Line-of-Sight Effects on Strong Lensing Time Delays

Daniel Johnson

Supervisors:

Dr. Lucia Marchetti, *Deparment of Astronomy, University of Cape Town, South Africa.*
Dr. Pierre Fleury, *Institut de Physique Théorique, Centre National de la Recherche Scientifique, France.*
Dr. Julien Larena, *Laboratoire Univers et Particules de Montpellier, Université de Montpellier, France.*

A minor dissertation submitted in partial fulfilment of the requirements for the degree M.Sc. in the Department of Astronomy, as part of the National Astrophysics and Space Science Programme

Astronomy Department
University of Cape Town
South Africa
February 2023

Abstract

The effects of matter along the line-of-sight in strong gravitational lensing systems are widely recognised to have a non-negligible effect on lensing observables. In this project, we explore the dominant-lens approximation, a new theoretical formalism designed to parameterise these effects, paying particular attention to degeneracies between line-of-sight effects and the properties of the main lens, and the influence of these effects on measurements of strong lensing time delays. We implement this formalism in the tidal regime as a new sub-package in the open-source lens modelling software package `lenstronomy`. Using this implementation, we generate mock lens images and measured time delay data, and investigate the prospects of measuring properties of the line-of-sight from these data. We also explore the consequences of under-parameterised line-of-sight effects on measurements of the Hubble constant from time delay cosmography. We find that images of lensed extended sources are those from which precise measurements of line-of-sight shears are most likely to be obtained, but that lensed point source images and time delay data can serve as valuable constraints on noisier images. For measurements of both the shear and the Hubble constant, under-parameterised lens models which exclude the foreground shear term lead to significant biases.

Contents

Introduction

On the 29th of May 1919, Dyson, Eddington, Davidson and their collaborators first observed the lensing of distant starlight by the gravitational field of the Sun, sending shockwaves through the world, and enshrining Albert Einstein as a household name for the next century [49]. The deflection of light by gravity had first been hypothesised by Newton himself in *Opticks*, but now served to usher in a new scientific paradigm. The observed deflection angle was around double the Newtonian value, and agreed spectacularly with Einstein's seminal General Theory of Relativity, cementing its place as our most elegant and reliable model of the Universe at all but its smallest scales. Though Einstein and many others doubted the role gravitational lensing could play in the study of the cosmos, believing it to only be observable in rare and unelucidating chance alignments of nearby objects, others remained intrigued by its promise, and in 1937, Fritz Zwicky first proposed the possibility of discovering the lensing of and by galaxies, suggesting that their mass and size might be enough to detect the splitting of light into multiple lensed trajectories [183]. The world would have to wait more than four decades more for such an observation to be made, when the doubly-lensed quasar SBS 0957+561 was discovered by a team at the Kitt Peak Observatory in 1979 [168].

Since the discovery of SBS 0957+561, the number of known multiply-image gravitational lens systems has grown to several thousand, with this number expected to grow by orders of magnitude in the next decade [39], and gravitational lensing has evolved from a theoretical curiosity to a well-established and rapidly expanding field. Through the magnifying effects of curved spacetime, lensing acts as a cosmic telescope, bringing objects from the furthest reaches of space into view [90, 164, 30, 169]. However, this is not a telescope which can be pointed, and we must instead hope for chance alignments which bring these stars and galaxies into view. Rather, it is what a light ray can tell us about the path it took from these sources to our telescopes that makes lensing observations so compelling. All light, as it travels from distant galaxies, experiences some degree of magnification, stretching and distortion before it reaches our solar system, due to the presence of matter along its path and the large-scale structure through which it passes. This phenomenon is known as weak lensing, and has been used to probe these matter distributions via massive imaging surveys and statistical analyses of the apparent shapes of galaxies [15]. If however, the lensing mass is compact enough, and the source and lens are sufficiently aligned, light from a single source position can reach the observer via multiple paths through this curved spacetime, and, just as in the case of SBS 0957+561, multiple separate images of the source may be visible from our telescopes. Such a source is said to have been strongly lensed, offering one of the most dramatic observational consequences of Einstein's theories of gravity, space and light. As the path of a lensed light ray depends only on the mass distribution of a lens and the cosmology in which it is embedded, and is independent on the lens' luminosity or composition, it can serve as a powerful probe of dark matter, as well as allowing well-constrained measurements of cosmological parameters such as the Hubble constant, cosmological constant and density parameter. Weak lensing studies have provided competitive constraints on the density and distribution of dark and baryonic matter throughout the Universe [73, 74, 161, 7], while strong lensing studies have contributed significantly to our knowledge of the mass structure of galaxies and clusters (see [160, 13] and references therein).

In the early years of gravitational lensing observations, strong and weak lensing were treated as separate fields aiming to answer distinct (though often complementary) questions. The formalisms developed to explore these phenomena required simplifying assumptions to separate these domains, with strong lenses being modelled by an isolated lensing potential, usually neglecting the contributions from matter along the line-of-sight, while weak lensing was restricted to tidal perturbations by the line-of-sight, without scope for higher-order effects [23]. However, it is now unanimously accepted that the influence of weak lensing on strong lensing cannot be ignored [174, 175, 40, 67, 56]. In fact, due to the magnifying effects which accompany strong lensing, a strongly lensed galaxy is visible at a much greater redshift than an unlensed one, and thus light from such a source propagates through a longer line-of-sight. As

a consequence, weak lensing has the potential to imprint itself more strongly on strong lensing images than elsewhere. Gravitational shear effects, which stretch and distort images, can arise from satellites of the main lens galaxy (or neighbouring galaxies in its vicinity) [95], from structure along the line-of-sight between observer, lens and source [95, 85], or from general large scale structure [68, 84, 85, 142, 12]. These shears can be as strong as the shear produced by the lens itself [12], and a few percent of the total distortion caused by the main lens [136]. The resulting distortions affect not only the shape of the image, but can also increase the lensing cross-section, and therefore a selection bias exists for observations of strong lens images perturbed by significant weak lensing [78, 136]. Though its importance had often been underestimated [106], external shear has now been identified as the most promising explanation for the overabundance of quadruply-lensed quasar images with typical ellipticities [78].

As a consequence of these effects and others, weak lensing effects must be included in models of strong lenses, and failing to do so can often make fitting impossible [83]. Even when a measurement appears successful, ignoring the contribution from matter outside the main lens results in poor constraints and systematic biases [83, 109], as shear effects cannot simply be accounted for by adding additional parameters to the radial mass distribution [88]. It is also known that simple lens models which are consistent with image positions are often unable to reproduce the observed flux ratios, and dark matter substructure in the relevant regions of galaxy lenses cannot account for this discrepancy [107, 3, 38]. It has been argued that the solution can be found in the form of line-of-sight structures [179, 150].

The mathematics of line-of-sight effects on strong gravitational lensing has a long history, and a number of formalisms have been developed for this purpose. Multi-plane lensing facilitates an arbitrary number of lens planes placed at different redshifts, allowing a ray to be traced back recursively through multiple deflections [25]. A common alternative to this is to supplement the lens equation with tidal convergence and shear corrections, which are intended to contain the integrated effects of all the matter outside of the main lens plane [100, 137, 141, 12]. The multi-plane and tidal approaches have been combined by [108, 109], in which an arbitrary number of main lens planes (treated exactly) and tidal planes (having only linear order effects) are placed at different redshifts, by [23], in which the so-called "critical-sheet Born approximation" allows the contributions of line-of-sight structures to be evaluated along a path deflected by a critical mass sheet with the same Einstein radius as the true lens (thus removing the need for recursion), and by [136], in which the multi-plane formalism is generalised to include tidal effects between successive lens planes. This generalisation was taken a step further by [58], in which one or more lens planes are placed in an arbitrary spacetime. The balance between the generality which is sometimes required to reproduce an observation and account for all potential sources of error and the simplicity needed to allow a model to be used in practical contexts and to enable measurements of physically-relevant quantities is what has motivated these efforts, and in particular the dominant-lens approximation, the formalism employed in this project [59]. In this framework, the lens equation and associated calculations are reduced to functions of the observed image positions only by assuming that one lens plane dominates over all the others. This framework can be further simplified by assuming that the non-dominant lenses are tidal, but can also facilitate the inclusion of higher order effects.

The perturbative effects of other structures beyond the main lens plane represent not only an important piece of a model aiming to constrain properties of the main lens plane or the cosmology into which it is embedded, but are now understood to make a measurable contribution to lensing observables in their own right. Several examples exist of individual dark matter haloes near the main lens or along the line-of-sight being detected via their imprints on strong lensing images [166, 167, 44, 143, 120], and [45] has shown that such halos produce observably different effects on strong lensing images when located near the main lens versus along the line-of-sight. However, [71] have warned via the use of cosmological simulations that the effects of individual halos are harder to disentangle than is often assumed, and false positives and overestimations of mass are often possible.

More recently, it has been proposed that weak lensing variables representing the integrated effects of the line-of-sight and large scale structure (as opposed to individual halos) could be measurable from Einstein rings, the result of a strongly lensed and well-aligned extended source. This was first proposed in [23] as an independent calibration of weak lensing shear measurements, in which simulated data was used to argue for the validity of the approximations employed. The paper claimed that cosmic shear effects could be measured accurately with a ~ 0.003 precision from an HST-like dataset. The measurability of line-of-sight effects was further explored in [59], in which a more degeneracy-resilient approach was developed. This approach was tested on mock data in [77], in which it was demonstrated that line-of-sight shear can, in general, be measured without systematic biases, even when higher order effects are present. Strong higher order perturbations contribute significantly to the uncertainty budget of these measurements, but their accuracy is maintained [77]. The work of [23, 19] was applied to real data in

[101], in which strong lensing images were analysed in combination with weak lensing surveys from the same field to measure the external shear.

Although these early studies have convincingly demonstrated the possibility of using Einstein rings for shear measurements, doubts remain as to the extent to which results are truly independent of complexities the structure of the main lens, in particular when the main lens model is simplified [64]. It has been demonstrated that an additional shear axis can be generated both by external perturbations and by misalignments between the visible, baryonic component of a galaxies mass distribution and its dark matter halo [88, 57]. Recently, [31] have shown that, when fitting an elliptical power law lens model plus external shear to mock data generated with realistically complicated mass distributions, a spurious shear can be measured with strengths up to 8 %, and therefore argue that reported measurements of the external shear may in fact be measurements of deviations from elliptical symmetry in the main lens convergence, particularly when simple lens models are used.

While the above discussion has focused on the role of imaging data in strong lensing, another crucial source of insight into lens systems and cosmological observables is found in the time delays between strong lens images. As a light ray traces a path in the presence a gravitational potential, its trajectory bends to follow the spacetime curvature induced by this potential. A key consequence of this to the field of strong gravitational lensing is that the length of this path is intrinsically linked to this perturbed spacetime, and therefore to the gravitational potential of the lens. Without independent knowledge of the scale of a lensing system, measuring the length of the null geodesic along which a single ray propagated is impossible, as the observables of such a ray are dimensionless. However, when a fortuitous alignment results in the strong lensing of a compact, variable source, the role of time in gravitational lensing takes centre stage. When a quasar or a supernova is multiply-imaged, photometry of each of these images over the course of several months allows characteristic peaks and dips of their light curves to be compared, and thus for the relative time delays between these images to be determined. In recent studies, these have been measured with the tightest reported uncertainties on the order of 1% [27, 111, 102, 154, 37, 148].

The expansion of the Universe is neatly summarised in the Hubble parameter $H(t)$, which governs the rate of change of the scale factor a, and hence the recession velocity of distant galaxies. Under the standard model, this parameter depends on the matter content, dark energy and spatial curvature of the Universe, and is anchored by the Hubble constant H_0, the value of H today. Well-constrained measurements of H_0 would serve as a crucial piece of the puzzle that is the dark energy driven accelerating cosmological expansion, and are therefore considered amongst the most important goals of this precision era.

Conventional measurements of H_0 are broadly divided into "late Universe" measurements from the distance ladder, and "early Universe" measurements from temperature anisotropies in the cosmic microwave background. It is from these latter measurements that the most precise constraints have been obtained, with the Planck mission reporting a value of $H_0 = 67.36 \pm 0.54$ km.s^{-1}.Mpc^{-1}, a 0.7% uncertainty [6], consistent with measurements from baryon acoustic oscillations via the Dark Energy Survey [1]. Alternatives to these methods are offered by late Universe techniques, whereby Type 1a supernovae and Cepheid variables with known luminosities provide a direct measurement of cosmological distances, assuming that they are well-calibrated. Measurements from Cepheid variables have given a value of $H_0 = 74.03 \pm 1.42$ km.s^{-1}.Mpc^{-1}, a 1.91% uncertainty [128]. Thus, early and late Universe measurements disagree by 4.4σ, a discrepancy known as the Hubble tension, which is widely considered to pose one of the largest challenges to the standard model of modern cosmology [162].

The use of strong gravitational lensing as a novel measurement of H_0 was first proposed by Refsdal in 1964 [125]. While the positions and magifications of lensed images alone cannot constrain the scale of a lensing system, the product between the Hubble constant and the relative time delays between images $H_0\Delta t$ depends only on the lens model and the relative positions between which these delays are calculated, and thus a measurement of Δt from a single well-modelled system can theoretically provide a well-constrained H_0 measurement [125, 126]. This method is attractive for several reasons: firstly, because of the magnifying effects of lensing, time delay cosmography works with sources at much larger redshifts than standard candle methods. As the latter are generally found within $\sim$ 100 Mpc, their peculiar velocities are often comparible with their velocities due to the Hubble flow, whereas the Hubble flow dominates the observed velocities of more distant lensed systems. Additionally, while many other methods require the use of distance ladders, lensing is independent of all of these, and thus serves as a useful verifier of other methods.

Despite its promise, time delay cosmography is faced with serious challenges. Relative time delays are notoriously difficult and costly to measure, requiring observations over periods of at least several months, and are plagued by uncertainties introduced by the extrinsic variability of microlensing effects on the

signal (see e.g. [154]). The measured time delay is the sum of the cosmological and microlensing time delays, the latter of which can be non-negligible [156, 36].

An even greater obstacle to tightly-constrained measurements is posed by the fact that a given set of image positions, relative magnifications and relative time delays is consistent with an essentially infinite family of degenerate mass models of the lens and line-of-sight, from which different values of the Hubble constant would be inferred. The most notorious of these degeneracies is known as the mass-sheet transformation, which was first identified in [51], and has since been recognised as the dominant source of uncertainty in time delay cosmography (see [66, 133, 178, 105, 138, 22, 62, 18], amongst many others). The mass-sheet degeneracy is often artificially removed with the specific choice of lens model [138]. As noted by [94, 64], lens models force a relationship between the surface density at the Einstein ring and lensing observables. If such a profile has insufficient degrees of freedom, constraints on this surface density and on H_0 may by very precise, but will likely be inaccurate. The mean slope of early-type galaxies has been demonstrated to be very well-approximated by elliptical power law models, which are usually able to reproduce all lensing observables. However, there is little reason to think that this averaged slope should be a perfect description of the inner regions of galaxies where lensing occurs [138]. It has been demonstrated in [163, 127, 33] that the mass within the Einstein radius is divided fairly evenly between dark and baryonic matter, neither of which follow the isothermal or more general elliptical power law profiles often used to model them. [163] argues that the assumption of a power law likely cannot be extrapolated to smaller or larger radii than the Einstein radius of a lens, and [145] argues that a minimum of three degrees of radial freedom in a lens model are required to obtain measurements of the time delay distance (from which H_0 is derived) which are accurate to within a few percent. The prospects for alleviating this degeneracy from lensing observables without poorly-justified choices of lens models are bleak. As the errors introduced by the MSD are systematic, they will not be controlled by averaging over multiple systems [94]. [178] notes that, for a fixed external shear in quadruply-lensed point sources, H_0 depends on the radial mass index independently of this shear or other lens observables. Thus any agreement in H_0 values between multiple lenses offers little reason to trust the accuracy of these measurements, and could merely indicate that this mass index is fairly consistent across these lenses. Flux ratios or magnifications are similarly unable to alleviate the MSD, adding more parameters than constraints, and introducing a significant degree of complication in the form of non-linear functionals of the lensing potential to the modelling [178]. Finally, while some correlation between quasar variabilities and luminosities exists which can be used to estimate the absolute (rather than relative) magnifications of images, this relationship is broad, and thus could, at best, be used by averaging over a large number of sources [16].

In order to constrain the MSD, it is therefore necessary to obtain independent constraints on the mass profile of the lens galaxy. The most common method for doing so makes use of velocity dispersion measurements. At a given radius, the stellar velocity is a function of the mass contained within that radius, and thus by measuring Doppler shifts at different positions in a lens galaxy, the mass profile can be constrained. This was first employed in 1995 [112], and has been used in [130, 159, 131, 97, 149, 148, 144] and many others. With velocity dispersion data to limit the MSD, H_0 measurements with uncertainties below 5% have been reported [177, 27]. However, [138, 94] have called such measurements into question, pointing out that, since the square of the stellar velocity dispersion is proportional to the mass of the galaxy, which in turn is proportional to the Hubble constant, the percentage error on H_0 should be roughly double the uncertainty in this velocity dispersion. They therefore argue that, given the uncertainties present in dispersion data, constraints tighter than 10% require assumptions on the mass profile beyond what can be justified from velocity dispersion measurements. Motivated by these concerns, [19] obtained measurements of $H_0 = 74.5^{+5.6}_{-6.1}$ with lenses from the TDCOSMO survey and $H_0 = 67.4^{+4.1}_{-3.2}$ with lenses from both TDCOSMO and SLACS, in which mass profiles were constrained using kinematics only, representing an important step towards the often cited goal of constraining H_0 to within 1% [62]. Nonetheless, concerns remain about the use of stellar kinematics given the subtlety of the degeneracies present in lensing, and [151, 62] argue that a more sophisticated approach is necessary for trustworthy constraints on density slopes from these methods.

Though considered to be a smaller contributor to the uncertainty budget, it has nonetheless been recognised that shear effects can also significantly affect time delay measurements [83, 23], and excluding these effects from models will often lead to unsuccessful fits [109]. Although it is known that the quadrupole structure of the lensing potential has a large influence on time delays in quadruply-lensed systems, the ability of imaging data to tightly constrain this structure is sometimes taken for granted, leaving uncertainties in the convergence as the only significant source of error [93]. However, first order perturbations to a power law lens can result in zeroth-order errors in the time delays [124], and [138] demonstrates that a mass profile which deviates only slightly from the power law used to model it can

lead to a $> 10\,\%$ error on the inferred time delays. [31] demonstrated that applying the standard EPL plus external shear treatment to a realistically complicated lens system leads to a systematic underestimation of relative time delays by up to $\sim 20\,\%$. The shear can in fact also corrupt velocity dispersion data [24, 69], which could lead to systematic errors on convergence constraints. [83] further found that these effects are unlikely to be mitigated with observations of the immediate neighbours of a lens, as these perturbations are dominated by matter along the line-of-sight. As other contributions to the uncertainties in H_0 measurements are constrained more and more tightly, careful modelling of the influence of line-of-sight shear will play an increasingly important role in efforts to constrain H_0 to within $1\,\%$ [19].

[138] has argued that time delay ratios, in their sensitivity to the slope of the mass profile [66, 93], can be used to alleviate the degeneracy between simple power law models and their more realistic counterparts even more effectively than velocity dispersion data. The loss that these shear effects represents to precision in H_0 measurements may well be a gain for the measurement of the line-of-sight shear, as a perturbation to measured relative time delays may imprint a measurable deviation from a shear-free model, thus providing an opportunity for additional constraints on the shear.

