in agreement with the expectations with slow-roll inflation. This implies that there are no nontrivial correlation functions beyond the two-point correlation function, or equivalently, that the statistical properties of the temperature map are determined entirely in terms of the power spectrum. The distortion induced by lensing, though, changes that and gives rise to a characteristic departure from Gaussianity in the observed map. Here's how it works.

Since trajectory of a CMB photon in a direction $\vec\theta$ is deflected by an angle $\vec\alpha(\vec\theta)$, the temperature $T(\vec\theta)$ observed in a given direction $\vec\theta$ is related to the primordial temperature $T^p(\vec\theta)$ in that direction by

$$T(\vec\theta) = T^p\left(\vec\theta - \alpha(\vec\theta)\right) \simeq T^p(\vec\theta) - \alpha(\vec\theta)\nabla_\theta T^p(\vec\theta) \simeq T^p(\vec\theta) - \nabla_\theta\psi(\vec\theta)\nabla_\theta T^p(\vec\theta). \tag{49}$$

In Fourier space, the lensed and unlensed maps are related, in the presence of a deflection potential $\psi(\vec\theta)$, by

$$\tilde T(\vec\ell) \simeq \tilde T^p(\vec\ell) + \int \frac{d^2\ell_1}{(2\pi)^2}\left[\vec\ell_1\cdot(\vec\ell-\vec\ell_1)\right]\tilde T^p(\vec\ell_1)\psi(\vec\ell-\vec\ell_1), \tag{50}$$

where we have used the fact that multiplication in configuration space becomes convolution in Fourier space.

Now consider the correlation between two different Fourier modes of wavevectors $\vec\ell$ and $\vec\ell'$. It is

$$\left\langle \tilde T(\vec\ell)\tilde T^*(\vec\ell')\right\rangle_{\vec\ell\neq\vec\ell'} \simeq \int \frac{d^2\ell}{(2\pi)^2}\left[\vec\ell_1\cdot(\vec\ell-\vec\ell_1)\right]\tilde\psi(\vec\ell-\vec\ell_1)\left\langle\left[\tilde T^p(\vec\ell')\right]^*\tilde T^p(\vec\ell_1)\right\rangle + c.c., \tag{51}$$

where c.c. is the complex conjugate. The expectation value in the integrand sets $\ell' = \ell_1$, resulting in,

$$\left\langle \tilde T(\vec\ell)\tilde T^*(\vec\ell')\right\rangle_{\vec\ell\neq\vec\ell'} \simeq f(\vec\ell,\vec\ell')\tilde\psi(\vec\ell), \tag{52}$$

with

$$f(\vec\ell,\vec\ell') = C_{\ell'}^{\rm TT,p}(\vec\ell'\cdot\vec L) + C_\ell^{\rm TT,p}(\vec\ell\cdot\vec L). \tag{53}$$

Now consider a single Fourier mode $\tilde{\psi}(\vec{L})$ of the deflection field of wavevector $\vec{L}$. This equation then tells us that any pair of observed temperature modes $\tilde{T}(\vec{\ell})$ and $\tilde{T}(\vec{\ell}')$ with $\vec{L} = \vec{\ell} - \vec{\ell}'$, can be combined to provide an estimator,

$$\widehat{\tilde{\psi}(\vec{L})}_{vec\ell,\vec{\ell}'} = \frac{\tilde{T}(\vec{\ell})\tilde{T}^*(\vec{\ell}')}{f(\vec{\ell},\vec{\ell}')} \tag{54}$$

for $\tilde{\psi}(\vec{L})$. This estimator is not perfect—it has a variance,

$$\left\langle \left(\widehat{\tilde{\psi}(\vec{L})}_{vec\ell,\vec{\ell}'} \right)^2 \right\rangle = \frac{C_\ell^{\rm TT,tot} C_{\ell'}^{\rm TT,tot}}{\left[f(\vec{\ell},\vec{\ell}') \right]^2}. \tag{55}$$

