Special Topic: Gravitational Lensing

The observational confirmation of light deflection, and in particular the confirmation that for an impact parameter $b \gg 2GM/c^2$ the deflection angle is $4GM/(bc^2)$ (rather than the Newtonian $2GM/(bc^2)$) made Einstein a global superstar in 1919. In the last few decades the uses of gravitational lensing as a probe of many phenomena have accelerated, so that now it plays a role in the study of quasars, galaxies, galaxy clusters, dark matter, the Hubble constant, dark energy, and exoplanets!

Given that we have just completed our study of the basics of general relativity and black holes, it seems a good time to study the fundamentals of gravitational lensing, as well as the applications of this phenomenon.

The lens equation and other basics

As we discussed during the black hole classes, light deflection sufficiently close to a compact object can be substantial; indeed, although it would be unstable, in principle light could orbit in a circle around a black hole (at $R_{\rm light} = 3GM/c^2$ around a nonrotating black hole, for example). However, most of the applications of lensing focus on weak lensing, in the sense that the photons we observe are always many gravitational radii (GM/c^2) away from masses.

In that limit, as we said above, the total deflection angle (i.e., the angle between the final propagation direction and the initial propagation direction) is

$$\hat{\alpha} = \frac{4GM}{bc^2}\ , \tag{1}$$

measured in radians. This forms the basis of gravitational lensing theory. Another implicit assumption is that the source of the photons is very far from the gravitating mass, as are we the observers. This means that lensing by itself does not change the specific intensity; any net redshifts or blueshifts from the source to us are the same as they would be if the lens were removed. As we noted in an earlier class, one consequence of this is that the surface brightness of the source (i.e., the flux per solid angle) is not changed by lensing; if lensing makes a source brighter, it also increases its solid angle.

Typical deflection angles are not large. A photon grazing the limb of the Sun is deflected by $\hat{\alpha} = 1.75''$. Star-star microlensing can produce deflections of milliarcseconds, so that even if multiple images are produced they cannot be distinguished. Lensing by giant galaxies, or by galaxy clusters, can produce separations of a few arcseconds, which *can* be separated. As we will discuss below, this means that detection of lensing can take various forms, which depend on the specific application.

To set up the fundamental lensing equation, we will first take a quick diversion into the different definitions of distances. We will do this because many lensing applications are cosmological, and when cosmological distances are involved care must be taken in the definitions.

Suppose you want to determine the distance to a friend of yours, who is walking along a sidewalk to you. There are different approaches you could take. You could, for example, measure the length of a segment of the sidewalk and count how many segments it is to your friend. You could close one eye, then the other, and see how much your friend appears to move against the distant background. You could use your knowledge of how tall your friend is, and compare that with the angular width your friend subtends. If your friend has some weird fetish for carrying around lit 100 watt light bulbs, you could see how much light you get from the bulb. At small distances (say, within our Galaxy) the equivalents of all these approaches would yield the same value for the distance, but over cosmological distances they differ. For the curious, these methods, in order, correspond to proper distance, parallax distance, angular diameter distance, and luminosity distance (see https://arxiv.org/pdf/astro-ph/9905116.pdf for a nice description of the different types of distance). For gravitational lensing, we need to use the angular diameter distance.

With that in hand, let us consider a usually-valid approximation: the *thin lens* approximation. In this, we assume that the lens has an extent along the line of sight that is much smaller than the distance from the lens to us, or from the lens to the source. Let D_d be the angular diameter distance from us to the lens, D_s be the angular diameter distance from us to the source, and D_{ds} be the angular diameter distance from the lens to the source. It will be useful in the following to consult the lensing diagram from the Wikipedia page on lensing formalism: https://en.wikipedia.org/wiki/Gravitational_lensing_formalism.

Suppose that we have defined an optic axis that gives us a two-dimensional (i.e., projected) coordinate system on the sky. We define a photon's path as $(\xi_1(\lambda), \xi_2(\lambda), r_3(\lambda))$ (where λ is an affine parameter and r_3 is the distance, and the 1 and 2 designations mean that these are the two-dimensional, projected, coordinates). The true position of our source is in a direction $\boldsymbol{\beta}$, where the boldfacing means that this is a two-dimensional vector. Due to lensing, we see it in the direction $\boldsymbol{\theta}$. The two are related by:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta}) \equiv \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \ . \tag{2}$$

There are many definitions here, so let's walk through them carefully.