The study of line-of-sight effects on strong gravitational lensing is therefore ripe for new scientific inquiry. Many of the theoretical foundations have been laid, but significant work remains to fine tune its application. As new surveys massively increase the dataset of strong lens image and time delay measurements [4, 182, 155, 158, 10], analysis of anything more than the smallest fraction of this dataset will require a much greater degree of automatisation, and a detailed understanding of both the promise held by different classes of systems for cosmography and the often subtle sources of bias and uncertainty when employing different models to reconstruct these systems. While the prospects for measuring the line-of-sight shear from Einstein rings have been given significant attention, the question remains as to the role that point sources and relative time delay measurements could play. Choices in how tidal line-of-sight effects are modelled have yet to be fully compared, and the extent to which simple parametric lens models can absorb the non-linear effects of external perturbers when attempting to measure the line-of-sight shear at first order remains to be seen. Finally, it is believed that the goal of a $1\,\%$ uncertainty H_0 measurement could be achieved if sources of systematic error can be tighly controlled [47, 104, 37]. While the question of the role of line-of-sight shears in this uncertainty budget has been posed, the errors introduced by their non-linear effects on time delays are not always fully appreciated, and investigating the systematics of this will be an important step towards the $1\,\%$ target. In a nutshell then, this project will seek to answer two basic questions: what is the influence of line-of-sight effects on measurements of strong lensing systems and their time delays, and to what extent can strong lensing systems and their time delays influence measurements of these line-of-sight effects?

This report is structured as follows: in chapter 1, we lay the theoretical foundations of this work, deriving the lens equation from Einstein's field equations, presenting the mathematics of lensing by a single deflector, and the time delays associated with such a lens. We go on to present the formalism of the Dominant Lens approximation in the tidal regime used to model weak lensing perturbations of a strong lensing image, explore the degeneracies present in these models and the methods used to alleviate them, and end by presenting the equations governing line-of-sight effects on strong lensing time delays. In chapter 2, we implement these equations into `lenstronomy`. In chapter 3, we investigate the possibility of measuring the line-of-sight shear from strong lensing images in the presence and absence of time delay data. Analysing mock datasets, we identify the classes of lensing systems which hold the most promise for shear measurements, the variables which are most consequential for the precision of these results, and the approaches to modelling these systems which best escape systematic biases. Finally, in chapter 4, we investigate the consequences of different parameterisations of line-of-sight effects on attempts to constrain the Hubble constant from strong lensing time delays.

Chapter 1

Theory

The field of strong gravitational lensing rests on extensively developed theoretical foundations, which exist at the intersection between gravitational physics, observational cosmology, optics, and galactic and extragalactic astrophysics. Fundamentally, the phenomenon of lensing is a direct consequence of the curvature introduced into null geodesics by the presence of matter. The distribution of this matter into vast filaments and voids, clumping into galaxies, groups and clusters, means that concentrations high enough to divert a single light beam along multiple trajectories are possible, and tracing these beams back through time and space reveals properties of this structure and the cosmological scaffolding onto which it is build which are otherwise hidden from view.

The theory of strong lensing is therefore built from the interactions between these fields, in ways which are sometimes explicit, but often subtle. In this chapter, we begin from first principles in section 1.1, demonstrating how the deflection angle experienced by a light ray travelling through a gravitational field in an otherwise-empty spacetime emerges as a consequence of Einstein's field equations, and solving the geodesic equation in this perturbed metric. This deflection angle, under certain key assumptions, is the cornerstone of the lens equation, which we present for the case of lensing by a single deflector. Associated with this equation are certain other key quantities from which the formalism of single plane lensing is constituted, and whose role in the theory of lensing we discuss. After introducing some of the fiducial mass profiles commonly used to model these lenses, we end the discussion of single plane lensing with some comments about the hidden influence of cosmological parameters on the lens equation, and the dependence of this influence on the choice of lens model. In section 1.2, we pivot from gravitational physics to optics, exploring the alternative conception of the lens equation offered by Fermat's Principle, not as the solution to the geodesic equation, but instead as the result of extremising the arrival time of the light ray. We present a derivation of the time delay experienced by a light ray in the presence of a single-plane gravitational lens, and explore the obfuscation of cosmological parameters within the expression of this time delay by certain key degeneracies in single-plane lensing.

In section 1.3, we generalise the work of section 1.1 to the case where an arbitrary number of thin lens planes are present between source and observer, where a recursive lens equation is necessary. Invoking the Dominant Lens approximation, assuming that the effects of one of these lenses dominates over the others, and then further assuming that the non-dominant lens have only tidal effects, we arrive at a lens equation which closely resembles that derived in section 1.1, but which is modified to account for the effects of matter distributed along the line-of-sight between the observer, lens and source, and discuss the degeneracies which appear in lensing observables between the source position, main lens, and line-of-sight effects. Finally, in section 1.4, we present the equations governing the time delays experienced by a lensed light ray in the presence of tidal line-of-sight effects. Throughout this report, we adopt the unit convention that the speed of light c is equal to 1.

1.1 Single plane lensing

1.1.1 Deriving the deflection angle

Strong lensing, expressed mathematically, is a transformation from the 2D observed coordinates $\boldsymbol{\theta}$ of a light ray seen by the observer in the image plane to the 2D coordinates $\boldsymbol{\beta}$ at which that same light ray would be seen in the absence of any deflectors. Key to this transformation is the deflection angle $\hat{\boldsymbol{\alpha}}$, which is the angular separation between the source and the projection of the image onto the source

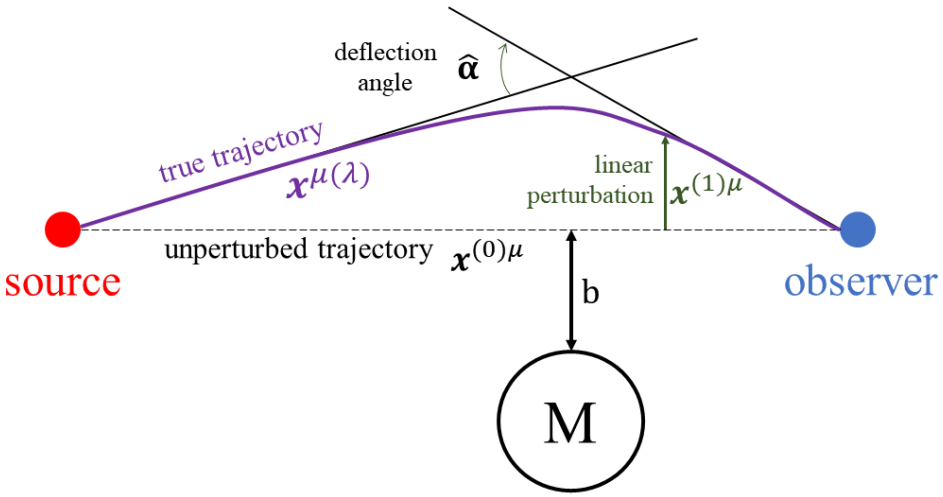

Figure 1.1: *A perturbed geodesic due to a mass M at nearest distance b from the unperturbed trajectory.. The true trajectory $\boldsymbol{x}^{\mu(\lambda)}$ follows an arced path, shaped by perpendicular linear perturbations along its length (shown in purple). To first order, it can be decomposed into the unperturbed trajectory $\boldsymbol{x}^{(0)\mu}$ and a linear perturbation $\boldsymbol{x}^{(1)\mu}$. The deflection angle is found by considering the overall angular deflection of these perturbations, as illustrated. [34]*

plane. Given an image position $\boldsymbol{\theta}$, a light ray can be traced back to its origin in the source plane $\boldsymbol{\beta}$ if the deflection angle $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ and the angular diameter distances D_{ds} and D_s are known. This angle relates to the gravitational field of the source, and can be derived by considering the lens as a perturbation to the background spacetime. In the following, we loosely follow the derivation found in [34].

The perturbed metric

Consider a Minkowski background spacetime perturbed by the presence of a static gravitational source, which we model as a perfect fluid with zero pressure (i.e. a cosmological "dust"). In the rest frame of this source, the 3-momentum of the source is zero, and so the energy-momentum tensor can be written as

$$T^{\alpha\beta} = \rho U^{\alpha} U^{\beta} = \begin{pmatrix} \rho c^2 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \tag{1.1}$$

where ρ is the density of the source and U^{α} its four-velocity. For the time being, we have assumed a Minkowski background, and the energy-momentum tensor for a source in motion can therefore be found by applying a Lorentz transformation to eq. (1.1). This limiting case breaks down when multiple sources with large relative velocities are present, and thus it is necessary to assume that the scale of the source is much less than the curvature scale or, more generally, all scales at which the velocity of the source might change appreciably. High relative velocities give rise to pressure terms, and thus the source can no longer be modelled as a dust, and $T^{\alpha\beta}$ changes. Nonetheless, it is reasonable to assume that the background spacetime is Minkowskian, as corrections due to the local curvature of the reference space-time are negligible on astrophysically-relevant scales [58].

If the above assumptions hold, and it is further assumed that the field is weak, the metric $g_{\mu\nu}$ can be decomposed into the metric of the background spacetime $g^{(0)}_{\mu\nu}$, plus a small perturbation $h_{\mu\nu}$

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| \ll 1. \tag{1.2}$$

Without loss of generality, this perturbation can be written as

$$h_{00} = -2\Phi, \tag{1.3}$$
$$h_{0i} = w_i, \tag{1.4}$$
$$h_{ij} = 2s_{ij} - 2\Psi\delta_{ij}, \tag{1.5}$$

where the h_{ij} have been expressed in terms of Ψ, from which the trace of h_{ij} is obtained, and s_{ij}, which is traceless.

$$\Psi = -\frac{1}{6}\delta^{ij}h_{ij}, \tag{1.6}$$
$$s_{ij} = \frac{1}{2}\left(h_{ij} - \frac{1}{3}\delta^{kl}h_{kl}\delta_{ij}\right). \tag{1.7}$$

If the background spacetime $g^{(0)}_{\mu\nu}$ is assumed to be Minkowskian (with metric $\eta_{\mu\nu}$), then we have

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = -(1+2\Phi)\mathrm{d}t^2 + w_i(\mathrm{d}t\mathrm{d}x^i + \mathrm{d}x^i\mathrm{d}t) + [(1-2\Psi)\delta_{ij} + 2s_{ij}]\mathrm{d}x^i\mathrm{d}x^j. \tag{1.8}$$

The Christoffel symbols for this metric are found according to

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2}g^{\sigma\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}), \tag{1.9}$$

which gives

$$\Gamma^0_{00} = \partial_0\Phi, \tag{1.10}$$
$$\Gamma^i_{00} = \partial_i\Phi + \partial_0 w_i, \tag{1.11}$$
$$\Gamma^0_{j0} = \partial_j\Phi, \tag{1.12}$$
$$\Gamma^i_{j0} = \partial_{[i}w_{j]} + \frac{1}{2}\partial_0 h_{ij}, \tag{1.13}$$
$$\Gamma^0_{jk} = -\partial_{(j}w_{k)} + \frac{1}{2}\partial_0 h_{jk}, \tag{1.14}$$
$$\Gamma^i_{jk} = \partial_{(j}h_{k)i} - \frac{1}{2}\partial_i h_{jk}. \tag{1.15}$$

To eliminate redundant degrees of freedom, we will adopt the transverse gauge, which is defined by

$$\partial_i s^{ij} = 0, \tag{1.16}$$
$$\partial_i w^i = 0. \tag{1.17}$$

Einstein's equations then become

$$G_{00} = 2\nabla^2\Psi = 8\pi G T_{00}, \tag{1.18}$$
$$G_{0j} = -\frac{1}{2}\nabla^2 w_j + 2\partial_0\partial_j\Psi = 8\pi G T_{0j}, \tag{1.19}$$
$$G_{ij} = (\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi - \Psi) - \partial_0\partial_{(i}w_{j)} + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij} = 8\pi G T_{ij}. \tag{1.20}$$

For static sources, all time-derivative terms vanish. Plugging in the expression for $T_{\mu\nu}$ from eq. (1.1), we have

$$\nabla^2\Psi = 4\pi G\rho, \tag{1.21}$$
$$\nabla^2 w_j = 0, \tag{1.22}$$
$$(\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi - \Psi) - \nabla^2 s_{ij} = 0. \tag{1.23}$$

The second and third equations feature gradients of the scalar terms w_j and s_{ij}. Consider such a field, f, in spherical coordinates. Solutions to the Laplace equation

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2} = 0, \tag{1.24}$$

take the general form

$$f(r,\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(A_{l,m} r^l + B_{l,m} r^{-l-1}\right) Y_{l,m}(\theta,\phi). \tag{1.25}$$

The $Y_{l,m}$ terms are the spherical harmonics. For any values of l and m, there will be some θ and ϕ which make $Y_{l,m} \neq 0$. We therefore consider the contributions of

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(A_{l,m} r^l + B_{l,m} r^{-l-1}\right). \tag{1.26}$$

In the limit $r \to \pm\infty$, this sum diverges, unless $A_{l,m} = 0$ for all $l > 0$. Similarly, in the limit $r \to 0$, the sum diverges, unless $B_{l,m} = 0$ for all values of l in the sum. Thus, for a field f such that $\nabla^2 f = 0$ (i.e. which satisfies Laplace's equation) to be non-singular and well behaved at infinity, the only remaining term must be constant in r. In order to recover a Minkowski spacetime at very large distances from the source (in other words, for the perturbation to be localised), in the limit as $r \to \infty$, the perturbation should disappear entirely, and so we set this constant term equal to zero. We will therefore impose that $\nabla^2 f = 0 \implies f = 0$.

Taking the trace and summing over the δ_{ij} in the third equation, since s_{ij} is traceless, we arrive at

$$2\nabla^2(\Phi - \Psi) = 0. \tag{1.27}$$

From which it follows that $\Phi = \Psi$. Einstein's equations therefore reduce to

$$\nabla^2 \Phi = 4\pi G\rho, \tag{1.28}$$
$$\nabla^2 w_j = 0, \tag{1.29}$$
$$\nabla^2 s_{ij} = 0, \tag{1.30}$$

with $w_j = s_{ij} = 0$. Substituting these results into our perturbed metric eq. (1.8),

$$\mathrm{d}s^2 = -(1+2\Phi)\mathrm{d}t^2 + (1-2\Phi)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2). \tag{1.31}$$

Solving the geodesic equation

Under the eikonal approximation, whereby it is assumed that any relevant length scales are much larger than the wavelength of the light beam [58], the path followed by photons through this perturbed metric can be found by solving the perturbed geodesic equation for null trajectories. We start by decomposing the null geodesic into a background path plus a perturbation

$$x^\mu(\lambda) = x^{(0)\mu}(\lambda) + x^{(1)\mu}(\lambda), \tag{1.32}$$

where λ is the affine parameter. $x^{(0)\mu}$ is the solution to the geodesic equation in the background (i.e. a null path in the Minkowski spacetime). Quantities in the perturbed metric are then evaluated along this background path to determine the perturbation to the path, $x^{(1)\mu}(\lambda)$. This approach requires that the potential Φ is approximately equal along the background and perturbed geodesics, i.e. $x^{(1)i}\partial_i\Phi \ll \Phi$. This can be ensured by considering very small paths, where the deviation $x^{(1)\mu}$ must be small, and then integrating these over the actual path $x^\mu(\lambda)$.

We introduce the unperturbed and perturbation wave vectors

$$k^\mu \equiv \frac{\mathrm{d}x^{(0)\mu}}{\mathrm{d}\lambda}, \tag{1.33}$$
$$l^\mu \equiv \frac{\mathrm{d}x^{(1)\mu}}{\mathrm{d}\lambda}. \tag{1.34}$$

The condition of a null path is

$$g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = 0. \tag{1.35}$$

Expanding this, with $eta_{\nu\nu}$, being the background Minkowsi metric,

$$\eta_{\mu\nu}k^\mu k^\nu + 2\eta_{\mu\nu}k^\mu l^\nu + \eta_{\mu\nu}l^\mu l^\nu + h_{\mu\nu}k^\mu k^\nu + 2h_{\mu\nu}k^\mu l^\nu + h_{\mu\nu}l^\mu l^\nu = 0. \tag{1.36}$$

Now, the null path condition must hold true for the zeroth (background) order in the absence of the perturbation terms, as in the unperturbed path, only this order exists. We therefore have $\eta_{\mu\nu}k^\mu k^\nu = 0$, or equivalently,

$$-(k^0)^2 + (k^i)^2 = 0, \tag{1.37}$$

allowing us to define the constant k

$$(k^0)^2 = (\vec{k})^2 \equiv k^2. \tag{1.38}$$

We are therefore able to eliminate the background contribution altogether, and can consider the next order (the first order) on its own. Here, we must have

$$2\eta_{\mu\nu}k^\mu l^\nu + h_{\mu\nu}k^\mu k^\nu = 0. \tag{1.39}$$

Using what we know about $\eta_{\mu\nu}$ and $h_{\mu\nu}$, as well as our definition of k,

$$-kl^0 + \vec{l}\cdot\vec{k} = 2k^2\Phi. \tag{1.40}$$

The geodesic equation is

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\rho\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda} = 0. \tag{1.41}$$

The Christoffel symbols are found by plugging in the previous results $w^i = 0$ and $h_{ij} = -2\Phi\delta_{ij}$ into the expression for the Christoffel symbols for the perturbed metric in eq. (1.15) before the adoption of the transverse gauge, as well as the condition that $\partial_0\Phi = 0$ (since we assume we have static potentials),

$$\Gamma^0_{00} = \Gamma^i_{j0} = \Gamma^0_{jk} = 0, \tag{1.42}$$

$$\Gamma^0_{0i} = \Gamma^i_{00} = \partial_i\Phi, \tag{1.43}$$

$$\Gamma^i_{jk} = \delta_{jk}\partial_i\Phi - \delta_{ik}\partial_j\Phi - \delta_{ij}\partial_k\Phi. \tag{1.44}$$

The Christoffel symbols are already first order in the perturbation, and so, at zeroth order, the geodesic equation is simply

$$\frac{\mathrm{d}^2x^{(0)\mu}}{\mathrm{d}\lambda^2} = 0, \tag{1.45}$$

which tells us that $x^{(0)\mu}$ is a straight trajectory through spacetime.

