1

Overview

Light passing by a point particle with mass M is deflected (Fig. 1.1) by an angle

$$\delta\theta = \frac{4MG}{bc^2} \tag{1.1}$$

where b is the distance of closest approach, the *impact parameter*; G is Newton's constant; and c is the speed of light.

Physically and dimensionally, this makes sense: the deflection angle is bound to increase as the mass of the deflector, or *lens*, gets larger and decrease as the light path strays farther from the lens. And, of course, the deflection is caused by gravity, so must be proportional to Newton's constant.

The factor of 4 is trickier and carries an interesting history. When developing his theory of gravity, general relativity, Einstein initially derived a result with coefficient equal to 2, a prediction that agreed with one obtained a century earlier that relied on Newtonian physics. Based on this result, he calculated the deflection angle of a light ray coming from a star aligned with the edge of the Sun and determined it to be a little less than an arcsecond. Einstein sent eminent astronomer George Hale a letter asking if a deviation of this size could be detected. Hale responded that the starlight would be dominated by light from the Sun unless the Sun were eclipsed. So, Einstein courted the German astronomer Erwin Freundlich to undertake an expedition to the site at which the next eclipse would be maximal: the Crimean Peninsula in Summer 1914. It turns out that August 1914 was not a very good time to be wandering around the continent, and Freundlich and his team were imprisoned. The good news is that they were released eventually; the apparent bad news is that they were released after the eclipse was over, so could not make the observations that Einstein deemed "a simply invaluable service to theoretical physics." This was only *apparent* bad news, because Einstein subsequently tweaked his theory and – as we will see – got a different coefficient for the deflection.

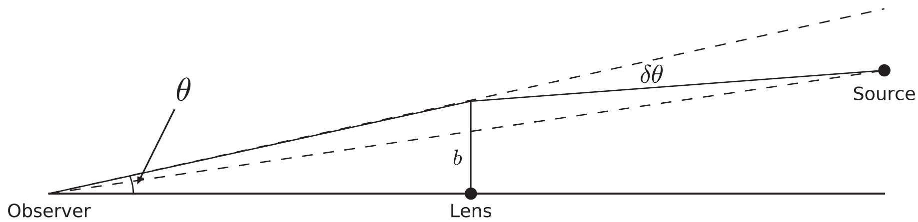

Figure 1.1 Light from a source (far right) is deflected by an angle $\delta\theta$ as it passes a projected distance b by a lens of mass M. The source appears to the observer to be an angular distance θ away from the line of sight connecting the observer to the lens.

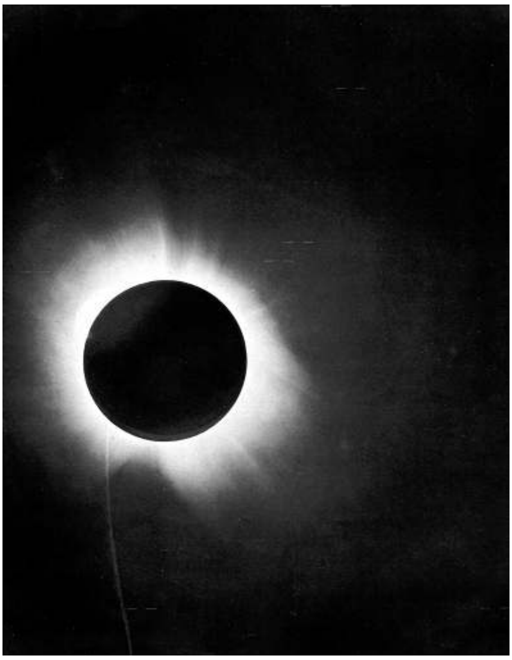

Figure 1.2 Photograph of the 1919 solar eclipse by Dyson et al. (1920).

In May 1919, Sir Frank Watson Dyson and Sir Arthur Eddington led an expedition to the island of Príncipe, off the west coast of Africa, and sent another team to Brazil in case the weather was cloudy. They were to observe the positions of stars in the Hyades cluster – which was situated right near the limb of the Sun – to see if they were shifted by the mass of the Sun. Their team took the photograph shown in Fig. 1.2. The positions of the background stars were indeed off from their expected positions by 1.9″, in agreement with the (improved) theory. Einstein became a celebrity and the theory of general relativity officially became the theory of gravity.