We can then consider all pairs of $\vec{\ell}$-$\vec{\ell}'$ modes with $\vec{\ell}+\vec{\ell}' = \vec{L}$. Each of these provides an independent estimator of $\tilde{\psi}(\vec{L})$ with some variance. We can then sum all of these to obtain the optimal (or minimum-variance) estimator for $\tilde{\psi}(\vec{L})$:

$$\widehat{\tilde{\psi}(\vec{L})} = A_L \sum_{\vec{\ell}} \tilde{T}(\vec{\ell}) \tilde{T}(\vec{L} - \vec{\ell}) F(\vec{e}ll, \vec{L} - \vec{\ell}), \tag{56}$$

with

$$F(\vec{\ell},\vec{\ell}') \equiv \frac{f(\vec{\ell},\vec{\ell}')}{2C_\ell^{\rm TT,tot} C_{\vec{\ell}'}^{\rm TT,tot}}, \qquad A_L = L^2 \left[\sum_\ell f(\vec{\ell}, \vec{L}-\vec{\ell}) F(\vec{\ell}, \vec{L}-\vec{\ell}) \right]^{-1}, \tag{57}$$

where $C_\ell^{\rm TT,tot}$ is the total observed power spectrum (signal plus noise). This then provides a straighforward recipe to take a temperature map $T(\vec{\theta})$ and from it reconstruct the (Fourier-space) projected potential $\tilde{\psi}(\vec{L})$, and from it, the real-space projected potential, and thus the projected surface-mass density. The effects of lensing of the CMB were first detected almost a decade ago, and one of the really nice results from Planck was the first all-sky projected mass map obtained from lensing.

6.1 Lensing-induced B modes

One of the most important consequences of lensing of the CMB is the B modes they induce. Generalizing the analysis of the temperature above gives us

$$\begin{pmatrix} T \\ Q \\ U \end{pmatrix}_{\rm obs.} (\boldsymbol{\theta}) = \begin{pmatrix} T \\ Q \\ U \end{pmatrix}_{\rm ls} (\boldsymbol{\theta} + \delta\boldsymbol{\theta}) \simeq \begin{pmatrix} T \\ Q \\ U \end{pmatrix}_{\rm ls} (\boldsymbol{\theta}) + \delta\boldsymbol{\theta} \cdot \nabla \begin{pmatrix} T \\ Q \\ U \end{pmatrix}_{\rm ls} (\boldsymbol{\theta}), \tag{58}$$

where $\delta\boldsymbol{\theta} = \nabla\Phi$ is the lensing deflection, and Φ is a projection of the three-dimensional gravitational potential $\Phi(\boldsymbol{x})$ along the line of sight $\hat{n}$.

The generation of B modes by lensing is most easily seen in the flat-sky limit. If there is no B mode at the surface of last scatter, then $\tilde{Q}(\boldsymbol{\ell}) = 2\tilde{E}(\boldsymbol{\ell}) \cos 2\varphi_{\boldsymbol{\ell}}$ and $U(\boldsymbol{\ell}) = -2E(\boldsymbol{\ell}) \sin 2\varphi_{\boldsymbol{\ell}}$. Thus,

$$\nabla Q(\boldsymbol{\theta}) = -2i \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} \, \tilde{E}(\boldsymbol{\ell}) \, \cos 2\varphi_{\boldsymbol{\ell}} \, \boldsymbol{\ell} \, e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\theta}}, \tag{59}$$

and similarly for $\nabla U(\boldsymbol{\theta})$ with $\cos \to -\sin$. The deflection angle is likewise

$$\tilde{\Phi}(\boldsymbol{\theta}) = -i \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2}\, \Phi(\boldsymbol{\ell})\, e^{-i\boldsymbol{\ell}\cdot\boldsymbol{\theta}}\, \boldsymbol{\ell}. \tag{60}$$

Thus, the perturbation to Q and U induced by gravitational lensing is

$$\delta Q(\boldsymbol{\theta}) = (\nabla Q)\cdot(\nabla\Phi) = \int \frac{d^2\boldsymbol{\ell}}{(2\pi)^2} e^{i\boldsymbol{\ell}\cdot\boldsymbol{\theta}} (\nabla Q \cdot \nabla \Phi)_{\boldsymbol{\ell}}, \tag{61}$$

where

$$\delta Q(\boldsymbol{\ell}) \equiv [(\nabla Q)\cdot(\nabla\Phi)]_{\boldsymbol{\ell}} = 2\int \frac{d^2\boldsymbol{\ell}_1}{(2\pi)^2}[\boldsymbol{\ell}_1\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}_1)]\tilde{E}(\boldsymbol{\ell}_1)\tilde{\Phi}(\boldsymbol{\ell}-\boldsymbol{\ell}_1)\cos 2\varphi_{\boldsymbol{\ell}_1}, \tag{62}$$

$$\delta U(\boldsymbol{\ell}) \equiv [(\nabla U)\cdot(\nabla\Phi)]_{\boldsymbol{\ell}} = -2\int \frac{d^2\boldsymbol{\ell}_1}{(2\pi)^2}[\boldsymbol{\ell}_1\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}_1)]\tilde{E}(\boldsymbol{\ell}_1)\tilde{\Phi}(\boldsymbol{\ell}-\boldsymbol{\ell}_1)\sin 2\varphi_{\boldsymbol{\ell}_1}. \tag{63}$$