At first order, using the definitions of l^μ and k^μ,

$$\frac{\mathrm{d}l^\mu}{\mathrm{d}\lambda} = -\Gamma^\mu_{\rho\sigma}k^\rho k^\sigma. \tag{1.46}$$

Note that, since the Christoffel symbol is first order in the perturbation, no l^μ factors appear on the right hand side of this equation. We consider the temporal and spatial components separately. The $\mu = 0$ components are

$$\frac{\mathrm{d}l^0}{\mathrm{d}\lambda} = -2\partial_i\Phi k^i k^0, \tag{1.47}$$

$$= -2k(\vec{k}\cdot\vec{\nabla}\Phi), \tag{1.48}$$

and the $\mu = i$ components are

$$\frac{\mathrm{d}\vec{l}}{\mathrm{d}\lambda} = -\partial_i\Phi k^0 k^0 - (\delta_{jk}\partial_i\Phi - \delta_{ik}\partial_j\Phi - \delta_{ij}\partial_k\Phi)k^j k^k, \tag{1.49}$$

$$= -k^0k^0\partial_i\Phi - k^kk^k\partial_i\Phi - k^ik^j\partial_j\Phi - k^ik^k\partial_k\Phi, \tag{1.50}$$

$$= -2k^2\vec{\nabla}\Phi - 2(\vec{k}\cdot\vec{\nabla}\Phi)\vec{k}. \tag{1.51}$$

By our definition of k^μ, it is the wavevector of the null trajectory in the background, i.e. it points in a straight line along the background null geodesic. We can interpret the second term as the gradient of the gravitational potential Φ along the background path, multiplied by $2k^2$, and define

$$\vec{\nabla}_{||}\Phi = \frac{(\vec{k}\cdot\vec{\nabla}\Phi)\vec{k}}{k^2}. \tag{1.52}$$

We then define the gradient transverse to the path $\vec{\nabla}_\perp$ as the difference between the gradient and the gradient along the path, which gives us

$$\frac{\mathrm{d}\vec{l}}{\mathrm{d}\lambda} = -2k^2(\vec{\nabla}\Phi - \vec{\nabla}_{||}\Phi), \tag{1.53}$$

$$= -2k^2\vec{\nabla}_\perp\Phi. \tag{1.54}$$

Next, we'll determine a value of the temporal wave vector perturbation l^0 by integrating eq. (1.48)

$$l^0 = \int \frac{\mathrm{d}l^0}{\mathrm{d}\lambda}\mathrm{d}\lambda, \tag{1.55}$$

$$= -2k\int(\vec{k}\cdot\vec{\nabla}\Phi)\mathrm{d}\lambda, \tag{1.56}$$

$$= -2k\int\left(\frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda}\cdot\vec{\nabla}\Phi\right)\mathrm{d}\lambda, \tag{1.57}$$

$$= -2k\int\vec{\nabla}\Phi\cdot\mathrm{d}\vec{x}, \tag{1.58}$$

$$= -2k\Phi. \tag{1.59}$$

The constant of integration is fixed to be zero by setting $l^0 = 0$ when $\Phi = 0$, i.e. there is no perturbation when there is no gravitational potential. We can plug this into the first order expression for the null geodesic condition, eq. (1.40), to obtain

$$\vec{l}\cdot\vec{k} = 2k^2\Phi + kl^0 = 0. \tag{1.60}$$

From this, we see that $\vec{l}$ and $\vec{k}$, i.e. the spatial components of the wave vector perturbation and the original wave vector respectively, must be orthogonal. In other words, when a light ray is perturbed by a static and weak gravitational field, the spatial component of this perturbation occurs orthogonally to the original direction in which the ray was travelling.

The deflection angle

As discussed in section 1.1.3 and as shown in fig. 1.1, the deflection angle $\hat{\boldsymbol{\alpha}}$ is the deflection experienced by the original spatial wave vector $\vec{k}$ between source to the observer. At each point along this journey, the spatial wave vector experiences a linear perturbation $\vec{l}$. All these deflections sum to give the total perturbation from the background path $\Delta\vec{l}$, and so the deflection angle is

$$\hat{\boldsymbol{\alpha}} = -\frac{\Delta\vec{l}}{k}. \tag{1.61}$$

The total perturbation (i.e. the rotation of the spatial wave vector) is found by integrating the infinitesimal wave vector perturbations

$$\Delta\vec{l} = \int\frac{\mathrm{d}\vec{l}}{\mathrm{d}\lambda}\mathrm{d}\lambda, \tag{1.62}$$

$$= -2k^2\int\vec{\nabla}_\perp\Phi\mathrm{d}\lambda. \tag{1.63}$$

To obtain the deflection angle, we integrate over the physical spatial distance travelled using $\mathrm{d}s = k\mathrm{d}\lambda$, and find

$$\hat{\boldsymbol{\alpha}} = 2\int\vec{\nabla}_\perp\Phi\mathrm{d}s, \tag{1.64}$$

where the integral runs over all photon paths (i.e. over all past-directed null geodesics), and $s = k\lambda$ is the physical distance travelled by the photon. $\nabla_\perp$ is the two-dimensional gradient perpendicular to the background path. $\mathrm{d}\hat{\alpha} = 2\vec{\nabla}_\perp\Phi\mathrm{d}s$ can be thought of as an infinitesimal deflection over a short section of the path of the light ray and, adding up all these deflections (i.e. integrating along the path of the light ray), we obtain the total deflection angle.

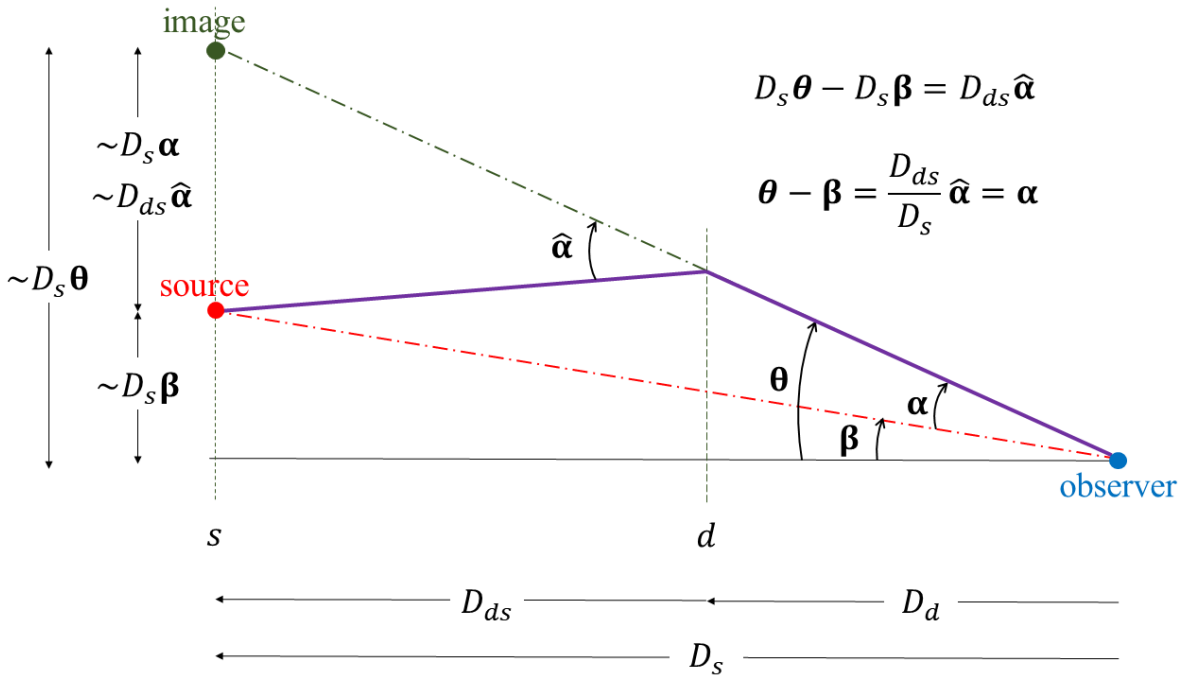

Figure 1.2: *The geometry of lens equation. The purple line shows the path followed by a light ray under the thin-screen approximation, while d and s refer to the deflector and source planes respectively. Light arrives at the observer at an angle* $\boldsymbol{\theta}$ *from the optical axis (the solid black line), which differs by an amount* $\boldsymbol{\alpha}$ *from the angle at which the light would arrive from the source in the absence of any deflectors,* $\boldsymbol{\beta}$*. The green and red dots show the projection of the image and source respectively onto the source plane. The difference between these projected positions allows the deflection angle* $\hat{\boldsymbol{\alpha}}$ *to be related to the reduced deflection angle (the 'displacement angle')* $\boldsymbol{\alpha}$*.*

1.1.2 The thin-screen approximation

The deflection angle in eq. (1.64) is the integrated effect of the perturbation to the null geodesic which traces out the path of a light ray through spacetime in the presence of a gravitational field Φ. Most of the deflection takes place in a region $\Delta z \sim \pm b$, where the impact parameter b is the minimum distance between the light ray and the centre of the lens. In the context of lensing by galaxies or clusters, the distances between observer, lens and source are typically orders of magnitude larger than Δz. The mathematics of strong lensing by a single lens can therefore be greatly simplified by treating the deflection as taking place at in a single plane. The actual three-dimensional mass distribution of the lens is projected into the so-called 'lens plane', orthogonal to the line-of-sight. Points within this plane are described by the two-dimensional vector $\boldsymbol{\xi}$ (measured from the optical axis), and the plane's surface density can be related to the density $\rho(\boldsymbol{\xi}, z)$ at redshift z and position $\boldsymbol{\xi}$ via

$$\Sigma(\boldsymbol{\xi}) = \int \rho(\boldsymbol{\xi}, z)\mathrm{d}z. \tag{1.65}$$

The deflection angle eq. (1.64) can be rewritten in terms of these quantities as

$$\hat{\boldsymbol{\alpha}} = 4G \int \frac{(\boldsymbol{\xi} - \boldsymbol{\xi}')\Sigma(\boldsymbol{\xi}')}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}\mathrm{d}^2\boldsymbol{\xi}'. \tag{1.66}$$

Under the thin-screen approximation for single-plane lensing, the light ray is imagined to travel along a straight line from the source plane to the lens plane, be instantaneously deflected by an angle $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ as it passes through the lens plane at $\boldsymbol{\xi}$, and then travel along a straight line from the lens plane to the observer.

1.1.3 The lens equation

From the definition of the angular diameter distance and assuming the angle is small, the physical separation between the source and the projection of the image onto the source plane is $\hat{\boldsymbol{\alpha}}D_{ds}$, where D_{ds}

is the angular diameter distance from the deflector to the source. However, as can be seen in fig. 1.2, this is the same as the separation between the observed image position and the origin of the coordinate system, less the separation between the source and the origin,

$$\boldsymbol{\theta} D_s - \boldsymbol{\beta} D_s = \hat{\boldsymbol{\alpha}} D_{ds}, \tag{1.67}$$

where D_s is the angular diameter distance from the observer to the source. Dividing throughout by this distance, we obtain

$$\boldsymbol{\theta} - \boldsymbol{\beta} = \frac{D_{ds}}{D_s} \hat{\boldsymbol{\alpha}}. \tag{1.68}$$

The difference between $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ is known as the the displacement angle $\boldsymbol{\alpha}$. From eq. (1.68), we see that

$$\boldsymbol{\alpha} = \frac{D_{ds}}{D_s} \hat{\boldsymbol{\alpha}}. \tag{1.69}$$

Equation (1.68), more typically written as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.70}$$

is the lens (or "ray-tracing") equation, which is the transformation from the image plane to the source plane. When strong lensing occurs, multiple image positions $\boldsymbol{\theta}$ correspond to a single source position $\boldsymbol{\beta}$.

1.1.4 The lensing potential and the shear matrix

The lensing potential

From eq. (1.64), we can write the displacement angle as

$$\boldsymbol{\alpha} = \frac{D_{ds}}{D_s} \hat{\boldsymbol{\alpha}} = 2 \frac{D_{ds}}{D_s} \int \vec{\nabla}_\perp \Phi \mathrm{d}s. \tag{1.71}$$

Thus, given that $\boldsymbol{\alpha}$ is found by integrating the gradient of the scalar field Φ, it must be curl-free, and can therefore be written in terms of a scalar potential

$$\boldsymbol{\alpha} = \nabla \psi. \tag{1.72}$$

The quantity ψ is two-dimensional, and is known as the lensing potential. These arguments are only valid for single-plane lensing, as multiple lenses produce a non-zero curl. This curl can be interpreted as the rotation of images and lens-lens coupling, thus a departure from the Born approximation. As such, the overall lensing potential cannot simply be calculated as the sum of the lensing potentials of each lens.

In the single lens case, this lensing potential is given by

$$\psi(\boldsymbol{\theta}) = \frac{D_{ds}}{D_d D_s} 2 \int \Phi(D_d \boldsymbol{\theta}, z) \mathrm{d}z. \tag{1.73}$$

The gradient of this potential gives the deflection angle:

$$\vec{\nabla}_\theta \psi = D_d \vec{\nabla}_\xi \psi = 2 \frac{D_{ds}}{D_s} \int \vec{\nabla}_\perp \Phi \mathrm{d}z = \boldsymbol{\alpha}, \tag{1.74}$$

which follows from eq. (1.71) and the fact that, assuming the deflection angle is small, we can integrate along the path of the undeflected light ray (and so $\mathrm{d}s$ becomes $\mathrm{d}z$). $\vec{\nabla}_\xi$ is the gradient with respect to the transverse coordinate ξ in the lens plane.

The Jacobian matrices of the lens mapping

The Jacobian matrix of the lens map eq. (1.70), typically known as the amplification matrix, is given by

$$\mathcal{A}_{ij} \equiv \frac{\partial \beta^i}{\partial \theta^j} = \delta_{ij} - \frac{\partial \alpha^i}{\partial \theta^j} = \delta_{ij} - \psi_{ij}, \tag{1.75}$$

where

$$\frac{\partial^2 \psi(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \equiv \psi_{ij}, \tag{1.76}$$

and, in this context, i and j refer to the component of a matrix or vector. Closely related to this matrix is the Jacobi matrix of the lens map $\boldsymbol{\theta} \mapsto \boldsymbol{\eta}(\boldsymbol{\theta})$, where $\boldsymbol{\eta}$ is a dimensional position in the source plane. This matrix is defined as

$$\boldsymbol{\mathcal{D}}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}}. \tag{1.77}$$

$\boldsymbol{\mathcal{D}}$ relates the morphology of an infinitesimal image patch $\mathrm{d}^2\boldsymbol{\theta}$ observed in direction $\boldsymbol{\theta}$ to the morphology of the corresponding source patch $\mathrm{d}^2\boldsymbol{\eta}$. It is related to the amplification matrix via [56]

$$\boldsymbol{\mathcal{A}} = D_s^{-1}\boldsymbol{\mathcal{D}}. \tag{1.78}$$

More generally, the Jacobi matrix can be defined along the path of the light ray via the affine parameter λ as

$$\boldsymbol{x}(\lambda) = \boldsymbol{\mathcal{D}}(\lambda)\frac{\mathrm{d}\boldsymbol{x}(0)}{\mathrm{d}\lambda} + \boldsymbol{x}(0), \tag{1.79}$$

where $\boldsymbol{x}(\lambda)$ is the screen-space separation between the physical and fiducial rays at λ. Within any thin screen along the light-ray's path, it is the separation between the actual ray and the fiducial ray.

$\boldsymbol{\mathcal{D}}(\lambda)$ therefore relates the value of $\dot{\boldsymbol{x}}$ at the position of the observer, to the value of $\boldsymbol{x}(\lambda)$ at some screen at λ (looking back along the ray). We can assume that the actual and fiducial rays converge at the observer position $\lambda = 0$, such that $\boldsymbol{x}(0) = 0$ and

$$\boldsymbol{x}(\lambda) = \boldsymbol{\mathcal{D}}(\lambda)\frac{\mathrm{d}\boldsymbol{x}(0)}{\mathrm{d}\lambda}. \tag{1.80}$$

The derivative of the separation with respect to the affine parameter at the observer position is related to the angular separation of the physical and fiducial rays (or in other words, the angular coordinates of the physical ray) via $\dot{\boldsymbol{x}}(0) = \omega_0\boldsymbol{\theta}$.

If any two rays of the reference space-time emerge from, or converge to, a point in plane a with angular separation $\boldsymbol{\theta}_a$, then their transverse separation $\boldsymbol{x}_b$ in another plane (b) reads

$$\boldsymbol{x}_b = \boldsymbol{\mathcal{D}}_{ab}\dot{\boldsymbol{x}}_a = \boldsymbol{\mathcal{D}}_{ab}\omega_a\boldsymbol{\theta}_a, \tag{1.81}$$

where ω_a is the fiducial photon's cyclic frequency as measured in the same rest frame where $\boldsymbol{\theta}_a$ was defined, and $\boldsymbol{x}_i$ is the value of $\boldsymbol{x}$ in plane i. $\boldsymbol{\mathcal{D}}_{ab}$ simply governs how light rays with angular separation $\boldsymbol{\theta}_a$ at point a will be separated at point $\boldsymbol{x}_b$, as a function of their separation. It is therefore closely related to the amplification matrix, which relates a small image-position change to a small source-position change, $\boldsymbol{\mathcal{A}} = \mathrm{d}\boldsymbol{\beta}/\mathrm{d}\boldsymbol{\theta}$. If the reference space-time can be treated as a weakly perturbed homogeneous isotropic FLRW model, then

$$\omega_a\boldsymbol{\mathcal{D}}_{ab} = D_{ab}\boldsymbol{\mathcal{A}}_{ab}. \tag{1.82}$$

Convergence

The convergence of a lens is defined as the dimensionless projected surface-mass density of the lens in units of the critical density

$$\kappa(\boldsymbol{\theta}) \equiv \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_\mathrm{c}}, \tag{1.83}$$

where the critical density Σ_c is given by

$$\Sigma_\mathrm{c} = \frac{D_s}{4\pi G D_d D_{ds}}. \tag{1.84}$$

Poisson's equation allows us to relate the projected surface-mass density to the Laplacian of the gravitational field Φ

$$\nabla_\xi^2\Phi = 4\pi G\rho, \tag{1.85}$$

thus the Laplacian of the lensing potential is simply

$$\nabla_\theta^2\psi = 2\frac{D_d D_{ds}}{D_s}\int \nabla_\xi^2\Phi \mathrm{d}z = 2\frac{D_d D_{ds}}{D_s}4\pi G\Sigma = 2\frac{\Sigma(\boldsymbol{\theta})}{\Sigma_\mathrm{c}} \equiv 2\kappa(\boldsymbol{\theta}). \tag{1.86}$$

Inverting the above, we can express the lensing potential and displacement angle in terms of the convergence,

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta}')\ln|\boldsymbol{\theta} - \boldsymbol{\theta}'|\mathrm{d}^2\theta', \tag{1.87}$$

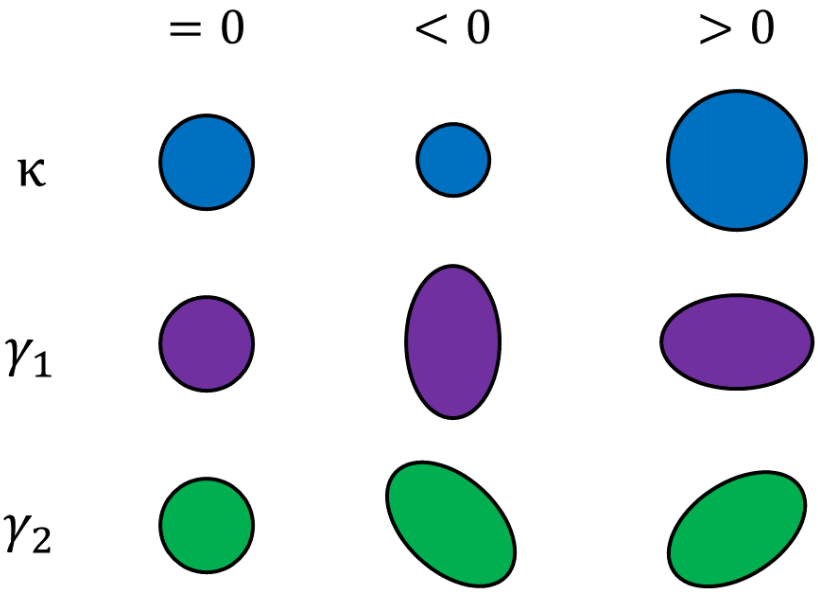

Figure 1.3: *The effects of convergence and shear on a circular image. The convergence κ causes a uniform change to the size of an image, while the shears γ_1 and γ_2 stretch and distort the image.*

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|}\mathrm{d}^2\theta'. \tag{1.88}$$

Using eq. (1.86) and the definition in eq. (1.76), the convergence can be written as

$$\kappa = \frac{1}{2}\nabla^2_\theta\psi = \frac{1}{2}(\psi_{11}+\psi_{22}) = \frac{1}{2}\mathrm{tr}\;\psi_{ij}. \tag{1.89}$$

At first order, the convergence determines the magnification of images.