The success of the 1919 expeditions is often heralded as a crowning achievement, in that it secured Einstein's fame and general relativity's place in the pantheon of fundamental physics laws. This is true, but this "game-ending" interpretation casts a shadow over perhaps an even more important ramification. The expeditions and the success of general relativity paved the way for a new field of astrophysics: gravitational lensing.

The path from 1919 to what might be called the first direct detection of a lensing event by Walsh et al. (1979) was so long in large part because instrumentation needed to catch up with theory. But there is another reason for the delay: since we usually do not know the true position of astronomical objects, it is difficult to understand how or when the deflection of light might be detected. If we don't know where things really are, then how can we tell that they do not appear where they should be? The case of the starlight passing by the Sun appears to be an almost unique case, where we know the actual location of the source and the lens and the mass of the lens. In general we will not have these advantages, so what are the observable effects of light deflection? The rest of this chapter provides a first glimpse at the answers to this question. The phenomena outlined here and explored in detail in the book have enabled gravitational lensing to play a powerful role in many areas of modern astronomy and cosmology.

Lensing phenomena vary in their complexity, ranging from the simple case of a point source lensed by a point mass (leftmost in Fig. 1.3) to a diffuse source lensed by a diffuse mass distribution (lower right). The outline in this chapter and the details in the ensuing chapters move gradually from the simple to the complex, with the aim of unifying these disparate phenomena and driving home the point that they all emerge from the same relatively simple physical law.

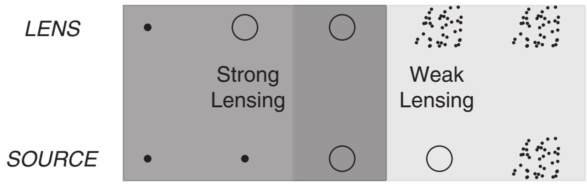

Figure 1.3 Cartoon depicting the range of phenomena spanned by gravitational lensing. The leftmost limit is when a point mass lenses a point source. Moving to the right, first the lens and then both the source and lens can be extended but are still single objects. The effects seen in these cases – multiple images, magnification, and microlensing – are associated with *strong lensing*. Moving further to the right, a single extended source can be distorted by an intervening diffuse set of lenses, or equivalently a fluctuating gravitational potential. The most extreme example of lensing is when the source itself is diffuse, for example, the cosmic microwave background. These less dramatic phenomena are detectable only statistically and fall in the domain of *weak lensing*.

1.1 Multiple Images

We begin our overview of lensing with the simple case of a point source lensed by a point mass, so that the light from the source is deflected by the amount given in Eq. (1.1). If the deflection angle is much smaller than the separation of the source from the lens-observer line of sight, there will be little observable effect. The more dramatic effects of lensing occur for objects with separation θ smaller than the deflection angle:

$$\theta < \delta\theta. \tag{1.2}$$

If the distance between the lens and us is denoted D_L, then the transverse distance b that represents the distance of closest approach corresponds to an angular size $\theta = b/D_L$[1] away from the line of sight. Therefore, the above requirement becomes

$$\theta < \frac{4MG}{D_L\theta c^2} \tag{1.3}$$

or

$$\theta < \sqrt{\frac{4MG}{D_L c^2}} = \theta_E, \tag{1.4}$$

where θ_E is called the Einstein radius. This is a good enough estimate for our purposes now, but in Chapter 2, we will derive a slightly modified expression for the Einstein radius of a point mass that accounts for the finite distance between us and the source. More generally, extended mass distributions have different coefficients, but Eq. (1.4) gives a good sense of when interesting effects will occur.

As shown in Fig. 1.4, if the source is within the Einstein radius of the lens, then multiple images can be observed. Most dramatic of all, if the source is directly behind the lens, then, as seen in Figure 1.5, the image will be a ring around the lens, a so-called *Einstein ring*, with radius equal to the Einstein radius.

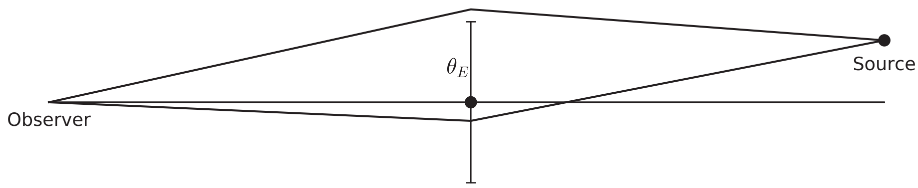

Figure 1.4 Light rays emanating in different directions from a source can be focused by an intervening lens on to the same point beyond the lens. The observer at this point will detect multiple images of the same object.