- In the thin-plane approximation, we only care about the photon's path very near the lens. Thus instead of $(\xi_1(\lambda), \xi_2(\lambda))$, we only have to worry about the projected location of the ray near the lens: $\boldsymbol{\xi}$.
- Suppose that we break the lens into a large number of infinitesimal mass elements, and

that a given mass element in the lens is at a projected location of $\boldsymbol{\xi}'$. Then the total deflection angle, where we use the integral form of $4GM/(bc^2)$, is

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2} \int d^2\boldsymbol{\xi}' \int dr_3' \rho(\xi_1', \xi_2', r_3') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} . \tag{3}$$

But given our thin-lens approximation, we can simplify a bit by integrating the lens mass density over the radial direction to form the surface mass density:

$$\Sigma(\boldsymbol{\xi}) \equiv \int dr_3 \rho(\xi_1, \xi_2, r_3) \tag{4}$$

which leads to

$$\hat{\boldsymbol{\alpha}} = \frac{4G}{c^2} \int d^2\boldsymbol{\xi}' \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} . \tag{5}$$

- The preceding gives the necessary definitions for the first part of Equation (2). For the second part, the "scaled deflection angle" is given by

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int d^2\boldsymbol{\theta}' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} , \tag{6}$$

where the dimensionless surface mass density is

$$\kappa(\boldsymbol{\theta}) \equiv \frac{\Sigma(D_d \boldsymbol{\theta})}{\Sigma_{\rm cr}} \tag{7}$$

with the critical surface mass density being

$$\Sigma_{\rm cr} \equiv \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}} . \tag{8}$$

This surface mass density is "critical" because if $\Sigma > \Sigma_{\rm cr}$ it is guaranteed that some source positions can lead to multiple images.

The lens equation (2) is fundamental to the study of gravitational lenses. Note that it can have more than one solution; that is, for a given actual source direction β, it is possible that there are several directions θ in which the source could be seen by an observer. For example, if the lens is spherical and the source is a point, then if $\Sigma > \Sigma_{\rm cr}$ and the source is directly behind the lens as seen by the observer, the observer will see the source spread out in a ring. Real lenses are often more complex than spheres (obviously!), which means that the morphology of multiple images can be complicated and therefore informative about the lens and sometimes about the source.

The Einstein radius and time delays

If the lens is a point mass (a star is a good approximation), then a characteristic angular scale is the "Einstein radius"

$$\begin{aligned} \theta_E &= \left(\frac{4GM}{c^2}\frac{D_{LS}}{D_L D_S}\right)^{1/2} \\ &\approx 2.85" \left(\frac{M}{10^{12}\ M_\odot}\frac{1\ \text{Gpc}}{D_L D_S/D_{LS}}\right)^{1/2}\ , \end{aligned} \tag{9}$$

where we have scaled by a ∼Milky Way scale dark matter halo and a typical cosmological distance. If instead we scale by typical stellar microlensing parameters, we get

$$\theta_E \approx 0.0016" \left(\frac{M}{1\ M_\odot}\frac{3\ \text{kpc}}{D_L D_S/D_{LS}}\right)^{1/2}\ . \tag{10}$$

Roughly, a ray needs to pass within θ_E of a point mass for the source to be multiply imaged, and the angular separation of multiple images is also of order θ_E. Note that $\theta_E \propto M^{1/2}$. That means that the solid angle for multiple imaging is $d\Omega = \pi\theta_E^2 \propto M$, which tells us that the total probability of lensing from a population of point masses is simply proportional to the total mass of the population. As a result, for example, the probability that a given star will have multiple-imaging microlensing by some other star is considerably greater than the probability that the lensing will be by some black hole, because the total mass in black holes is much less than the total mass in stars. More about microlensing in a bit. The "probability proportional to mass" also means that if the lens is a galaxy, the lensing is dominated by the stellar and gas mass rather than by the presence of a supermassive black hole in the center (given that supermassive black holes comprise only a tiny fraction of the total mass of the galaxy).