Shear

The shear associated with a strong lensing transformation is a dimensionless quantity which describes the local distortion of lensed images. There are two shear components, γ_1 and γ_2, which are given by

$$\gamma_1(\boldsymbol{\theta}) = \frac{1}{2}(\psi_{11}-\psi_{22}) \equiv \gamma(\boldsymbol{\theta})\cos\left[2\phi(\boldsymbol{\theta})\right], \tag{1.90}$$

$$\gamma_2(\boldsymbol{\theta}) = \psi_{12} = \psi_{21} \equiv \gamma(\boldsymbol{\theta})\sin\left[2\phi(\boldsymbol{\theta})\right], \tag{1.91}$$

where

$$\gamma = (\gamma_1^2+\gamma_2^2)^{\frac{1}{2}}, \tag{1.92}$$

and ϕ is the angle between the displacement angle $\boldsymbol{\alpha}$ and the x-axis.

The reduced shear is defined component-wise as

$$g_1 = \frac{\gamma_1}{1-\kappa}, \tag{1.93}$$

$$g_2 = \frac{\gamma_2}{1-\kappa}, \tag{1.94}$$

or as a single term as $g = \gamma/(1-\kappa)$.

The shear matrix

As a matrix of derivatives, the amplification matrix $\boldsymbol{\mathcal{A}}$ is symmetric, and can be decomposed into a diagonal matrix (which contains the trace), and a symmetric, trace-free matrix (which contains the off-diagonal symmetric component and the trace-free diagonal element). These three degrees of freedom are,

respectively, the convergence κ and the two shear components γ_1 and γ_2.

$$\mathcal{A}_{ij} = \delta_{ij} - \psi_{ij}, \tag{1.95}$$

$$\kappa = \frac{1}{2}(\psi_{11} + \psi_{22}), \tag{1.96}$$

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}), \equiv \gamma \cos(2\phi)\,, \tag{1.97}$$

$$\gamma_2 = \psi_{12} = \psi_{21} \equiv \gamma \sin(2\phi)\,. \tag{1.98}$$

Given the above, we can rewrite the transformation matrix as

$$\boldsymbol{\mathcal{A}} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{1.99}$$

$$= \begin{pmatrix} 1-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} - \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}, \tag{1.100}$$

$$= (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}. \tag{1.101}$$

As we have seen, the lens equation maps the position $\boldsymbol{\theta}$ in the image plane to the position $\boldsymbol{\beta}$ in the source plane, via the scaled deflection angle $\boldsymbol{\alpha}$. Given a light ray originating from a point $\boldsymbol{\theta}$ in the image plane, it determines the point $\boldsymbol{\beta}(\boldsymbol{\theta})$ where that specific point would have originated in the source plane. As the matrix of derivatives of this transformation, $\boldsymbol{\mathcal{A}}$ describes how an infinitesimal change in the image position affects the corresponding source position. $\boldsymbol{\mathcal{A}}(\boldsymbol{\theta})$ gives a linear approximation to the transformation of the area of points surrounding a specific point of interest $\boldsymbol{\theta}$ by the lens mapping. It describes the local isotropic magnification of a source (from the convergence κ to first order) and its distortion (from the shear components γ_1 and γ_2).

The amplification matrix $\boldsymbol{\mathcal{A}}$ is often rewritten in terms of the shear matrix $\boldsymbol{\Gamma}$,

$$\boldsymbol{\mathcal{A}} = \mathbf{1} - \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\theta}} = \mathbf{1} - \boldsymbol{\Gamma}, \tag{1.102}$$

where

$$\boldsymbol{\Gamma} = \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & \kappa-\gamma_1 \end{pmatrix}. \tag{1.103}$$

We also introduce the reduced shear matrix,

$$\boldsymbol{G} = \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \tag{1.104}$$

allowing the amplification matrix to be rewritten as

$$\boldsymbol{\mathcal{A}} = (1-\kappa)\boldsymbol{G}. \tag{1.105}$$

Magnification

The inverse matrix of $\mathcal{A}$,

$$\boldsymbol{\mathcal{A}}^{-1} = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\beta}} = \boldsymbol{\mathcal{M}}, \tag{1.106}$$

is the Jacobian of the transformation from the source plane to the image plane, and is known as the magnification matrix.

Now, the determinant of any linear transformation gives the factor by which the length/volume/area is stretched out. The Jacobian is the linear approximation to the lens equation transformation, and so its determinant, to linear order (or to the point source limit), is the factor by which areas are stretched. If the Jacobian determinant of f at the point p is positive, f preserves orientation, whereas if it is negative, it reverses orientation.

The magnification of an image is defined as the ratio of its area to that of the source producing the image. In the point source limit, i.e. very close to point, where the transformation is effectively linear

and described by the Jacobian, the local magnification μ is given by the determinant of the magnification tensor,

$$\mu = |\boldsymbol{\mathcal{M}}| = \frac{1}{|\boldsymbol{\mathcal{A}}|} = \frac{1}{|(1-\kappa)^2 - \gamma^2|}. \tag{1.107}$$

The magnification is determined by the convergence κ only to the first order, i.e. when κ and γ are small. For extended sources, the observed magnification depends on the surface brightness distribution of the source as well as the magnification matrix.

When the determinant of the inverse magnification matrix vanishes, the magnification becomes formally infinite. The loci of formally infinite magnification in the image plane are referred to as **critical lines**, while their corresponding loci in the source plane are known as **caustics**.

Compact sources located close to a caustic can be magnified by very large factors, up to two orders of magnitude, although the total observed flux is always finite for astrophysical sources of finite angular size.

If the background spacetime is Minkowskian, a lensed image cannot be demagnified relative to its unlensed counterpart. However, if the background ray is defined in an FLRW spacetime, negative density planes are possible, and the background ray may therefore be more magnified than a lensed ray travelling through these negative density planes [141]. This is discussed in more detail in section 1.3.1.

Key to understanding the magnification of images by gravitational lenses is the conservation of surface brightnesses. While the apparent solid angle of the source changes from source to observer, the flux received per unit solid angle remains constant. Thus the apparent brightness increases in proportion to the size of the image. This is a consequence of Etherington's reciprocity law.

Etherington's Reciprocity Law and conservation of surface brightness

Etherington's reciprocity law relates the Jacobi matrix from eq. (1.77) to the matrix of the reverse mapping,

$$\boldsymbol{\mathcal{D}}(\lambda_1 \leftarrow \lambda_2) = -\boldsymbol{\mathcal{D}}^\top(\lambda_2 \leftarrow \lambda_1), \tag{1.108}$$

which has the consequence that

$$\det \boldsymbol{\mathcal{D}}(\lambda_1 \leftarrow \lambda_2) = \det \boldsymbol{\mathcal{D}}(\lambda_2 \leftarrow \lambda_1). \tag{1.109}$$

Now, the number of photons emitted during a short time interval $\delta\tau_s$ in a source frame which can be received by the observer's detector is

$$\delta N = \frac{L_s}{\hbar\omega_s} \times \delta\tau_s \times \frac{\Omega_s}{4\pi}. \tag{1.110}$$

The first term $L_s/\hbar\omega_s$ is the conversion between the luminosity of the source (in units of energy) to the number of photons emitted per unit time by the source (since a photon emitted at frequency ω_s will have energy $\hbar\omega_s$). Multiplying this term by the time interval $\delta\tau_s$ gives the number of photons emitted in that time interval, and finally multiplying this by the ratio of Ω_s, the solid angle in the source frame of the light beam which is intercepted by the observer, to the solid angle of the entire sky 4π, gives δN. In the last step, we have assumed that the light is emitted isotropically.

The number of photons received by the observer is

$$\delta N = \frac{F_o}{\hbar\omega_o} \times \delta\tau_o \times A_o. \tag{1.111}$$

F_o is the observed flux, ω_o is the observed frequency, and so the first term converts the observed intensity (energy per unit time per unit area) into a number of photons per unit time per unit area. To convert this quantity into the actual number of photons received by the detector, it must be multiplied by τ_o, the corresponding interval in the observer's frame in which that fraction of photons is received, and A_o, the area of the detector.

The ratio $\delta\tau_o/\delta\tau_s$ is the same as the ratio between the observed and emitted periods of the light signal, which in turn is the inverse ratio of their frequencies, ω_s/ω_o. This, in turn, is simply the redshift $1+z$.

$$\frac{\delta\tau_o}{\delta\tau_s} = \frac{\omega_s}{\omega_o} = 1 + z. \tag{1.112}$$

If the the total number of photons is assumed to be conserved, and if none are absorbed along their trajectories, we can set eqs. (1.110) and (1.111) equal, and see that

$$\frac{L_s}{4\pi F_o} = \frac{\omega_s}{\omega_o} \times \frac{\delta\tau_o}{\delta\tau_s} \times \frac{A_o}{\Omega_s}, \tag{1.113}$$

$$= (1+z)^2 \frac{A_o}{\Omega_s}. \tag{1.114}$$

The luminosity distance is defined in terms of the source luminosity L_s and observed flux F_o as

$$D_{\rm L} = \sqrt{\frac{L_s}{4\pi F_o}}, \tag{1.115}$$

from which it follows that

$$D_{\rm L} = (1+z)\sqrt{\frac{A_o}{\Omega_s}}. \tag{1.116}$$

The determinant of the Jacobi matrix (modulo a frequency factor, which is needed to convert from the frame-independent affine parameter to physical distances in the relevant frame) is the Jacobian of the transformation from the observed angular position of the source to its physical position in the source plane, and thus gives the ratio between the physical area of the source A_s and its observed solid angle Ω_o,

$$\det[\omega_o \boldsymbol{\mathcal{D}}(\lambda_s \leftarrow \lambda_o)] = \frac{A_s}{\Omega_o} \equiv D_{\rm A}^2, \tag{1.117}$$

where λ is the affine parameter along null geodesics and $D_{\rm A}$ is the angular diameter distance. Note that, since ω_o is a constant which multiplies the matrix $\boldsymbol{\mathcal{D}}$, and since $\boldsymbol{\mathcal{D}}$ is a 2×2 matrix, we can write $\det(\omega_o \boldsymbol{\mathcal{D}}) = \omega_o^2 \det \boldsymbol{\mathcal{D}}$.

In similar fashion, the ratio of A_o to Ω_s is related to the Jacobian of the reverse mapping,

$$\det\left[\omega_s \boldsymbol{\mathcal{D}}(\lambda_o \leftarrow \lambda_s)\right] = \frac{A_o}{\Omega_s}. \tag{1.118}$$

Substituting this expression into eq. (1.116) and dividing on both sides by $D_{\rm A}$ (as given in eq. (1.117)), the ratio of the luminosity distance to the angular diameter distance is

$$\frac{D_{\rm L}}{D_{\rm A}} = (1+z)\frac{\omega_s}{\omega_o}\frac{\sqrt{\det \boldsymbol{\mathcal{D}}(\lambda_o \leftarrow \lambda_s)}}{\sqrt{\det \boldsymbol{\mathcal{D}}(\lambda_s \leftarrow \lambda_o)}}. \tag{1.119}$$

From Etherington's reciprocity law in eq. (1.109) and the definition of the redshift in eq. (1.112), we arrive at the distance duality relation

$$D_{\rm L} = (1+z)^2 D_{\rm A}. \tag{1.120}$$

With this result in mind, we turn to the relationship between the emitted and observed intensities, I_s and I_o, which are given by the fluxes divided by the solid angle into which those fluxes are emitted,

$$I_o = \frac{F_o}{\Omega_o}\frac{\Omega_s}{F_s} I_s. \tag{1.121}$$

Furthermore, the flux from the source which could be received by the observer's detector is given by multiplying the total flux from the source by the fractional solid angle into which the flux received by the observer is emitted,

$$F_s = \frac{L_s}{A_s}\frac{\Omega_s}{4\pi}. \tag{1.122}$$

Substituting eq. (1.122) into eq. (1.121),

$$I_o = \frac{A_s}{\Omega_o}\frac{4\pi F_o}{L_s} I_s. \tag{1.123}$$

The definitions of $D_{\rm L}$ in eq. (1.115) and of $D_{\rm A}$ in eq. (1.117) are easily recognised in the above, and so we can write

$$I_o = \frac{D_{\rm A}^2}{D_{\rm L}^2} I_s. \tag{1.124}$$

Finally, from the distance duality relationship in eq. (1.120), we can write

$$I_o = \frac{1}{(1+z)^4} I_s. \tag{1.125}$$

Equation (1.125) tells us that I_o, the intensity received, relates to the intensity of the source only via the redshift, and the area distance plays no part. This is a direct consequence of Etherington's reciprocity law and the definitions of D_{L} and D_{A}, and is valid in any Riemann spacetime in which photons travel along null geodesics and are conserved, and in which a global scale factor can be defined.

Stated differently, the intensity received by an observer depends only on the source intensity and the redshift to the source, regardless of distortions to the shape and size of an image resulting from the passage of the light beam through a curved spacetime. Thus, if an image is magnified, the integrated flux received will scale with the increase in the size of the image.

1.1.5 Simple lens models

Real gravitational lenses are highly complicated gravitational fields arising from the contributions of the baryonic and dark matter components of a galaxy, coupled with the effects of additional structure in the vicinity of the galaxy, matter along the line-of-sight, and the large scale structure of the Universe. The resulting lensing potential is, in its exact form, irreducibly complex, and these different contributions are, in many cases, highly degenerate with one another. Simple parametric lens models, while failing to capture this complexity, have nonetheless been used to great effect to reproduce key lensing observables. These models must be employed with a healthy dose of caution, and conclusions drawn from them should be treated similarly. In their simplicity, however, analytical solutions to important calculations involving these models are often possible, and they therefore offer a fertile source of insight into a wide array of lensing phenomena. When modelling real lenses, hierarchical models are often adopted, with simple models forming a foundation by which key first order characteristics of a lens are determined, and onto which successive degrees of complexity are added.

In this subsection, after introducing some alternative notation, we discuss some of the simplest parametric lens models, in particular those which will be used in subsequent chapters. Calculations not performed explicitly in this subsection can be be found in appendix A.1.

Complex notation

Up to this point, we have worked in vector notation, which is typically chosen as an intuitive representation of quantities the two-dimensional plane of the sky. While this offers an intuitive picture, many vector calculations are greatly simplified by mapping the x and y components of vector quantities to the real and imaginary components respectively of complex numbers. In what follows, $\underline{a}$ will refer to the complex counterpart of the vector $\boldsymbol{a}$, and $\underline{a}^*$ to its complex conjugate. Under this transformation, following Section 2.3.2 of [59],

$$\boldsymbol{\theta} = \theta_1 \boldsymbol{e}_1 + \theta_2 \boldsymbol{e}_2 \mapsto \underline{\theta} = \theta_1 + \mathrm{i}\theta_2. \tag{1.126}$$

When a matrix $\boldsymbol{\Gamma}$ is parameterised according to eq. (1.246), the result of multiplying a vector $\boldsymbol{u}$ by $\boldsymbol{\Gamma}$ in complex notation is simply

$$\boldsymbol{v} = \boldsymbol{\Gamma}\boldsymbol{u} \mapsto \underline{v} = (\kappa + \mathrm{i}\omega)\underline{u} + \gamma \underline{u}^*. \tag{1.127}$$

A key advantage of the complex formalism is that, rather than being treated as functions of the components θ_1 and θ_2, a complex number $\underline{a}$ is treated as a function of the position $\underline{\theta}$ and its conjugate $\underline{\theta}^*$. Partial derivatives with respect to these variables take the form

$$\frac{\partial}{\partial \underline{\theta}} \equiv \frac{1}{2}\left(\frac{\partial}{\partial \theta_1} - \mathrm{i}\frac{\partial}{\partial \theta_2}\right), \tag{1.128}$$

$$\frac{\partial}{\partial \underline{\theta}^*} \equiv \frac{1}{2}\left(\frac{\partial}{\partial \theta_1} + \mathrm{i}\frac{\partial}{\partial \theta_2}\right). \tag{1.129}$$

This must be taken into account when Taylor expanding,

$$\underline{\alpha}(\underline{\theta} + \delta\underline{\theta}) = \underline{\alpha}(\underline{\theta}) + \frac{\partial \underline{\alpha}}{\partial \underline{\theta}}\delta\underline{\theta} + \frac{\partial \underline{\alpha}}{\partial \underline{\theta}^*}\delta\underline{\theta}^* + \mathcal{O}\left(\delta\theta^2\right), \tag{1.130}$$

and when applying the chain rule,

$$\frac{\partial}{\partial \theta}\left\{\underline{\alpha}[\beta(\underline{\theta})]\right\} = \frac{\partial \underline{\alpha}}{\partial \underline{\beta}}\frac{\partial \underline{\beta}}{\partial \underline{\theta}} + \frac{\partial \underline{\alpha}}{\partial \underline{\beta}^*}\frac{\partial \underline{\beta}^*}{\partial \underline{\theta}}. \tag{1.131}$$

In the following, we will use the vector notation where straightforwardly applicable, but switch to complex notation to alleviate analytical complexity where necessary.

The point mass lens

Perhaps the simplest gravitational lens model is that of a point mass, for which the gravitational potential takes the form

$$\Phi(r) = -\frac{GM}{r}. \tag{1.132}$$

Using eq. (1.64), this gives a deflection angle of

$$\hat{\boldsymbol{\alpha}} = \frac{4GM}{b}\hat{\boldsymbol{b}}, \tag{1.133}$$

where $\boldsymbol{b}$ is the impact parameter described previously, b its magnitude, and $\hat{\boldsymbol{b}}$ is a unit vector in the direction of $\boldsymbol{b}$. This simple case illustrates some intuitive features of lensing by compact objects, with the deflection increasing with the mass of the object, and decreasing with the distance of closest approach.

Substituting the above result into the lens equation,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\frac{4GM}{b}\hat{\boldsymbol{b}}. \tag{1.134}$$

It is common practice to define the optical axis such that it intersects with the centre of the lens. In this case, $\boldsymbol{b}$ is also the position of the image, since it's where the light ray crosses the lens plane, and so we have $\boldsymbol{b} = \theta D_d$. The lens equation can then be rewritten as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}\frac{4GM}{\theta D_d}\hat{\boldsymbol{\theta}}, \tag{1.135}$$

where $\hat{\boldsymbol{\theta}}$ is a unit vector in the direction of $\boldsymbol{\theta}$. For a source located at the intersection between the optical axis and the source plane, the resulting image position is known as the Einstein radius θ_{E}. Setting $\boldsymbol{\beta} = 0$ in the lens equation, we see that, for the point mass,

$$\theta_{\mathrm{E}} = \sqrt{4GM\frac{D_{ds}}{D_d D_s}}. \tag{1.136}$$

The lens equation can therefore be expressed as

$$\boldsymbol{\theta} - \boldsymbol{\beta} - \frac{\theta_{\mathrm{E}}^2}{\theta}\hat{\boldsymbol{\theta}} = \mathbf{0}, \tag{1.137}$$

from which we see that each of the three terms must lie on a line, and we can express the image position corresponding to a given source position via the scalar equation

$$\theta = \frac{1}{2}\left[\beta \pm \sqrt{\beta^2 + 4\theta_{\mathrm{E}}^2}\right]. \tag{1.138}$$

A source at position $\boldsymbol{\beta}$ produces two images, one on the opposite side of the point mass and the other on the same side. The Einstein radius defines a characteristic scale for the deflection caused by the lens.