[1] Astronomy almost always works in the small angle approximation wherein $\sin\theta \simeq \tan\theta \simeq \theta$.

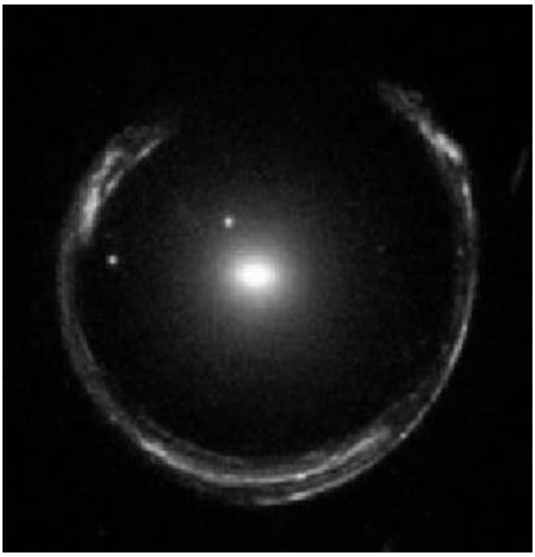

Figure 1.5 Einstein ring first observed by the Sloan Digital Sky Survey and then followed up with the Hubble Space Telescope. Foreground lens is a galaxy at the center, while the background object is almost perfectly aligned so is seen as a ring (Credit: ESA/Hubble & NASA). (See color plates section.)

Let's estimate a typical value for the Einstein radius. Normalize to a solar mass lens a distance of 1 kpc from us, so that

$$\theta_E = 1.4 \times 10^{-8} \left(\frac{\text{kpc}}{D_L}\right)^{1/2} \left(\frac{M}{M_\odot}\right)^{1/2}. \tag{1.5}$$

So every object has an Einstein radius, which increases with its mass and decreases the farther away it is from us. Translating to arcseconds, this becomes

$$\theta_E = 0.0028'' \left(\frac{\text{kpc}}{D_L}\right)^{1/2} \left(\frac{M}{M_\odot}\right)^{1/2}. \tag{1.6}$$

Typical angular resolutions for optical telescopes are in the arcsecond range, so a star in the Milky Way does lens distant objects but not in any way that can produce observable multiple images. At cosmological distances $D_L \sim 10^6$ kpc, a galaxy with mass of $10^{12} M_\odot$ has an Einstein radius of order an arc second, so can potentially produce multiple images of a distant point source.

1.2 Time Delay

Multiple images and most other lensing phenomena stem from the perpendicular deflection of light as it traverses past masses, or equivalently through a varying gravitational potential. But gravity causes not just changes in the perpendicular direction of the photons but also distortions in the propagation along the line of sight, i.e., time delays. Roughly, the time delay is of order the gravitational potential divided by the speed of light squared (MG/Rc^2 for a point mass) multiplied by the time spent in the neighborhood of the potential, R/c. So, we expect a time delay of order

$$\delta t \sim \frac{MG_{\odot}}{c^3} \sim 5\,\mu\text{sec} \tag{1.7}$$

for a light ray passing by the Sun. There are various factors of order unity that push this up to about 200 μsec, an effect dubbed the Shapiro time delay.

In the case of multiple images, the different light rays that reach us take different paths and therefore arrive at different times. If the source has a time dependence (i.e., is variable), the time delay can be observed, thereby obtaining information about both the lens and the various distances. This has the potential to be a powerful cosmological tool. Roughly, galaxies have masses of order $10^{12}M_{\odot}$ so delays of order $\delta t \sim 5 \times 10^6$ sec, or a few months, are expected and have been observed.

1.3 Magnification

Even if multiple images cannot be resolved, we might still observe the effect in the form of magnification since we receive flux from multiple directions. More generally, lensing conserves surface brightness (Exercise (1.3)), so the flux from an object depends on its apparent size. The magnification is given by the ratio of the lensed to unlensed sizes, as depicted in Fig. 1.6. For sources whose light passes within an Einstein radius of the lens ($\theta < \theta_E$) and whose true distance β from the line of sight connecting the observer to the lens is smaller than the Einstein radius ($\beta < \theta_E$), the magnification is of order

$$\mu \simeq \frac{\theta_E}{\beta}. \tag{1.8}$$

Therefore, even if multiple images cannot be observed by resolving them, they might contribute to an observed magnification of an object.