If there are multiple images then they take different paths to go from the source to the observer, which means that (barring special symmetry) the propagation time is different as well. The characteristic time difference is the light-crossing time of the Schwarzschild radius of the lensing mass:

$$\Delta t \sim 2GM/c^3 \approx 10^{-5}\ \text{s}\ (M/M_\odot) \approx 4\ \text{months}\ (M/10^{12}\ M_\odot)\ . \tag{11}$$

The actual time delay depends on details, but it is roughly of this order.

Magnification and shearing

Even if the ray path is not sufficiently close to the lens to produce multiple images, lensing can still have an effect. One such effect is magnification of the brightness of a source. The closer the lens comes to the source-observer direction, the larger the magnification. In contrast, if a lens is angularly far enough away, it can bend light *away* from the source, and actually dim it compared to how it would look without the lens. One consequence of this is that the brightness of cosmologically distant sources is affected in a probabilistic way by lensing; even if all sources were intrinsically identical, we'd see a distribution of brightness from sources at a fixed redshift. This has implications for the analysis of, for example, Type Ia supernovae (the "standard bombs" that facilitated the discovery of the accelerated expansion of the universe). If we want to get more precise information about the expansion of the universe, this effect must be taken into account.

Magnification also means that, in lucky cases where a source is strongly magnified, it can be studied with far greater signal to noise than would be possible otherwise. This allows us to see rare sources at greater redshift, and with much better precision, than if we had no lensing. Because lensing conserves surface brightness, brightened sources also have larger solid angles than they would without lensing. This means that extended sources, such as galaxies, can be expanded in their angular detail.

This, however, brings us to the next effect: shearing and distortion. Even if there isn't any multiple imaging, the contours of a lensed image will typically have a different shape than the contours of the unlensed image. This often takes the form of shearing; the iconic image of multiply imaged galaxies behind the galaxy cluster Abell 2218 show this. "Weak lensing", which does not produce multiple images (in contrast to "strong lensing", which does), involves an attempt to quantify the degree of shear of images of galaxies. It is an increasingly important tool for the study of large-scale structure. Note, though, that because we do not know in advance the unlensed shape of the galaxy, such analyses require high-powered statistical techniques rather than being able to get information one image at a time.

Applications: Quasars

Quasars are effectively point sources, and they vary a lot. These two characteristics make them ideal for certain types of lensing studies. The first example was the quasar QSO 0957+0561 (Walsh, Carswell, and Weymann 1979, Nature, 279, 381), which is at a redshift of $z = 1.4$ and which has two images separated by 5.7". Because the quasar varies, it is possible to correlate the two images to determine the time delay, which is approximately 417 days. As an example of the difficulties of such analysis, Sazhin et al. (2003, MNRAS, 343, 353) reported what appeared to be a pair of images of the same galaxy, with characteristics that could only be explained by the lens being a cosmic string(!), but later observations showed that in fact this system is just two *similar* galaxies in chance projection (which Sazhin et al. had also suggested as a possibility).

Given that the time delay depends on angular diameter distances, and that the redshifts of both the lensing galaxy and the distant quasar are known very accurately in cases of this type, such systems would seem to be excellent probes of cosmography, e.g., measurements of the Hubble constant and the like. However, the substantial natural variability of quasars and the fact that the different images pass through different parts of the lensing galaxy mean that time delay analysis is extremely difficult. One byproduct (in a sense) of these studies has been the discovery of microlensing of quasar images by the stars in the lensing galaxy. The angular scale of the microlensing is so small (microarcseconds at these distances) that their Einstein radii are *smaller* than the accretion disks of the quasars. Thus the microlensing events are chromatic: they lens part of the disk more than others, and this progresses during the event. Given the large number of stars, again major statistical analysis is needed, but this is providing us with new information about quasar disks.