Singular and non-singular isothermal sphere lenses

While a useful starting point, the point source mass is poorly suited to model lensing by extended galaxies. A spherically-symmetric extended alternative often employed in simplified illustrated models is the singular isothermal sphere [99]. Here, we imagine a galaxy consisting of mass (such as stars) which behaves as in ideal gas, in thermal equilibrium (in other words, with an isothermal stellar gas and hence

velocity dispersion) and in hydrostatic equilibrium. A simple mass distribution which satisfies these conditions is the singular isothermal sphere,

$$\rho(r) = \frac{\sigma_v^2}{2\pi G}\frac{1}{r^2}. \tag{1.139}$$

Projecting this along the line of sight, the corresponding surface density is found to be

$$\Sigma(\boldsymbol{\xi}) = \frac{\sigma_v^2}{2G\xi}. \tag{1.140}$$

The deflection angle for this surface density is

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = 4\pi\sigma_v^2\hat{\boldsymbol{\xi}}, \tag{1.141}$$

and thus the lens equation is

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ds}}{D_s}4\pi\sigma_v^2\hat{\boldsymbol{\theta}}. \tag{1.142}$$

An interesting feature of the singular isothermal sphere model is that the deflection angle $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ does not depend at all on the distance from the centre of the lens, and depends on $\boldsymbol{\xi}$ only in that the deflection will be parallel to the position vector in the lens plane.

As for the point mass lens, the Einstein radius is the image position of a source located at $\boldsymbol{\beta} = 0$, and can easily be seen to be

$$\theta_{\mathrm{E}} = \frac{D_{ds}}{D_s}4\pi\sigma_v^2 = |\boldsymbol{\alpha}|. \tag{1.143}$$

Plugging this into the lens equation,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \theta_{\mathrm{E}}\hat{\boldsymbol{\theta}}. \tag{1.144}$$

As for the point mass lens, it is clear that $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ lie on the same line, and the lens equation can be written in scalar form as

$$\beta = \theta - \theta_{\mathrm{E}}\frac{\theta}{|\theta|}. \tag{1.145}$$

From this equation, the importance of the Einstein radius to lensing by a singular isothermal sphere is clear. Sources with $|\boldsymbol{\beta}| > \theta_{\mathrm{E}}$ will correspond to a single image. However, for a source with $|\boldsymbol{\beta}| < \theta_{\mathrm{E}}$, there are two solutions to eq. (1.145), and thus two images will form.

One of the more unphysical characteristics of the singular isothermal sphere model is that its surface mass density is singular at its centre (whence the name). The nonsingular isothermal sphere (NIS) model address this with a surface mass density given by,

$$\Sigma(\boldsymbol{\xi}) = \frac{\sigma_v^2}{2G}\frac{1}{\sqrt{\xi^2 + \xi_c^2}} = \frac{\Sigma_0}{\sqrt{1 + \xi^2/\xi_c^2}}, \tag{1.146}$$

in which $\Sigma_0 = v^2/(2G\xi_c)$ is the central surface mass density, and ξ_c is the size of the core. Expressed in terms of the convergence,

$$\kappa(\boldsymbol{x}) = \frac{1}{2\sqrt{\theta^2 + \theta_c^2}}, \tag{1.147}$$

with $\theta_c \equiv \xi/\xi_c$.

The singular isothermal ellipse

While SIS and NIS models provide better approximations to certain extended circularly-symmetric gravitational lenses than the point mass lens, almost all observed galaxy lenses feature a degree of ellipticity. The simplest lens model which is generally able to reproduce the main features of observed galaxy lenses is the singular isothermal ellipse (SIE), which is a generalisation of the SIS. [99, 160]

In what follows, we will consider a general ellipse with the major axis in the ξ_2 direction, and the minor axis in the ξ_1 direction. The axis ratio (minor axis over major axis) will be denoted f. The surface mass distributions for the SIS and NIS models, eqs. (1.140) and (1.146), can be generalised for elliptical mass distributions by replacing the impact vector ξ with

$$\zeta \equiv \sqrt{f^2\theta_1^2 + \theta_2^2}. \tag{1.148}$$

This is constant on ellipses with minor axis ζ, major axis ζ/f, and thus axis ratio f. The surface mass density of the singular isothermal ellipsoid (SIE) is defined to be

$$\Sigma(\boldsymbol{\xi}) = \frac{\sqrt{f}\sigma_v^2}{2GD_d}\frac{1}{\zeta}, \tag{1.149}$$

which follows from substituting ζ in place of ξ in the equation for the surface mass density of an SIS, and multiplying by a factor $\sqrt{f}$. The normalisation of this is chosen such that the mass inside an elliptical iso-density contour for fixed Σ is independent of f, which is shown in appendix A.1.3.

Introducing

$$\varphi = \arctan\left(\frac{\xi_2}{\xi_1}\right), \tag{1.150}$$

$$\Delta(\varphi) = \sqrt{f^2\cos^2\varphi + \sin^2\varphi}, \tag{1.151}$$

$$f' = \sqrt{1-f^2}, \tag{1.152}$$

the lensing potential can be written as

$$\psi(\theta,\varphi) = \frac{\sqrt{f}\theta}{f'}\left[|\sin\varphi|\operatorname{arccosh}\frac{\Delta}{f} + |\cos\varphi|\arccos\Delta\right], \tag{1.153}$$

the gradient of which gives the displacement angle,

$$\boldsymbol{\alpha} = \frac{\sqrt{f}}{f'}\left[\arcsin\left(f'\cos\varphi\right)\hat{\boldsymbol{\xi}}_1 + \operatorname{arcsinh}\left(\frac{f'}{f}\sin\varphi\right)\hat{\boldsymbol{\xi}}_2\right]. \tag{1.154}$$

in the limit $f \to 1$, this reduces to the deflection angular of the SIS.

The non-singular version is obtained by introducing the core radius ζ_c, with dimensionless value defined by $b_c = \zeta_c/D_d$, the surface mass density becomes

$$\Sigma(\zeta) = \frac{v^2}{2G}\frac{\sqrt{f}}{\sqrt{\zeta^2+\zeta_c^2}} = \frac{\Sigma_0}{\sqrt{1+\zeta^2/\zeta_c^2}}, \tag{1.155}$$

with $\Sigma_0 = \sqrt{f}v^2/(2G\zeta_c)$, i.e.

$$\kappa(b) = \frac{\sqrt{f}}{2\sqrt{b^2+b_c^2}}. \tag{1.156}$$

This gives a non-singular isothermal ellipsoid (NIE). It fulfills the same 'constant-mass' property as the singular model if the core radius is kept fixed.

Isothermal lenses have infinite total mass, which is of course highly unphysical. To address this, the mass density beyond some tidal radius is set to drop more rapidly than ζ^{-2}. If this is much larger than the typical length scale ξ_0 where strong lensing occurs, this cut-off has no effect on the strong lensing properties of the deflector.

Elliptical power laws

The most commonly used simple parameterised lens model is the elliptical power law (EPL) lens, which adds a further degree of flexibility to the SIE [153]. The surface mass density of an EPL is given by

$$\kappa(\zeta) = \frac{3-\gamma_{\rm PL}}{2}\left(\frac{\theta'_E}{\zeta}\right)^{\gamma_{\rm PL}-1}, \tag{1.157}$$

where $\gamma_{\rm PL}$ is the negative power-law slope of the 3D mass distribution, θ'_E is the circularised Einstein radius, and ζ as defined in eq. (1.148).

The complex deflection angle can be written as

$$\underline{\alpha}(\zeta,\varphi) = \frac{2\theta'_E}{1+f}\left(\frac{\theta'_E}{\zeta}\right)^{\gamma_{\rm PL}-2} e^{i\varphi_{\rm ell}}\,{}_2F_1\left(1;\frac{\gamma_{\rm PL}-1}{2};\frac{3-\gamma_{\rm PL}}{2};-\frac{1-f}{1+f}e^{i2\varphi_{\rm ell}}\right), \tag{1.158}$$

where ${}_2F_1(a;b;c;z)$ is the Gaussian hypergeometric function, and $\varphi_{\rm ell}$ is the elliptical angle, defined by

$$\varphi_{\rm ell} = \arctan\left(\frac{\theta_2}{f\theta_1}\right). \tag{1.159}$$

The lensing potential (in complex coordinates) is then

$$\psi(z) = \frac{1}{3-\gamma_{\rm PL}}\frac{z\alpha^*(z)+z^*\alpha(z)}{2}. \tag{1.160}$$

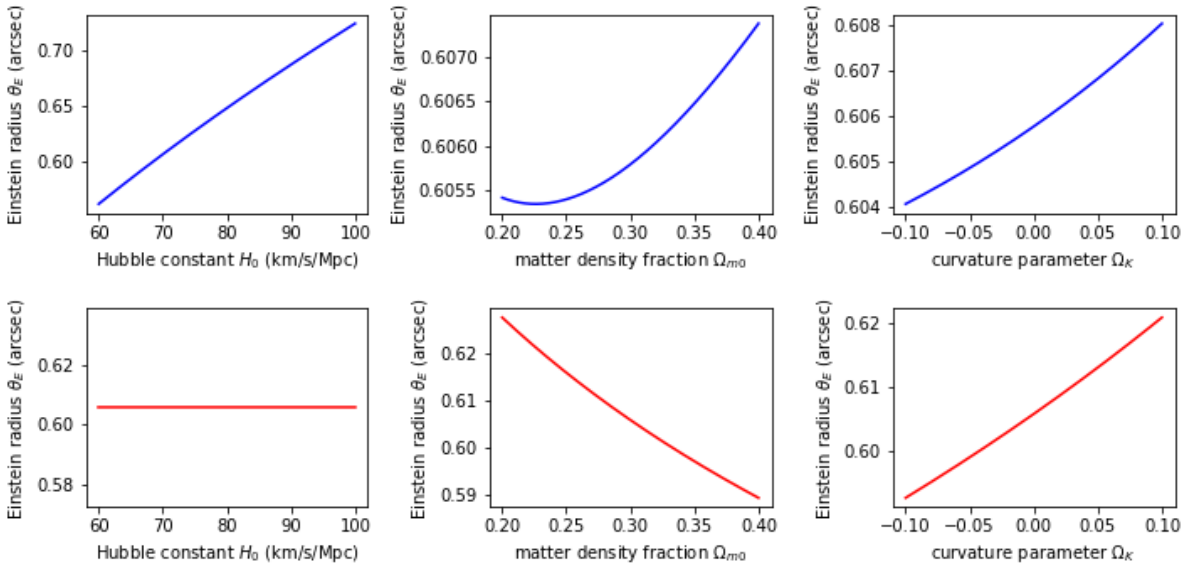

Figure 1.4: *The effects of cosmological parameters on Einstein radius for a point mass lens (top, in blue) and a singular isothermal sphere (bottom, in red). Both lenses are placed at redshift $z_d = 0.5$, with a source located at $z_s = 1.5$. The point lens has a mass of 10^{11} solar masses, and the velocity dispersion of the SIS is chosen such that its Einstein radius matches that of the point mass when $H_0 = 70$ km/s/Mpc, $\Omega_{m0} = 0.3$ and $\Omega_K = 0$.*

1.1.6 Light propagation in an FLRW Universe

Though deceptively simple, eq. (1.70) depends on various cosmological parameters in numerous subtle ways, even under the assumption of a simple homogeneous and isotropic FLRW background spacetime. Perhaps the most obvious of these is in the angular diameter distance factors which convert physical quantities such as $\boldsymbol{\xi}$ to angular quantities such as $\boldsymbol{\theta}$, and scale the deflection angle $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ to the displacement angle $\boldsymbol{\alpha}(\boldsymbol{\theta})$. These distances, in turn, relate to the redshifts and comoving distances to lens and source, the latter of which is found by integrating the inverse of the Hubble parameter, $H^{-1}(z)$. Thus the Hubble constant H_0, the curvature parameter K, the matter and vacuum energy density fractions Ω_m and Ω_Λ and the redshifts z_d and z_s can each impact lensing observables. However, these effects can be highly dependent on the lens model.

The most directly observable strong lensing parameter is the Einstein radius, which is plotted as a function of the Hubble constant H_0, the matter density fraction Ω_{m0} and the curvature parameter Ω_K for the point mass lens and the singular isothermal sphere in fig. 1.4. For the point mass, H_0 has by far the most significant impact of θ_E. However, for the SIS, its contribution is cancelled out (as can be seen in eq. (1.143)), and thus has no effect whatsoever on θ_E.

1.2 The Fermat potential and strong-lensing time delays

The lens equation presented in section 1.1 determines the position/s $\boldsymbol{\theta}$ at which a light ray travelling through a perturbed spacetime would be observed, given a source position $\boldsymbol{\beta}$ and a displacement angle $\boldsymbol{\alpha}(\boldsymbol{\theta})$. However, the distortions arising from the presence of a gravitational field affect not only the position of the ray, but also the length of the path through which it travelled, and the gravitational potential it experienced while doing so. When a beam is strongly lensed, the arrival time for the multiple trajectories of the beam will not necessarily be equal.

The time delay experienced by ray between emission and observation is governed by the Fermat potential, which serves both as a dimensionless measure of the arrival time and as a potential from which the lens equation can be rederived.

1.2.1 Fermat's principle

Fermat's Principle - Let S be an event (emission of light at the source) and ℓ a time-like world line (the "observer") in a spacetime $(M, g_{\alpha\beta})$. Then a smooth null curve γ from S to ℓ is a light ray (null geodesic) if, and only if, its arrival time τ on ℓ is stationary under first-order variations of γ within the set of smooth null curves from S to ℓ, i.e. $\delta\tau = 0$ [137].

A null curve is simply a path through spacetime for which any infinitesimal segment $\mathrm{d}s$ consists of equal spatial and temporal components, i.e. $\mathrm{d}\boldsymbol{x} = \mathrm{d}t$ in units where $c = 1$. Fermat's Principle states that, given the set of all possible null curves originating at the same point in space and time and terminating on the worldline of the observer, paths which are actually followed by a light ray, in other words null geodesics, must extremise the arrival time measured by that observer. Thus, by minimising an expression for the arrival time for a light ray passing through a particular spacetime, we have an alternative derivation of the lens equation.

1.2.2 The Fermat potential

The lens equation

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\psi(\boldsymbol{\theta}), \tag{1.161}$$

where as before, ψ is the lensing potential, can be rewritten as

$$\boldsymbol{\beta} = \nabla_{\boldsymbol{\theta}} \left[\frac{1}{2}\boldsymbol{\theta}^2 - \psi(\boldsymbol{\theta}) \right]. \tag{1.162}$$

By a simple rearrangement of this equation, $\boldsymbol{\beta}$ can be absorbed into the derivative, and so the lens equation in turn, can be expressed in terms of the scalar function

$$\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}), \tag{1.163}$$

as

$$\nabla_{\boldsymbol{\theta}}\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = 0. \tag{1.164}$$

ϕ is the Fermat potential, and relates to the light travel time of light rays. In fact, eq. (1.164) is the formulation of Fermat's principle applied to gravitational lensing, as shall be seen.

Fermat's principle offers a reformulation of the physics of gravitational lensing. In section 1.1.1, the lens equation was derived from the geodesic equation, a differential equation which relates a change in λ to a change in the tangent vector which traces out the path of the light ray. However, this analysis can also be done in terms of a least action principle (Fermat's principle). The path followed by a light ray as it travels from the source to the observer is extremised.

In the case of the geodesic equation, a highly complex, infinite-variable problem was greatly simplified by invoking the Born approximation, and treating all the deflection to the light ray as taking place in a single plane. The multi-plane formalism is a discrete version of the geodesic equation, and for an infinite number of infinitesimal planes, we arrive back at the geodesic equation.

The same simplification is invoked when starting from Fermat's principle. Normally, the trajectory must be extremised amongst an infinite number of possible paths. However, this is made much simpler by imagining that all the deflection takes place in a single plane. The optimisation problem reduces to one of a simple potential, which depends on just one variable.

1.2.3 The Shapiro time delay

Origins of the time delay

Two effects contribute to differences in the light travel times along different paths. Firstly, there is a geometrical time delay, due to the fact that the deflection causes the light to travel a curved (and hence longer) path. Secondly, as the light passes through the gravitational field of the deflector, there will be a gravitational time delay. For weak gravitational fields, with a metric given by eq. (1.31), this is found by setting $\mathrm{d}s^2 = 0$, and solving for $\mathrm{d}t$

$$(1 + 2\Phi)\mathrm{d}t^2 = (1 - 2\Phi)\mathrm{d}l^2, \tag{1.165}$$

$$\mathrm{d}t \approx (1 - 2\Phi)\mathrm{d}l. \tag{1.166}$$

Thus the integral of -2Φ along the ray gives the gravitational time delay. The distinction between the geometric and gravitational time delays is not fundamental, as discussed at the end of this section.

A useful consequence of Etherington's Reciprocity Theorem

Recall Etherington's principle

$$\det \boldsymbol{\mathcal{D}}(\lambda_1 \leftarrow \lambda_2) = \det \boldsymbol{\mathcal{D}}(\lambda_2 \leftarrow \lambda_1), \tag{1.167}$$

and the fact that

$$\omega_o^2 \det \boldsymbol{\mathcal{D}}(\lambda_s \leftarrow \lambda_o) \equiv D_{\mathrm{A}}^2. \tag{1.168}$$

From this, we can write

$$D_{\mathrm{A}}(\lambda_1 \leftarrow \lambda_2) = \frac{\omega_2}{\omega_1} D_{\mathrm{A}}(\lambda_2 \leftarrow \lambda_1). \tag{1.169}$$

From the definition of redshift,

$$\frac{\omega_i}{\omega_o} = 1 + z_i, \tag{1.170}$$

from which it follows that

$$\frac{\omega_2}{\omega_1} = \frac{1+z_2}{1+z_1}. \tag{1.171}$$

This allows us to rewrite eq. (1.169) as

$$\frac{D_{\mathrm{A}}(\lambda_1 \leftarrow \lambda_2)}{1+z_2} = \frac{D_{\mathrm{A}}(\lambda_2 \leftarrow \lambda_1)}{1+z_1}, \tag{1.172}$$

or equivalently,

$$\frac{D_{\mathrm{A}}(z_1, z_2)}{1+z_1} = \frac{D_{\mathrm{A}}(z_2, z_1)}{1+z_2}, \tag{1.173}$$

where $D_{\mathrm{A}}(z_1, z_2)$ is the angular diameter distance to redshift z_2 as seen from z_1.

Deriving the Shapiro Time Delay

Suppose an observer at the origin sees two images of a source, located at $\boldsymbol{\xi}^{(1)}$ and $\boldsymbol{\xi}^{(2)}$ in the image plane. Now, suppose the position of the source changes along a curve $\boldsymbol{\eta}(\lambda)$. Then, the image positions will also change, moving along some curve $\boldsymbol{\xi}^{(i)}(\lambda)$ for $i = 1, 2$. At every point along these curves, i.e. at every value of λ, the lens equation must be satisfied, provided $\boldsymbol{\eta}(\lambda)$ does not cross a caustic curve (in which case the number of images would change).

Consider the case where two source positions differ by $\mathrm{d}\boldsymbol{\eta}$, and wavefronts travelling from a pair of images from the original source position intersect the observer position simultaneously. The same wavefronts from the second source position will intersect the position of the observer at different times.