1.4 Microlensing

The discussion of magnification leads to the conclusion that a moving object passing between us and a distant star magnifies the star roughly for as long as the object is within the Einstein radius of the line of sight connecting us to the star. This short-term magnification is called *microlensing*. To get an estimate for the timescales, an object moving perpendicular to the line of sight with velocity v spends a time

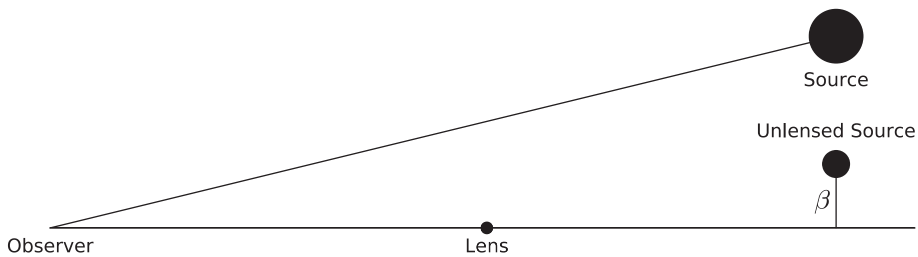

Figure 1.6 Magnification of an image is due to the increased area of a source. If the unlensed distance between the source and the line of sight connecting the observer to the lens β is much smaller than the Einstein radius, then this magnification is of order θ_E/β.

$$t \sim \frac{D_L \theta_E}{v} \tag{1.9}$$

within the Einstein radius of the line of sight. So the flux from the background star will be magnified for that brief amount of time. Let's plug in typical numbers for a Galactic lens: $D_L \sim 10$ kpc, $v \sim 200$ km/sec, and $M \sim M_\odot$. Then, the magnification will occur for a time

$$t \sim 6.8 \times 10^6 \,\text{sec} \left(\frac{D_L}{10\,\text{kpc}}\right)^{1/2} \left(\frac{M}{M_\odot}\right)^{1/2}, \tag{1.10}$$

a few months.

Although the lens could be a star, a much more intriguing possibility is that the lens is something that could not be seen directly. For example, Fig. 5.5 shows an example of a background star lensed by a dark foreground object. A lens like this in our Galaxy is called a MAssive Compact Halo Object (MACHO). MACHOs were proposed as a candidate for the dark matter in the Galaxy, with the name a lighthearted reference to an alternative candidate, Weakly Interacting Massive Particles (WIMPs). We have observed many such events – Fig. 5.5 shows one – and an important question that has apparently been resolved in the negative is whether MACHOs might be the sole component of dark matter in the Galaxy.

If we have detected a background star lensed by a moving foreground object in the manner described above, we can hunt for planets revolving around the foreground star. The effect is more pronounced than one might imagine: the planet does not simply add to the mass of the host star when it is aligned; rather, it leads to a significant bump in the magnification for the short period of time that the aligned system is within the Einstein radius of the planet. A Jupiter-sized planet has an Einstein radius 30 times smaller than the Sun (since $\theta_E \propto M^{1/2}$) so, from Eq. (1.10), the blip it produces in the light curve (as shown in Fig. 5.11) will last for a period of time on the order of a day. An Earth-sized planet will produce a signal shorter by another factor of $1000^{1/2}$, or roughly an hour. Astronomers have monitored lensed stars in this way and derived the remarkable fact that there are more planets than stars in the Galaxy.

1.5 Extended Lenses

The first way to generalize beyond point masses is to envision an extended lens (but still a single lens), characterized not by a mass M but by a 2D surface density (Exercise (1.4)) $\Sigma(R)$. The mass enclosed within a radius R is then roughly $\pi\Sigma(R)R^2$, so the generalization of Eq. (1.4) would then be

$$\theta_E \simeq \sqrt{\frac{4\pi G\Sigma R^2}{D_L c^2}} = \sqrt{\frac{4\pi G\Sigma D_L}{c^2}}\,\theta. \tag{1.11}$$

We can rewrite this as

$$\left(\frac{\theta_E}{\theta}\right)^2 = \frac{\Sigma}{\Sigma_{\rm cr}}, \tag{1.12}$$

where the critical surface density is approximately[2]

$$\Sigma_{\rm cr} \simeq \frac{c^2}{4\pi D_L G}. \tag{1.13}$$

So light passing by regions in which the density is greater than the critical surface density will be within the Einstein radius of the lens and will therefore produce the types of dramatic changes described above. Plugging in the cosmological distance $D_L \simeq 1$ Gpc leads to a cosmological value of $\Sigma_{\rm cr} \simeq 2\times 10^{13} M_\odot/(100\text{kpc})^2$. So galaxy clusters, which are roughly this massive and above, can dramatically impact images that lie close to their centers. An example is shown in Fig. 1.7.