Another application that hasn't quite panned out yet is the use of quasar strong lensing statistics to measure aspects of dark energy. The idea is that the probability that a quasar will be strongly lensed (i.e., multiply imaged) depends, among other things, on the fraction of the total energy budget of the universe that is in dark energy. The problem is that strong lensing magnifies the source, which makes strongly lensed sources easier to see; thus the fraction of quasars that are *seen* to be strongly lensed is much larger than the fraction of quasars that are *actually* strongly lensed, because we are obviously biased to see brighter sources. Because quasars that would have been below our threshold of observation thus become visible, we would need to know about the brightness distribution of the sub-threshold quasars to use lensing statistics to tell us about dark energy. Pity!

Applications: Clusters and dark matter

Because lensing depends only on the presence of mass-energy, rather than its type, lensing can be used to weigh lenses. This has been particularly useful for galaxy clusters. Galaxy clusters, which we'll revisit later in the course, are somewhat poorly named: perhaps 2% of a typical cluster's mass is in the stars in galaxies, with $10\times$ as much in hot gas and the remaining $\sim 80\%$ in dark matter. They can be weighed by looking at the orbits of their galaxies, or by looking at the temperature of the gas (assuming it is in virial equilibrium). But another, independent, technique uses weak lensing of background galaxies. This tells us that the mass estimates are all approximately consistent with each other. A particularly high-profile result of this type involved the analysis of the so-called "Bullet Cluster". This is actually two clusters that are passing through each other at high speed, and the point is that the lensing map of the mass corresponds to the positions of the galaxies (which are small and don't interact with each other) rather than to the location of the hot gas (which collides and shocks with itself). This suggests that most of the mass is in fact collisionless, as expected for cold dark matter.

Applications: Microlensing, dark matter, and exoplanets

There is significant indirect evidence to suggest that most of the ordinarily-gravitating matter in the universe (as opposed to the repulsive dark energy) is of a collisionless "dark matter" type. What it is, however, is unknown. If dark matter is in Massive Compact Halo Objects (MACHOs, to counteract Weakly Interacting Massive Particles, or WIMPs; maybe a bit too much testosterone flying around!) then background stars would be lensed frequently enough to be seen. As a result, microlensing monitoring projects such as OGLE and EROS were started a couple of decades ago. At first blush this seems crazy, because millions of stars would have to be monitored for years to see a reasonable number (say, tens) of them lensed, and it is guaranteed that a much larger fraction of those millions of stars will be naturally variable stars. Luckily, it is possible to distinguish a gravitational lensing event from natural variation by the achromaticity of the lensing event: that is, all wavelengths are brightened by exactly the same amount, and are also dimmed on the way out of the event in a way that is independent of wavelength. Moreover, an event caused by the passage of a lensing star is symmetric, i.e., the ingress and egress are mirrors of each other. In contrast, variable stars change colors as they brighten and dim, and the light curve doesn't have the shape of a lensing event. This allows the needle (lensing) to be picked out of the haystack (the much larger number of variable stars).

Because the Einstein radius scales as the square root of the lens mass, if a lens (star, black hole, whatever) passes in front of a source star, the duration of the event will scale as the square root of the lens mass (assuming that the relative angular speed is fixed). This has allowed the microlensing community to place limits on the fraction of dark matter that could be in compact objects of various types; the best limits are at about 10% of total dark matter, and there are interesting constraints that range from $\sim 10^{-8}\ M_\odot$ to $10\ M_\odot$. Some candidate black hole events have been seen, that have lasted for more than a year, but there is also the possibility that the lens just happened to have a small relative angular velocity. Incidentally, when the events have such a long duration, effects such as parallax help break some of the degeneracies.

There have even been exoplanets detected by microlensing! An example is the event OGLE-2005-BLG-390. The signature is of a smooth, achromatic rise and a smooth, achromatic decline that has a short bump on either the rise or decline caused by the influence of the planet. It has been argued that microlensing is the only way to find analogs to Uranus and Neptune, whose orbital periods are too long to claim confident detections via radial velocity variations or transit.

All in all, lensing has come a long way! It is an important tool for many fields in astronomy.