The time delay occurs on the path between the source positions and the observer positions. These paths are shown in fig. 1.5. The pair of wavefronts from $\boldsymbol{\xi}^{(1)}$ and $\boldsymbol{\xi}^{(2)}$ could be assumed to arrive simultaneously at the observer, and then the time difference when these wavefronts are instead arriving from the new image positions as a result of the source moving from $\boldsymbol{\eta}$ to $\boldsymbol{\eta}+\mathrm{d}\boldsymbol{\eta}$ could be calculated. However, this would be exactly equivalent to considering the same wavefronts departing simultaneously from $\boldsymbol{\eta}$, but departing with some time delay from $\boldsymbol{\eta}+\mathrm{d}\boldsymbol{\eta}$, related to the observed time delay via the redshift of the source. This is shown on the right of fig. 1.5. If the two images, located at $\boldsymbol{\xi}^{(1)}$ and $\boldsymbol{\xi}^{(2)}$, are separated by an angle $\boldsymbol{\sigma}$ as seen from the source plane, then so too are their wavefronts (as they are perpendicular to the light rays). The time delay at the source can be seen from the geometry of the situation to be

$$\mathrm{d}(\Delta t_s) = \boldsymbol{\sigma} \cdot \mathrm{d}\boldsymbol{\eta}. \tag{1.174}$$

The time delay is larger if the angular separation between the images is larger, or if the change in observer position is larger, or if the change in source position is closer to parallel to the image separation. There is no time delay if the source moves perpendicularly to the image separation. Because we assume $\mathrm{d}\boldsymbol{\eta}$ to be small, the change in the angular separation does not appear in the infinitesimal.

Now, note that we have been working with the time delay at the source, $\mathrm{d}t_s$. To obtain the time delay at the position of the observer, we must use the definition of redshift,

$$\frac{\mathrm{d}t_o}{\mathrm{d}t_s} = 1 + z_s, \tag{1.175}$$

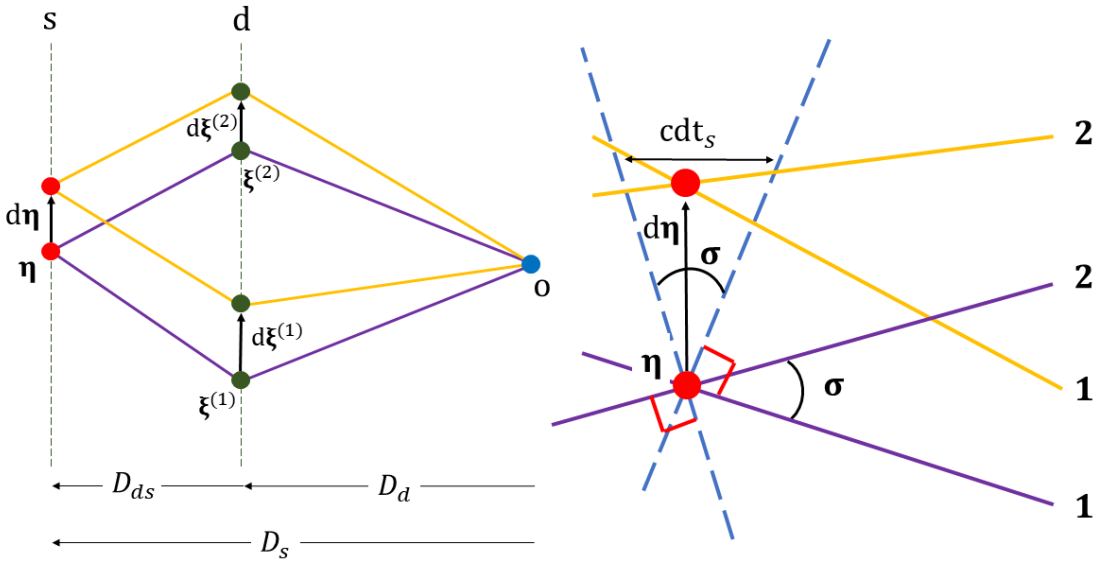

Figure 1.5: *The relevant geometry for calculating the time delay by considering a source moving from position* $\boldsymbol{\eta}$ *to* $\boldsymbol{\eta} + \mathrm{d}\boldsymbol{\eta}$*. As a result of this movement, the image positions change by* $\mathrm{d}\boldsymbol{\xi}$*, and wavefronts which would have reached the observer simultaneously no longer do so. This is equivalent to saying that two wavefronts depart simultaneously from* $\boldsymbol{\eta}$*, while the same wavefronts would depart non-simultaneously from* $\boldsymbol{\eta} + \mathrm{d}\boldsymbol{\eta}$*, separated by a time* $\mathrm{d}t_s$*. This is shown on the right of the figure.*

and so

$$\mathrm{d}(\Delta t) = (1 + z_s)\boldsymbol{\sigma} \cdot \mathrm{d}\boldsymbol{\eta}. \tag{1.176}$$

The total time delay if the source moves from $\boldsymbol{\eta}_0$ to $\boldsymbol{\eta}$ can be found by integrating the expression along the curve $\boldsymbol{\eta}(\lambda)$,

$$\Delta t(\boldsymbol{\eta}) = (1 + z_s)\left[\int_{\boldsymbol{\eta}_0}^{\boldsymbol{\eta}} \boldsymbol{\sigma}(\boldsymbol{\eta}') \cdot \mathrm{d}\boldsymbol{\eta}' + \Delta t(\boldsymbol{\eta}_0)\right]. \tag{1.177}$$

$\boldsymbol{\eta}_0$ is the point at which they are simultaneous, which is arbitrary. This equation tells us the time delay for any source position $\boldsymbol{\eta}$.

The dimensionless separation between the two images is related to their positions and the distance to the lens via

$$\boldsymbol{\sigma} = \frac{1}{D_{sd}}\left(\boldsymbol{\xi}^{(2)} - \boldsymbol{\xi}^{(1)}\right) = \frac{D_d}{D_{sd}}\left(\boldsymbol{\theta}^{(2)} - \boldsymbol{\theta}^{(1)}\right). \tag{1.178}$$

Note that the relevant distance here is D_{sd}, i.e. the angular diameter distance to the lens, as seen from the source. In order to express this in terms of D_{ds}, the angular distance to the source as seen by the lens, we make use of the relationship given in eq. (1.173), from which we have

$$D_{sd} = \frac{1 + z_s}{1 + z_d} D_{ds}, \tag{1.179}$$

and so

$$\boldsymbol{\sigma} = \frac{1 + z_d}{1 + z_s}\frac{D_d}{D_{ds}}\left(\boldsymbol{\theta}^{(2)} - \boldsymbol{\theta}^{(1)}\right). \tag{1.180}$$

The dimensionless source displacement $\mathrm{d}\boldsymbol{\beta}$ that corresponds to $\mathrm{d}\boldsymbol{\eta}$ is

$$\mathrm{d}\boldsymbol{\beta} = \frac{1}{D_s}\mathrm{d}\boldsymbol{\eta}. \tag{1.181}$$

Substituting eq. (1.180) and eq. (1.181) into eq. (1.177), the integral becomes

$$\Delta t_s(\boldsymbol{\beta}) = (1 + z_d)\left[\frac{D_s D_d}{D_{ds}}\int_{\boldsymbol{\beta}_0}^{\boldsymbol{\beta}} \mathrm{d}\boldsymbol{\beta}' \cdot \left[\boldsymbol{\theta}^{(2)}(\boldsymbol{\beta}') - \boldsymbol{\theta}^{(1)}(\boldsymbol{\beta}')\right] + \Delta t(\boldsymbol{\beta}_0)\right]. \tag{1.182}$$

To evaluate the integral, note $\boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\beta} = \mathrm{d}(\boldsymbol{\theta} \cdot \boldsymbol{\beta}) - \boldsymbol{\beta} \cdot \mathrm{d}\boldsymbol{\theta}$. From the lens equation in the form $\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\psi(\boldsymbol{\theta})$, we can rewrite $\boldsymbol{\beta} \cdot \mathrm{d}\boldsymbol{\theta} = \boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\theta} - \mathrm{d}\psi(\boldsymbol{\theta})$ to obtain

$$\boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\beta} = \mathrm{d}(\boldsymbol{\theta} \cdot \boldsymbol{\beta}) - \boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\theta} + \mathrm{d}\psi(\boldsymbol{\theta}). \tag{1.183}$$

From eq. (1.163),

$$\mathrm{d}\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = (\boldsymbol{\theta} - \boldsymbol{\beta})(\mathrm{d}\boldsymbol{\theta} - \mathrm{d}\boldsymbol{\beta}) - \mathrm{d}\psi(\boldsymbol{\theta}), \tag{1.184}$$

$$= \boldsymbol{\beta} \cdot \mathrm{d}\boldsymbol{\beta} - \mathrm{d}(\boldsymbol{\theta} \cdot \boldsymbol{\beta}) + \boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\theta} - \mathrm{d}\psi(\boldsymbol{\theta}), \tag{1.185}$$

and so we can write

$$\boldsymbol{\theta} \cdot \mathrm{d}\boldsymbol{\beta} = \mathrm{d}\left[\frac{\boldsymbol{\beta}^2}{2} - \phi(\boldsymbol{\theta}, y)\right]. \tag{1.186}$$

Since the $\boldsymbol{\theta}^{(2)}(\boldsymbol{\beta}')$ and $\boldsymbol{\theta}^{(1)}(\boldsymbol{\beta}')$ terms in the integral in eq. (1.182), the $\boldsymbol{\beta}^2/2$ terms cancel each other out when the relationship given above is used, leaving

$$\int_{\boldsymbol{\beta}_0}^{\boldsymbol{\beta}} \mathrm{d}\boldsymbol{\beta}' \cdot \left[\boldsymbol{\theta}^{(2)}(\boldsymbol{\beta}') - \boldsymbol{\theta}^{(1)}(\boldsymbol{\beta}')\right] = \left[\phi(\boldsymbol{\theta}^{(1)}, \boldsymbol{\beta}) - \phi(\boldsymbol{\theta}^{(2)}, \boldsymbol{\beta})\right] \tag{1.187}$$

$$- \left[\phi(\boldsymbol{\theta}_0^{(1)}, \boldsymbol{\beta}_0) - \phi(\boldsymbol{\theta}_0^{(2)}, \boldsymbol{\beta}_0)\right]. \tag{1.188}$$

$\boldsymbol{\theta}_0^{(i)}$, $i = 1, 2$ are the image positions for the source at $\boldsymbol{\beta}_0$.

Now, if $\boldsymbol{\beta}_0$ is on a caustic of the source's light cone, then the two images can fuse as the source moves to $\boldsymbol{\beta}_0$. Then $\boldsymbol{\theta}_0^{(1)} = \boldsymbol{\theta}_0^{(2)}$, $\Delta t(\boldsymbol{\beta}_0) = 0$, and we arrive at

$$\Delta t_s(\boldsymbol{\beta}) = (1 + z_d)\frac{D_d D_s}{D_{ds}}\left[\phi(\boldsymbol{\theta}^{(1)}, \boldsymbol{\beta}) - \phi(\boldsymbol{\theta}^{(2)}, \boldsymbol{\beta})\right]. \tag{1.189}$$

This expression is independent of the position $\boldsymbol{\beta}_0$ on the caustic curve. The prefactor is typically expressed as the so-called time delay distance τ_{ds},

$$\tau_{ds} = (1 + z_d)\frac{D_d D_s}{D_{ds}}, \tag{1.190}$$

which serves as a proportionality factor between the Fermat potentials and a physical time delay.

The time delay for a single ray relative to a reference ray travelling in the absence of a gravitational potential is simply found by replacing one of the potential terms in eq. (1.189) with the expression in eq. (1.163), and setting the other potential term to zero, which would be the potential in the absence of a gravitational field for a ray crossing the lens plane at $\boldsymbol{\theta} = \boldsymbol{\beta}$,

$$t(\boldsymbol{\theta}) = \tau_{ds}\left[\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta})\right]. \tag{1.191}$$

The quantity given in eq. (1.189) is observable, but the quantity in eq. (1.191) is not.

The resulting expression is sometimes interpreted as meaning that the total time delay is sum of the extra path length from the deflection and the gravitational time delay. While an appealing distinction, it is important to keep in mind that these contributions are not distinct, and the expressions were not arrived at with any such distinction coming into play. Fundamentally, it is the geometry of spacetime in the presence of a gravitational source, through which the light ray passes, which gives rise to this time delay.

From eq. (1.191), the Fermat potential (multiplied by the time delay scale, involving a redshift factor) can be interpreted as the time delay between the lensed signal, and an unlensed signal travelling through an FLRW background from the same source. Fermat's Principle tells us that a physical light ray will travel along the path which minimises this additional time of travel.

1.2.4 Degeneracies in single-plane lensing

As the formalism of gravitational lensing has been developed, several key degeneracies have been identified, arising from mathematical transformations which leave lensing observables unchanged while systematically altering inferred properties of the progenitor lens, source and cosmology [66, 133, 178, 105].

These degeneracies are often related to substructure in the lens galaxy which cannot be resolved by lensing observables alone, and to the fact that the true source position is fundamentally unknowable, and can only be determined by tracing light rays backwards in spacetime through the gravitational potential of the lens. As almost all lensing observables are dimensionless, hidden degeneracies abound when attempting to extract measurements of absolute quantities, with various rescalings being possible which remain consistent with observed features of a lens. Because the source position cannot be directly observed, any simultaneous transformations of the lensing potential and source position which leave lensing observables unaffected cannot be precluded.

One such transformation manifests as the monopole degeneracy [133]. In essence, if convergence is simultaneously added to and subtracted from any circular region of the lens plane, provided region does not contain an image and the change is circularly symmetric, then image positions, magnifications and time delays are unaffected. While this does not affect the inference of H_0, it greatly limits what can be inferred about the true mass distribution of a lens.

Perhaps the most important, notorious and well-studied degeneracy in gravitational lensing is the mass-sheet transformation (MST), first identified in [51]. Under this transformation, a constant sheet of mass with convergence $\lambda^{\rm ms}$ is added to the lens, while the lens convergence κ is simultaneously rescaled by $(1-\lambda^{\rm ms})$ (note this is typically just denoted as λ, but we include the superscript to distinguish from the affine parameter). This convergence sheet amplifies or dampens deflection angles, while the rescaling of the convergence changes the gradient of the lensing profile. While the affect on deflection angles cancels out, time delays are unaffected by the sheet of mass, but the quantity $H_0\Delta t$ is reduced by a shallower density profile (and increased by a steeper one). Without independent constraints on the mass profile of the lens, an underestimated $\lambda^{\rm ms}$ would therefore lead to H_0 being overestimated, and vice versa [178].

A range of interpretations of this effect are possible. As pointed out in [145, 94], lensing observables tend to depend on the ratio $\psi'''/(1-\psi'')$, which is unaffected by a mass-sheet transform. [156] notes that relative time delays provide constraints on $\Delta t \propto H_0^{-1}(1-\kappa_{\rm a})$, where $\kappa_{\rm a}$ is the convergence in the annulus between two images, and thus any uncertainty in $\kappa_{\rm a}$ translates to an uncertainty in H_0. In the language of [133], the MST and other lensing degeneracies can be interpreted as a rescaling of the arrival time surface. In the same paper, it is noted that, while an infinite uniform sheet of convergence is certainly an unphysical proposition, the monopole transformation and the MST can manifest simultaneously in the "mass-disk degeneracy". Any redistribution of mass inwards of the observed lensing images and any change in mass beyond these images are both unobservable, provided they are circularly symmetric. From the perspective of lensing observables, there is no difference between an infinite sheet and the much more physically realisable circular disk, which cannot be eliminated by boundary conditions on the convergence.

The mass-sheet degeneracy, the gauge invariance of the lensing potential, and the effects of a constant source position transform can be illustrated together as follows. Consider a lensing potential $\psi(\theta)$, from which the modified potential $\psi_\lambda(\theta)$ is defined as

$$\psi_\lambda(\boldsymbol{\theta}) = \frac{\lambda^{\rm ms}}{2}\boldsymbol{\theta}^2 + \boldsymbol{s}\cdot\boldsymbol{\theta} + c + (1-\lambda^{\rm ms})\psi(\boldsymbol{\theta}), \tag{1.192}$$

where $\boldsymbol{s}$ is a constant 2D vector and c and $\lambda^{\rm ms}$ are constants. c defines the zero point of the potential, since lensing observables depend only on derivatives of ψ, it can be freely set to 0.

The displacement angle which corresponds to this new lensing potential is given by

$$\boldsymbol{\alpha}_\lambda(\boldsymbol{\theta}) = \nabla\psi_\lambda = \lambda^{\rm ms}\boldsymbol{\theta} + \boldsymbol{s} + (1-\lambda^{\rm ms})\boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.193}$$

and the convergence by

$$\kappa_\lambda(\boldsymbol{\theta}) = \frac{1}{2}\nabla^2\psi_\lambda = \lambda^{\rm ms} + (1-\lambda^{\rm ms})\kappa(\boldsymbol{\theta}). \tag{1.194}$$

Under this transformation, the lens equation becomes

$$\boldsymbol{\beta}_\lambda = \boldsymbol{\theta} - \boldsymbol{\alpha}_\lambda(\boldsymbol{\theta}) = \boldsymbol{\theta} - \lambda^{\rm ms}\boldsymbol{\theta} - \boldsymbol{s} - (1-\lambda^{\rm ms})\boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{1.195}$$

The amplification matrix corresponding to ψ_λ is then

$$\boldsymbol{\mathcal{A}}_\lambda = \frac{{\rm d}\boldsymbol{\beta}_\lambda}{{\rm d}\boldsymbol{\theta}} = (1-\lambda^{\rm ms})\mathbf{1} - (1-\lambda^{\rm ms})\boldsymbol{\Gamma}(\boldsymbol{\theta}) = (1-\lambda^{\rm ms})\boldsymbol{\mathcal{A}}, \tag{1.196}$$

and therefore the magnification, shears and convergences take the form

$$\mu_\lambda = \frac{\mu}{(1-\lambda^{\rm ms})^2}, \tag{1.197}$$

$$\gamma_\lambda(\boldsymbol{\theta}) = (1-\lambda^{\rm ms})\gamma(\boldsymbol{\theta}), \tag{1.198}$$

$$(1-\kappa_\lambda) = (1-\lambda^{\rm ms})(1-\kappa). \tag{1.199}$$

The reduced shear is invariant under the transformation,

$$g_\lambda = \frac{\gamma}{1-\kappa} = g. \tag{1.200}$$

Now, dividing eq. (1.195) throughout by the $(1-\lambda^{\rm ms})$ term,

$$\frac{\boldsymbol{\beta}_\lambda}{1-\lambda^{\rm ms}} + \frac{\boldsymbol{s}}{1-\lambda^{\rm ms}} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \boldsymbol{\beta}. \tag{1.201}$$

Equation (1.201) illustrates the degeneracy between $\psi(\boldsymbol{\theta})$ and $\psi_\lambda(\boldsymbol{\theta})$. From a lensed image located at $\boldsymbol{\theta}$, there is no direct imaging information available to distinguish between the lensing potentials ψ and ψ_λ, and hence the source positions $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_\lambda = (1-\lambda^{\rm ms})\boldsymbol{\beta} + \boldsymbol{s}$ respectively. However, the Fermat potential is affected. Substituting eqs. (1.192) and (1.201) into eq. (1.163),

$$\phi_\lambda(\boldsymbol{\theta}, \boldsymbol{\beta}_\lambda) = \frac{1}{2}\left(\boldsymbol{\theta} - \boldsymbol{\beta}_\lambda\right)^2 - \psi_\lambda(\boldsymbol{\theta}), \tag{1.202}$$

$$\phi_\lambda(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1}{2}\left(\boldsymbol{\theta} - \boldsymbol{\beta} + \boldsymbol{s} + \lambda^{\rm ms}\boldsymbol{\beta}\right)^2 - \frac{\lambda^{\rm ms}}{2}\boldsymbol{\theta}^2 - \boldsymbol{s}\cdot\boldsymbol{\theta} - (1-\lambda^{\rm ms})\psi(\boldsymbol{\theta}) - c, \tag{1.203}$$

$$= \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 + (\boldsymbol{\theta} - \boldsymbol{\beta})(\boldsymbol{s} + \lambda^{\rm ms}\boldsymbol{\beta}) + \frac{1}{2}(\boldsymbol{s} + \lambda^{\rm ms}\boldsymbol{\beta})^2 - \frac{\lambda^{\rm ms}}{2}\boldsymbol{\theta}^2 - \boldsymbol{s}\cdot\boldsymbol{\theta} - (1-\lambda^{\rm ms})\psi(\boldsymbol{\theta}) - c, \tag{1.204}$$

Only terms which depend on $\boldsymbol{\theta}$ will affect relative time delays, and we can therefore write

$$\phi_\lambda(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \frac{\lambda^{\rm ms}}{2}\boldsymbol{\theta}^2 + \lambda^{\rm ms}\boldsymbol{\theta}\cdot\boldsymbol{\beta} - \frac{\lambda^{\rm ms}}{2}\boldsymbol{\beta}^2 - (1-\lambda^{\rm ms})\psi(\boldsymbol{\theta}) + \text{const.}, \tag{1.205}$$

$$= (1-\lambda^{\rm ms})\left[\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta})\right] + \text{const.}, \tag{1.206}$$

$$= (1-\lambda^{\rm ms})\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) + \text{const.} \tag{1.207}$$

Relative time delays between image positions $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$ will therefore take the form

$$\Delta t = (1-\lambda^{\rm ms})\tau_{ds}\left[\phi(\boldsymbol{\theta}_1, \boldsymbol{\beta}) - \phi(\boldsymbol{\theta}_2, \boldsymbol{\beta})\right]. \tag{1.208}$$

Because the $(1-\lambda^{\rm ms})$ term is independent of the specific image $\boldsymbol{\theta}_i$, a measured relative time delay of Δt is proportional to both $(1-\lambda^{\rm ms})$ and τ_{ds}, and the effects of these terms are indistinguishable.