Although the original map had (by assumption) no curl, the lensed map does:

$$B(\boldsymbol{\ell}) = \frac{1}{2}[\sin 2\varphi_{\boldsymbol{\ell}}\,\delta Q(\boldsymbol{\ell}) - \cos 2\varphi_{\boldsymbol{\ell}}\,\delta U(\boldsymbol{\ell})] = \int \frac{d^2\boldsymbol{l_1}}{(2\pi)^2}[\boldsymbol{\ell}_1\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}_1)]E(\boldsymbol{\ell}_1)\Phi(\boldsymbol{\ell}-\boldsymbol{\ell}_1)\sin 2\varphi_{\boldsymbol{\ell}_1}. \tag{64}$$

If the power spectrum for the projected potential is $C_\ell^{\Phi\Phi}$, then the B-mode power spectrum from lensing is,

$$C_\ell^{\rm BB} = \int \frac{d^2\boldsymbol{l_1}}{(2\pi)^2}[\boldsymbol{\ell}_1\cdot(\boldsymbol{\ell}-\boldsymbol{\ell}_1)]^2 \sin^2 2\varphi_{\boldsymbol{\ell}_1} C^{\Phi\Phi}_{|\boldsymbol{\ell}-\boldsymbol{\ell}_1|} C^{\rm EE}_{\ell_1}. \tag{65}$$

6.2 Smoothing of peaks

A similar analysis for the temperature power spectrum gives rise to a change,

$$\Delta C_\ell^{\rm TT} = \sum_{\vec{\ell}_1} C^{\psi\psi}_{\vec{\ell}'} \left\{ C^{\rm TT}_{|\vec{\ell}-\vec{\ell}_1|} \left[\vec{\ell}_1\cdot(\vec{\ell}-\ell_1)\right]^2 - C_\ell^{\rm TT}(\vec{\ell}\cdot\vec{\ell}_1)^2 \right\}, \tag{66}$$

in the temperature power spectrum. To derive this result (which is second order in ψ), one needs to expand the observed temperature to second order in ψ; i.e.,

$$T(\vec{\theta}) = T^p\left(\vec{\theta} - \alpha(\vec{\theta})\right) \simeq T^p(\vec{\theta}) - \alpha(\vec{\theta})\nabla_\theta T^p(\vec{\theta}) + \frac{1}{2}\alpha_i\alpha_j \frac{\partial^2 T^p}{\partial\theta_i\partial\theta_j} + \cdots. \tag{67}$$

There is then a contribution that comes from the cross term between the second-order and zero-th terms, in addition to the contribution (as we saw for B modes) from the square of the first-order term.

There are two terms in the integrand of Eq. (68. The first involves a convolution of the temperature and projected-potential power spectra. The acoustic peaks in the CMB power spectrum arise at $\ell \gtrsim 200$, while the lensing power spectrum peaks at $\ell \sim 50$. We can thus take $\ell' \ll \ell$ in the integrand in the first term, and replace $C^{\rm TT}_{|\vec{\ell}-\vec{\ell}_1|}$ therein by a smoothed power spectrum $\bar{C}_\ell^{\rm TT}$ which is the temperature power spectrum smoothed over a region of with roughly $\Delta\ell \simeq 50$ about ℓ. The change can then be approximated,

$$\Delta C_\ell^{\rm TT} = \left(\bar{C}_\ell^{\rm TT} - C_\ell^{\rm TT}\right) \sum_{\vec{\ell}_1} C_{\vec{\ell}'}^{\psi\psi} (\vec{\ell} \cdot \vec{\ell}_1)^2. \tag{68}$$

From this expression, we see that near peaks (where C_ℓ is higher than its smoothed value), the power spectrum is reduced by lensing, while near troughs (where the smoothed power spectrum is higher), it is increased. The observed power spectrum is thus smoothed be lensing.

There is a simple heuristic explanation for this smoothing. When we look at different $\sim 2^\circ \sim 2^\circ$ regions of the sky, we are seeing a patch of the CMB that has been lensed by over- or under-densities in such a way that the observed angular scale is either amplified or de-amplified. In each such patch, the peaks in the angular power spectrum are thus shifted to either slightly smaller or slightly larger values of ℓ. When we average over the entire sky, the peaks are then smoothed.