1.6 Extended Sources

The distortions evident in Fig. 1.7, background galaxies that appear highly elongated, provide a segue to the next level of generalization: the case where not only the lens but the source as well is extended. When an individual object serves as a lens for background galaxies, the resulting pattern is characteristic: ellipticities oriented tangentially to the line of sight connecting them to the lens center. This *tangential shear*, we will see, is directly related to the surface density; roughly,

$$\gamma_t(\theta) \sim \frac{\Sigma(\theta)}{\Sigma_{\rm cr}}. \tag{1.14}$$

[2] The exact expression will be derived in Chapter 2 and includes a ratio of distances.

Figure 1.7 The shapes of background galaxies made highly elliptical by the foreground Abell cluster (Credit: Gravitational lensing in galaxy cluster Abell 2218: NASA, A. Fruchter and the ERO Team, STScl). (See color plates section.)

It is worth walking through the reason why the background galaxies in Fig. 1.7 are so elongated. A simple way to understand this is to remember that a source directly behind a lens leads to an Einstein ring, which is pretty close to what is observed in Fig. 1.7. Another way to think about why lenses make background objects more elliptical is to look at the cartoon in Fig. 1.8.

We are now in the middle of our progression across Fig. 1.3, with an extended source and lens. This is the region where both strong and weak lensing are important. Dramatic effects, characteristic of *strong lensing*, occur when $\Sigma > \Sigma_{\rm cr}$, while small distortions, or *weak lensing*, occur in the low surface density limit.

It is illuminating to understand the reason why this case, single extended source and lens, allows for both strong and weak lensing. As mentioned above, the critical surface density for an object at a cosmological distance is of order $\Sigma_{\rm cr} = 2 \times 10^{13} M_\odot/(100\text{kpc})^2$. As a simple example, let us assume that the density profile of a galaxy cluster is isothermal, so that

$$\rho(r) = \frac{\sigma^2}{2\pi G r^2}, \tag{1.15}$$

where σ is the velocity dispersion, which is typically of order 1000 km/sec. The surface density of this cluster is then (Exercise (1.5)):

Figure 1.8 Background galaxies appear elliptical due to the deflection by a foreground lens. Left panel shows background circular galaxy with mass clump denoted by dots between us and the background galaxy. Middle panel shows that the light that passes closest to the mass will be deflected most, so the ensuing image we observe will look like the right panel: elliptical.

$$\Sigma(R) = \frac{\sigma^2}{2RG}. \tag{1.16}$$

So $\Sigma(R)/\Sigma_{\rm cr} = 2\pi D_L(\sigma/c)^2/R$. Since $\sigma/c \sim 1/300$, this ratio is unity when $R \simeq 2\pi \times 10^{-5}$ Gpc or 60 kpc. For a cluster with radius of order a Mpc, then, only a small fraction of the total area will be in the region where $\Sigma(R) > \Sigma_{\rm cr}$; since area scales as R^2, that fraction will be $(60/1000)^2 \sim 0.004$. That means that, of all the background galaxies within a projected radius of 1 Mpc from the lens, fewer than one percent will be strongly lensed. All the others will be only weakly distorted by the foreground cluster. So, yes, there will be strong lensing phenomena, as is apparent in Fig. 1.7, but there will also be many, many galaxies in the weak lensing regime.

Another way of looking at this is to compute the total number of background galaxies that might be observed in the strong lensing regime. Taking the radius within which strong lensing occurs to be 60 kpc and a typical distance to a cluster to be 1000 Mpc, then the angular radius is a little more than $10''$. There are simply not that many galaxies in a region of this size. A final look at this is to think about adding the signal from every background galaxy, the signal being the tangential shear γ_t. The signal will typically be small compared to the noise, so the way to win is to beat down the noise: the more galaxies you measure, the smaller the noise becomes. We will encounter this often throughout the text, with the reduction in noise scaling as the square root of the number of objects sampled.