Expanding this prefactor in terms of the redshifts and angular diameter distances, we have that

$$(1-\lambda^{\rm ms})\tau_{ds} = (1-\lambda^{\rm ms})(1+z_d)\frac{D_{od}D_{os}}{D_{ds}}, \tag{1.209}$$

These angular diameter distances are calculated according to

$$D_{ij} = \frac{f_K(\chi_j - \chi_i)}{1+z_j}, \tag{1.210}$$

where χ_i and χ_j are the comoving distances of i and j from the observer, and

$$f_K(\chi) \equiv \frac{\sin(\sqrt{K}\chi)}{\sqrt{K}}, \tag{1.211}$$

with K being the FLRW spatial-curvature parameter, and

$$\chi(z) = \int_0^z \frac{{\rm d}z'}{H(z')}. \tag{1.212}$$

$H(z)$ is the Hubble parameter at redshift z, which is given in terms of this redshift, the Hubble constant H_0, the density parameters for dark energy Ω_Λ, matter Ω_m, and curvature Ω_K as

$$H(z) = H_0\sqrt{\Omega_\Lambda + \Omega_m(1+z)^3 + \Omega_K(1+z)^2}. \tag{1.213}$$

Note that, if i were not comoving, equation 1.210 would need to be corrected by the ratio $\omega_{i,o}/\omega_{i,b}$ between the observed and background frequencies.

Now, since $|K| \ll 1$, it follows from the above that each of the angular diameter distances is approximately proportional to H_0^{-1}, and so

$$H_0 \propto \frac{1-\lambda^{\rm ms}}{\Delta t}. \tag{1.214}$$

Equation 1.214 illustrates in explicit terms the problem posed by the mass sheet degeneracy for measurements of H_0 from relative time delays in gravitational lensing. Any given set of strong lens imaging observables is consistent with the addition of a uniform sheet of convergence $\lambda^{\rm ms}$ and a corresponding rescaling of the lens' convergence by $(1-\lambda^{\rm ms})$, which scales the inferred value of H_0 by the same factor of $(1-\lambda^{\rm ms})$. Without further information to constrain the mass distribution of the main lens, precise measurements of H_0 are impossible. This problem is further compounded by the effect of matter along the line-of-sight, and is discussed further in section 1.3.4.

1.3 Multi-plane lensing and the dominant-lens approximation

The idealised setting of single-plane lensing serves as a crucial stepping stone towards a full understanding of the physics behind the increasingly large catalogue of strong gravitational lensing images. However, real lenses do not exist in isolation, but are instead immersed in the messy and complex mass distribution of the local Universe. With the resolution made possible by modern telescopes, neighbouring galaxies, matter along the line-of-sight and large-scale structure can have a significant impact on strong lensing systems, and must be taken into account when modeling these systems [12, 88, 78, 174, 83, 109, 23, 136].

In the language of [58], the spacetime through which a light beam passes can be thought of as the superposition of *smooth* and *rough* gravitational fields. As a beam passes through the smooth background spacetime, it experiences the continuous distorting effects of this field, but is never split. Embedded in this smooth background are rough-field regions (typically large galaxies), which are concentrated enough to result in strong lensing, splitting the beam into multiple rays. The tendency of matter in the Universe to clump into galaxies and galaxy clusters, near which the majority of the deflection of a lensed light beam occurs, means that much of the formalism of single-plane lensing introduced in section 1.1 can be preserved and adapted for the more general theoretical framework of multi-plane lensing [25]. Here, the thin-lens approximation is applied to each of an arbitrary number of rough-field lenses embedded in an arbitrary background spacetime. As the light beam propagates through these lenses, it experiences deflection in each plane, and the position at which the effects of each successive lens much be evaluated depends recursively on the ones proceeding it.

While essential for modelling systems in which two or three galaxies are clearly identifiable in good alignment with the source, the multi-plane lensing formalism is almost impossible to implement when describing subtler perturbations to strong lensing images in its full recursive form. In most typical lensing systems, the effects of one lens dominates over all the others, and the non-dominant lenses will only weakly lens the signal. This assumption is known as the Dominant-Lens approximation, and allows for a much more easily applied version of multi-plane lensing [59].

In this section, we present the mathematics of multi-plane lensing in the notation of [59], in which the single-plane lensing formalism is generalised to an arbitrary number of thin lens planes located at different redshifts. Invoking the Dominant-Lens approximation, we then summarise the work by Fleury, Larena and Uzan in [59, 58], who introduced a new formalism to describe the effects of matter along the line-of-sight on strong lensing images. We will pay particular attention to the case where all but one of the lenses are treated under the tidal approximation.

1.3.1 Multi-plane lensing

In multi-plane lensing, a light ray is assumed to propagate from source s to observer o through an FLRW spacetime, in which an arbitrary number of thin lens planes lie orthogonal to the optical axis. In the limit as the number of lens approaches infinity, the exact mass distribution through which the ray passes is reconstructed. As for the single-plane lens, we will assume that the light is only deflected in these lens planes.

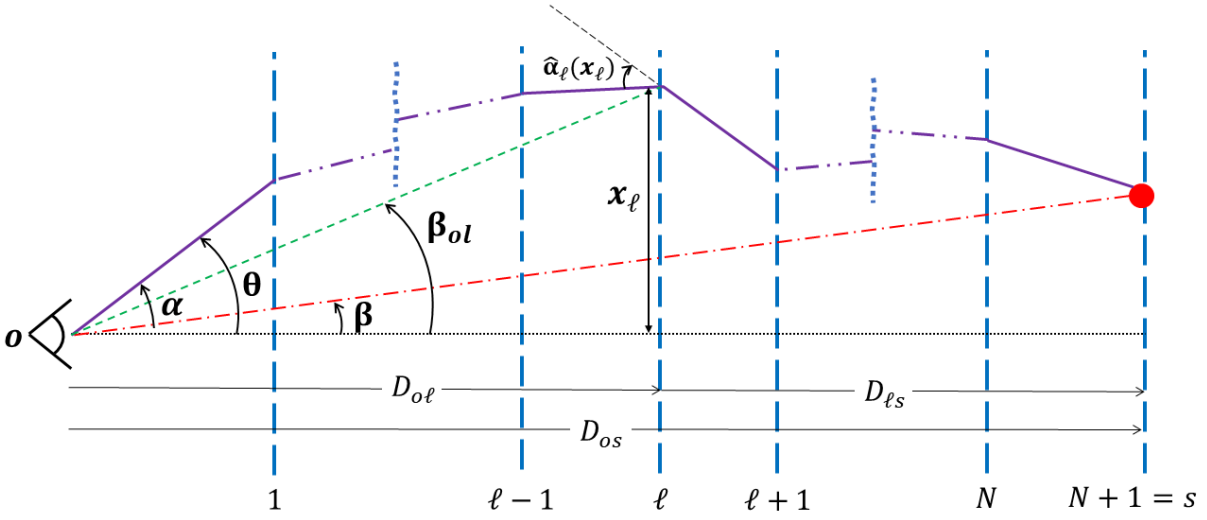

Figure 1.6: *Multi-plane lensing in an FLRW background. The dotted black line is the optical axis, which defines the origin of each of the N lens planes. The light ray leaves the source at an angular position $\boldsymbol{\beta} = \boldsymbol{\beta}_{os}$ (which is where the source would be observed in the absence of any lenses), and is deflected by an angle $\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l)$ at a physical position $\boldsymbol{x}_l$ at each lens l for $l = 1, ..., N$. $\boldsymbol{x}_l$ corresponds to an angular position $\boldsymbol{\beta}_{ol} = \boldsymbol{x}_l/D_{ol}$, where D_{ol} is the angular-diameter distance to l as seen by the observer, in the absence of lenses (note, this has nothing to do with the specific path of a light ray, but rather must be specified because the presence of lenses lead to a magnification effect, which in turn affects the angular diameter distance). The ray crosses the lens plane closest to the observer at an angular position $\boldsymbol{\beta}_{o1} = \boldsymbol{\theta} = \boldsymbol{\beta} + \boldsymbol{\alpha}$. The dash-dotted lines passing through the wavy dotted blue lines between plane 1 and $l-1$, and between $l+1$ and N, represent the arbitrary number of lens planes and hence deflections which may be present between these planes.*

Figure 1.6 illustrates the path of a light ray from s to o through N thin lens planes. As in the case of single-plane lensing, we will use $\boldsymbol{\theta}$ to refer to the observed angular position of an image in the lens plane, $\boldsymbol{\beta}$ to the angular position at which it would be seen in the absence of these lens planes, and the displacement angle $\boldsymbol{\alpha}$ to the difference between these quantities. The subscript l, with $1 \leq l \leq N$, will refer to an arbitrary lens within the set of lens planes. $\boldsymbol{x}_l$ denotes the physical position at which the light beam crosses lens plane l, and $\hat{\boldsymbol{\alpha}}(\boldsymbol{x}_l)$ the deflection experienced by a light beam crossing plane l at $\boldsymbol{x}_l$, in the rest frame of the lens. The angular diameter distances from observer to lens D_{ol}, from observer to source D_{os} and from lens to source D_{ls} are defined in the absence of any lens planes.

The lensing potential

In the case of single-plane lensing, we arrived at an expression for the deflection angle by considering perturbations to the metric. Invoking the thin-screen approximation, we assumed that all the deflection took place in a single plane. From the geometry of lensing, we related this deflection angle $\hat{\boldsymbol{\alpha}}$ to the angular position of the image $\boldsymbol{\theta}$ and of the source $\boldsymbol{\beta}$, and defined the displacement angle $\boldsymbol{\alpha}$ as the difference between these.

In the case of multi-plane lensing, the overall displacement of the beam depends on its deflection at each of the lens planes. Because of the coupling between these multiple deflections, the displacement angle can no longer be written in terms of a simple potential. Instead, using $\Sigma_l(\boldsymbol{x})$ to refer to the surface density of plane l at position $\boldsymbol{x}$, we begin by defining the lensing potential of a lens plane as

$$\hat{\psi}_l(\boldsymbol{x}) \equiv \int \mathrm{d}^2\boldsymbol{y} 4G\Sigma_l(\boldsymbol{y}) \ln|\boldsymbol{x} - \boldsymbol{y}|. \tag{1.215}$$

By this definition, $\hat{\psi}$ satisfies the Poisson equation for the projected surface density,

$$\nabla^2_{\boldsymbol{x}} \hat{\psi}_l = 8\pi G \Sigma_l, \tag{1.216}$$

as it is twice the two-dimensional projected gravitational potential of the mass Σ_l.

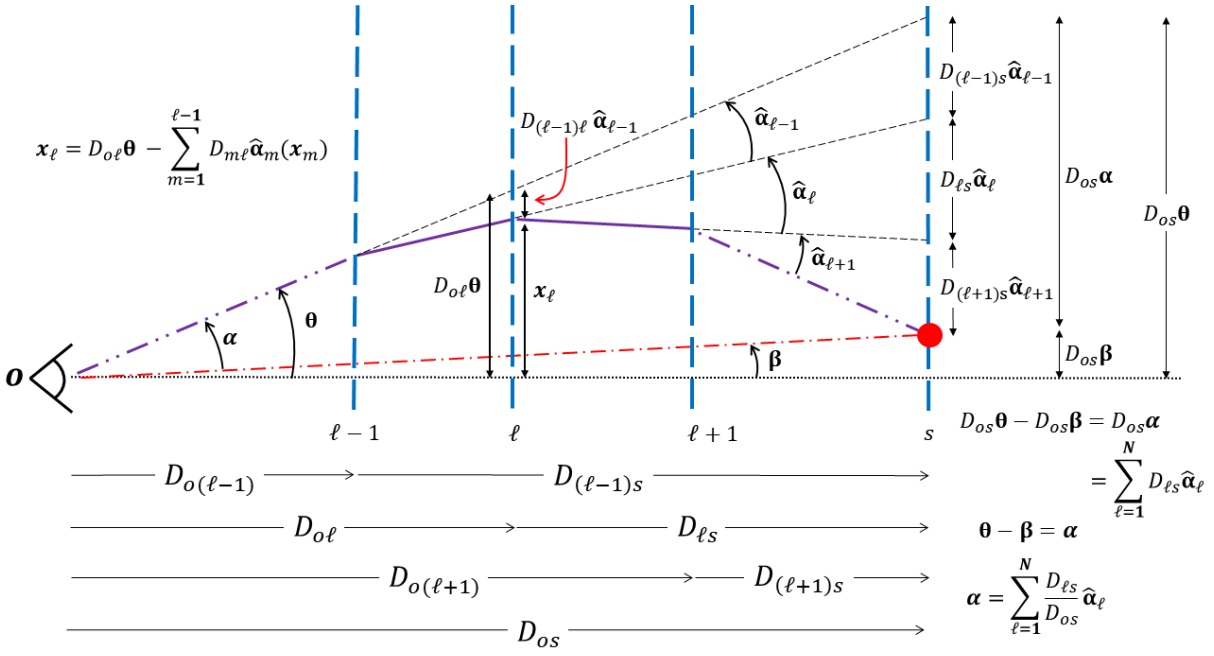

Figure 1.7: *The geometry of the multi-plane lens equation. The image and source positions $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are separated by an angle $\boldsymbol{\alpha}$, which can be seen to be the sum from $l = 1$ to $l = N$ of terms which take the form $(D_{ls}/D_{os})\hat{\boldsymbol{\alpha}}_l$, as can be seen in the diagram. Also illustrated in the diagram is the calculation of $\boldsymbol{x}_l$ from the deflection angles of all the lenses $m \leq l-1$, the displacement angle $\boldsymbol{\theta}$, and the appropriate angular distances. Three lens planes illustrated, but the recursive relationships are trivially generalised to an arbitrary number of planes.*

In contrast with the lensing potential $\psi(\boldsymbol{\theta})$ in 1.87, which relates to the overall displacement angle $\boldsymbol{\alpha}$ of the single plane lens, $\hat{\psi}(\boldsymbol{\theta})$ relates to the deflection angle $\hat{\boldsymbol{\alpha}}_l$ via

$$\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l) \equiv \frac{\mathrm{d}\hat{\psi}_l}{\mathrm{d}\boldsymbol{x}_l} = \int \mathrm{d}^2\boldsymbol{x} 4G\Sigma_l(\boldsymbol{x})\frac{\boldsymbol{x}_l - \boldsymbol{x}}{|\boldsymbol{x}_l - \boldsymbol{x}|^2}. \tag{1.217}$$

As discussed in section 1.1.6, the background spatially-averaged FLRW spacetime contributes to the lensing of light. The lens planes described above are a means by which perturbations from this spacetime are modelled, and thus may have positive or negative densities [58]. These negative density lenses would correspond to under-dense regions of the true spacetime. While the physical distribution of matter is continuous, matter in the Universe tends to clump together, and is thus reasonably approximated by a series of discrete lens planes.

A negative value of the projected surface mass density Σ would produce a negative deflection angle (eq. (1.217)), which corresponds to a reduced difference between $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. This would take the form of a deflection away from the optical axis, because such a deflection would have the overall result of a light path which was closer to the unlensed path.

The multi-plane lens equation

As light propagates through multiple lens planes, it experiences a deflection $\hat{\boldsymbol{\alpha}}_l$ at each plane l. Since this deflection angle is evaluated at the position $\boldsymbol{x}_l$ at which it reaches the plane, which in turn depends on the deflections by each the proceeding planes, the resulting displacement of the light beam must be generated recursively. From fig. 1.7, it can be seen that the multi-plane lens equation takes the form

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}, \qquad \boldsymbol{\alpha} = \sum_{l=1}^{N} \frac{D_{ls}}{D_{os}}\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l). \tag{1.218}$$

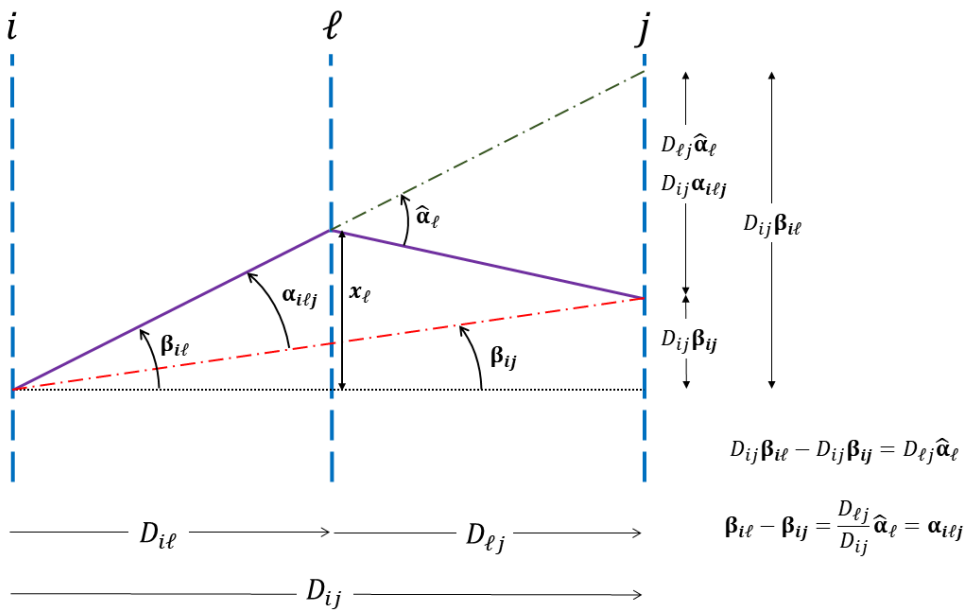

Figure 1.8: *The geometrical interpretation of $\boldsymbol{\alpha}_{ilj}(\boldsymbol{x}_l)$, which related the angular position $\boldsymbol{\beta}_{il}$ of x_l as seen from i in the absence of lenses, to the angular position $\boldsymbol{\beta}_{ij}$ of x_j as seen from j in the absence of lenses. Note these are not necessarily the same x_l and x_js which would appear anywhere in the multi-plane lensing equation, as this assumes that there are no lenses between i and j except for l. Setting $i = o$ and $j = s$ tells us how light would be deflected if l was the only lens between the observer and the source in an otherwise FLRW spacetime.*

Given the dependence of $\boldsymbol{x}_l$ on the previous lens planes, it must also be determined recursively as in fig. 1.7,

$$\boldsymbol{x}_l = D_{ol}\boldsymbol{\theta} - \sum_{m=1}^{l-1} D_{ml}\hat{\boldsymbol{\alpha}}_m(\boldsymbol{x}_m) \qquad \forall l \in \{1, ..., N+1\}. \tag{1.219}$$

Here, we generalise our notation for the angular diameter distances, such that D_{ij} now refers to the angular diameter distance from plane i to plane j, calculated according to eq. (1.210).

Partial displacement

In what follows, it will be useful to define the partial displacement

$$\boldsymbol{\alpha}_{ilj}(\boldsymbol{x}_l) \equiv \frac{D_{lj}}{D_{ij}}\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l), \tag{1.220}$$

such that each term in the sum in eq. (1.218) is expressed as $\boldsymbol{\alpha}_{ols}$. The geometric interpretation of this is shown in fig. 1.8. This notation is applicable to any three planes with $i < l < j$.

We also introduce

$$\boldsymbol{\beta}_{ij} \equiv \frac{\boldsymbol{x}_j}{D_{ij}}, \tag{1.221}$$

which is the angular position at which the point $\boldsymbol{x}_j$ in plane j would be seen from plane i in the absence of any lenses, and under the assumption that plane i has no peculiar velocity (i.e. is comoving with the Hubble flow). To account for a nonzero peculiar velocity in any such calculation which affects the frequency of light, we must include a correction factor of the ratio of the observed frequency to the background frequency of the light at i, $\omega_i/\bar{\omega}_i$.

The partial distortion matrix

The gradient of $\boldsymbol{\alpha}_{ilj}$ with respect to $\boldsymbol{\beta}_{il}$ is denoted

$$\boldsymbol{\Gamma}_{ilj} \equiv \frac{\mathrm{d}\boldsymbol{\alpha}_{ilj}}{\mathrm{d}\boldsymbol{\beta}_{il}}. \tag{1.222}$$

From eqs. (1.217), (1.220) and (1.221), this can be expressed as

$$\boldsymbol{\Gamma}_{ilj} = \frac{D_{il}D_{lj}}{D_{ij}}\frac{\mathrm{d}^2\hat{\psi}_l}{\mathrm{d}\boldsymbol{x}_l^2} = \frac{\boldsymbol{\Sigma}_l}{\Sigma_{ilj}^{\mathrm{crit}}}, \tag{1.223}$$

with

$$\boldsymbol{\Sigma}_l \equiv \frac{\mathrm{d}^2}{\mathrm{d}\boldsymbol{x}_l^2}\int \frac{\mathrm{d}^2\boldsymbol{y}}{\pi}\Sigma_l(\boldsymbol{y})\ln|\boldsymbol{x}_l - \boldsymbol{y}|, \qquad \Sigma_{ilj}^{\mathrm{crit}} \equiv \left(\frac{4\pi G D_{il}D_{lj}}{D_{ij}}\right)^{-1}. \tag{1.224}$$

The amplification matrix $\boldsymbol{\mathcal{A}}_{ilj} = 1 - \boldsymbol{\Gamma}_{ilj}$ is a partial distortion matrix, describing the effect of lens l on an infinitesimal source in plane j, as observed from plane i.

While the displacement angle $\boldsymbol{\alpha}$ in eq. (1.218) could be obtained from a summation of the partial displacement angles $\boldsymbol{\alpha}_{ilj}$, the relationship between the amplification matrix $\boldsymbol{\mathcal{A}}$ (characterising the overall distortions from source to observer) and the partial distortion matrices $\boldsymbol{\mathcal{A}}_{ilj}$ is less straightforward.

As in the single lens case, the shear matrix $\boldsymbol{\Gamma}_{ilj}$, being a matrix of derivatives, is symmetric, and can therefore be decomposed as

$$\boldsymbol{\Gamma}_{ilj} \equiv \kappa_{ilj}\mathbf{1} + \begin{bmatrix} \mathrm{Re}(\gamma_{ilj}) & \mathrm{Im}(\gamma_{ilj}) \\ \mathrm{Im}(\gamma_{ilj}) & -\mathrm{Re}(\gamma_{ilj}) \end{bmatrix}, \tag{1.225}$$

from which the partial convergence κ_{ilj} and partial complex shear γ_{ilj} are defined.

Now, $\boldsymbol{\Sigma}_l$ describes the distribution of matter in the plane l. From equation eq. (1.223), it is equal to $\boldsymbol{\Gamma}_{ilj}$ multiplied by a constant. It is therefore a quadrupole matrix, i.e. it is invariant under rotations by integer multiples of π, with a trace given by

$$\mathrm{tr}\left[\boldsymbol{\Sigma}_l\right] = \mathrm{tr}\left[\frac{\mathrm{d}^2}{\mathrm{d}\boldsymbol{x}_l^2}\int \frac{\mathrm{d}^2\boldsymbol{y}}{\pi}\Sigma_l(\boldsymbol{y})\ln|\boldsymbol{x}_l - \boldsymbol{y}|\right]. \tag{1.226}$$

$\mathrm{d}^2/\mathrm{d}\boldsymbol{x}_l^2$ is the matrix of derivatives, the trace of which is the Laplacian, giving

$$\mathrm{tr}\left[\boldsymbol{\Sigma}_l\right] = \Delta\int \frac{\mathrm{d}^2\boldsymbol{y}}{\pi}\Sigma_l(\boldsymbol{y})\ln|\boldsymbol{x}_l - \boldsymbol{y}|, \tag{1.227}$$

$$= \frac{1}{4\pi G}\Delta\hat{\psi}_l, \tag{1.228}$$

$$= 2\Sigma_l. \tag{1.229}$$

It can therefore be similarly decomposed to yield

$$\boldsymbol{\Sigma}_l \equiv \Sigma_l\mathbf{1} + \begin{bmatrix} \mathrm{Re}(Q_l) & \mathrm{Im}(Q_l) \\ \mathrm{Im}(Q_l) & -\mathrm{Re}(Q_l) \end{bmatrix}, \tag{1.230}$$

where Σ_l is the homogeneous matter distribution, and the complex number Q_l is the quadrupole of the projected density.

The proportionality factor $\Sigma_{ilj}^{\mathrm{crit}}$ arises from the dependence of $\boldsymbol{\Gamma}_{ilj}$ on the positions of j and i, from which the ray originates and is observed, respectively. Given the particular combination of planes i, l and j, it characterises the density scale at which these planes could result in strong lensing (for which the combination is referred to as "critical"). From the expansions of $\boldsymbol{\Gamma}_{ilj}$ and $\boldsymbol{\Sigma}_l$, it is easy to see that

$$\kappa_{ilj} = \frac{\Sigma_l}{\Sigma_{ilj}^{\mathrm{crit}}}, \qquad \gamma_{ilj} = \frac{Q_l}{\Sigma_{ilj}^{\mathrm{crit}}}. \tag{1.231}$$

1.3.2 The dominant lens approximation

The multi-plane lensing formalism allows, in theory, for the effects of an arbitrary number of lens planes to be included when modelling a strongly lensed image. In a handful of cases, lensing systems have been observed for which multiple lens profiles at different redshifts are necessary in models of the system [35, 61]. However, such alignments are rare, and in most cases there is a single lens which is responsible for the majority of observed lensing effects [136]. Nonetheless, the effects of the other lens planes cannot be omitted when attempting to reconstruct lensed images.

To exploit this fact, the dominant-lens (DL) approximation was introduced in [59], an assumption under which the overall displacement angle $\boldsymbol{\alpha}$ reduces to a function of $\boldsymbol{\theta}$ only, eliminating the need for recursion.

Defining the dominant-lens approximation

The dominant-lens approximation is the assumption that only one of the N lens planes d is super-critical ($\kappa_{ods} > 1$), and thus could result in strong lensing effects. All other lens planes $l \neq d$ are treated as perturbations, and are assumed to be highly sub-critical in the relevant area of the sky, with

$$\forall l \neq d \;\; \forall i < l < j \quad \kappa_{ilj} \ll 1. \tag{1.232}$$

We introduce the parameter ϵ to quantify these perturbations, such that, for a plane $l \neq d$ with a localised mass distribution, $\kappa_{ilj} = \mathcal{O}(\epsilon^2)$. ϵ can be though of as being proportional to the typical Einstein radius of the non-dominant lens. To understand this, a localised mass distribution is one which is concentrated enough that there is some region of sky in which $k_{ilj} > 1$ for that plane. For such a distribution, its mass M will increase in proportion to the average value of κ_{ilj} in that plane. Equivalently, the average mass M of a lens plane will increase in proportion to the average value of k_{ilj} across the non-dominant lenses of an entire lens system. Since the Einstein radius of a lens varies with $\sqrt{M}$, it follows that the typical Einstein radius of the lens planes will be proportional to ϵ.

Since $\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l)$ involves an integral over Σ_l (eq. (1.217)) and so this follows that $\boldsymbol{\alpha}_{ilj}$ will also be on the order of κ_{ilj} (with some additional prefactors). The same must be true for the derivative of $\boldsymbol{\alpha}_{ilj}$ (the symmetric part of which is κ_{ilj} by definition). The DL approximation will therefore allow us to expand the complicated recursive relationships of the fully general multi-plane formalism, and to omit all but zeroth and first order terms in ϵ^2.

The lens equation in the DL regime

The multi-plane lens equation in eq. (1.218) can be split into the displacement from the foreground ($l < d$), dominant ($l = d$) and background ($l > d$) lenses,

$$\boldsymbol{\theta} - \boldsymbol{\beta} = \boldsymbol{\alpha} = \underbrace{\sum_{l<d} \boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol})}_{\text{foreground}} + \underbrace{\boldsymbol{\alpha}_{ods}(\boldsymbol{\beta}_{od})}_{\text{dominant}} + \underbrace{\sum_{l>d} \boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol})}_{\text{background}}. \tag{1.233}$$

The angles $\boldsymbol{\beta}_{ol} = \boldsymbol{x}_l / D_{ol}$, in terms of which the deflection angles have been expressed, are related to each other via the recursion relation which follows from eq. (1.219),

$$\boldsymbol{\beta}_{ol} = \boldsymbol{\theta} - \sum_{m<l} \boldsymbol{\alpha}_{oml}(\boldsymbol{\beta}_{om}). \tag{1.234}$$

As can be seen from eqs. (1.233) and (1.234), the arguments of the partial deflection angles is whence the complexity of multi-plane lensing arises. However, by considering the foreground, background and dominant contributions in turn, expanding these expressions and neglecting terms of order higher than ϵ^2, we can reduce this to a single equation of the single observable $\boldsymbol{\theta}$. The details of this derivation are given in appendix A.2.1. Eventually, we arrive at the lens equation under the DL approximation,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.235}$$

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \underbrace{\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta})}_{\text{dominant}} - \underbrace{\boldsymbol{\Gamma}_{ods}(\boldsymbol{\theta}) \sum_{l<d} \boldsymbol{\alpha}_{old}(\boldsymbol{\theta})}_{\text{post Born for dominant}} + \underbrace{\sum_{l<d} \boldsymbol{\alpha}_{ols}(\boldsymbol{\theta})}_{\text{foreground}} + \underbrace{\sum_{l>d} \boldsymbol{\alpha}_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]}_{\text{background}} + \mathcal{O}(\epsilon^4). \tag{1.236}$$

The consequences of this result are shown in fig. 1.9, which captures the essence of the DL approximation. The contributions of the foreground perturbers are evaluated not at the true position through which the physical ray passes, but rather at $\boldsymbol{\theta}$ (the angular position observed from o), projected back through the planes. The contribution from the dominant lens, however, must be evaluated at a position which accounts for these foreground positions to first order. Finally, the contribution from the background lenses must be evaluated at positions modified by the deflection of the light ray in the dominant lens. However, to first order in ϵ^2, the effect of the main lens in the argument of the background lenses need only be evaluated at $\boldsymbol{\theta}$.

The precise way in which this manifests can be understood term by term in eq. (1.236). The first term is the deflection by the dominant lens, evaluated at a position which is unaffected by the foreground lenses (i.e. conforming to the Born approximation). This position is the intersection of the green dotted line with the dominant lens plane in fig. 1.9.

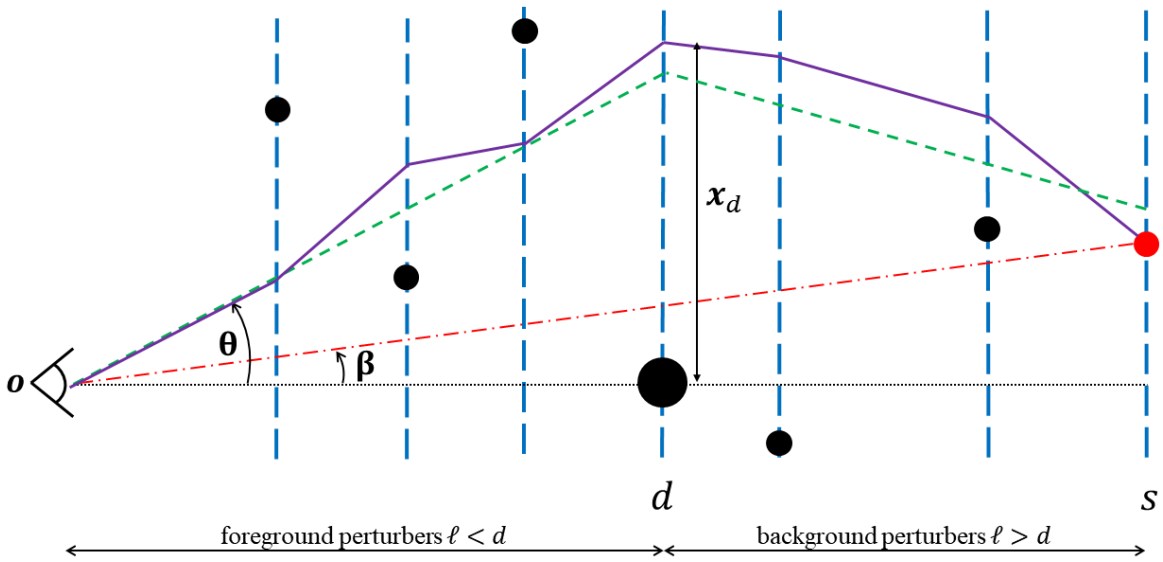

Figure 1.9: *Multi-plane lensing with a dominant lens. The lens at d dominates the deflection of the light ray, while the effects of the other lenses ($l \neq d$) are perturbative. The purple line is the actual path followed by the light ray, and the dot-dashed red line is the path it would take in the absence of any lens. The green dotted line is the path that would be followed by a ray if d was the only lens present. The dotted black line is the optical axis, which is defined so as to pass through the observer position and the centre of the dominant lens d. In the DL approximation, the effects of the perturbing lens planes are evaluated along the green dashed line (the dominant-lens-only ray), while the deflection caused by the dominant lens is corrected by a term which accounts for the foreground perturbers. As illustrated, deflections away from the optical axis are also possible, which can arise either from a lens plane with positive surface density, centred further from the optical axis than $\boldsymbol{x}_l$, or from a lens plane with negative surface density (not shown). The latter only holds if the relevant angular diameter distances are calculated under the assumption of a non-vacuum background spacetime (typically FLRW).*

The second term is a correction to this, which accounts for the fact that the light ray intersects the dominant lens plane at a position which is slightly offset from $\boldsymbol{\theta}$, because of the contributions of the foreground lenses contained in $\boldsymbol{\alpha}_{old}$. The terms within the sum account for the difference between $\boldsymbol{\theta}$ and $\boldsymbol{\beta}_{od}$, and the shear matrix multiplies this to map this difference in angular position in the lens plane to a difference in the deflection angle, to first order in ϵ^2.

The third term is the sum of the deflections by the foreground lenses in the Born approximation, i.e. evaluated along the path of the observed light ray as if there were no foreground lenses.

The fourth term is the contribution from the background lenses, which must be evaluated along a light ray emitted from the source before being deflected by the dominant lens. In other words, for a given background lens l, the observed angular position $\boldsymbol{\theta}$ must be adjusted by the deflection of the dominant lens, evaluated at $\boldsymbol{\theta}$ and as viewed from the background lens l. The deflection at this background lens $\boldsymbol{\alpha}_{ols}$ is then evaluated at this adjusted point (the green dotted line in fig. 1.9). Note that the post Born correction to the dominant lens does not feature in the argument of the partial deflection angles in the background, to first order in ϵ^2.

Put together, these terms allow an incoming light ray observed at $\boldsymbol{\theta}$ to be traced back to its point of origin, to first order in ϵ^2, given functions for the deflection angles $\hat{\boldsymbol{\alpha}}_l$ at each plane l, and the redshifts of each plane (from which the angular diameter distances needed to convert these deflection angles into the displacement angles appearing in eq. (1.236) can be determined, under the assumption of a certain cosmology).

The shear matrix

The partial shear matrices $\boldsymbol{\Gamma}_{ilj} = \mathrm{d}\boldsymbol{\alpha}_{ilj}/\mathrm{d}\boldsymbol{\beta}_{il}$ defined in section 1.3.1 characterise the distortions caused by a lens l to an infinitesimal source placed at j as observed from i in the absence of any other lenses. The total shear matrix $\boldsymbol{\Gamma}(\boldsymbol{\theta})$, which governs the overall distortions to an infinitesimal image at s as seen by o, is defined in terms of total displacement angle $\boldsymbol{\alpha}(\boldsymbol{\theta})$ as

$$\boldsymbol{\Gamma}(\boldsymbol{\theta}) \equiv \frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}\boldsymbol{\theta}}. \tag{1.237}$$

Taking the derivative of the total displacement angle in eq. (1.236), we arrive at an expression for the shear matrix as a function of $\boldsymbol{\theta}$ only,

$$\boldsymbol{\Gamma}(\boldsymbol{\theta}) = [\mathbf{1} - \boldsymbol{\Gamma}_{ds}(\boldsymbol{\theta})]\boldsymbol{\Gamma}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})][\mathbf{1} - \boldsymbol{\Gamma}_{od}(\boldsymbol{\theta})] + \boldsymbol{\Gamma}_{os}(\boldsymbol{\theta}), \tag{1.238}$$

where

$$\boldsymbol{\alpha}_{od}(\boldsymbol{\theta}) \equiv \sum_{l<d} \boldsymbol{\alpha}_{old}(\boldsymbol{\theta}) \tag{1.239}$$

is the displacement of a source located in plane d by the foreground perturbers,

$$\boldsymbol{\Gamma}_{od}(\boldsymbol{\theta}) \equiv \sum_{l<d} \boldsymbol{\Gamma}_{old}(\boldsymbol{\theta}). \tag{1.240}$$

is the distortion to an infinitesimal source in plane d by the foreground perturbers,

$$\boldsymbol{\Gamma}_{os}(\boldsymbol{\theta}) \equiv \sum_{l<d} \boldsymbol{\Gamma}_{ols}(\boldsymbol{\theta}) + \sum_{l>d} \boldsymbol{\Gamma}_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]. \tag{1.241}$$

are the distortions to an infinitesimal source in plane s by all perturbers $l \neq d$ as would be observed from o, with a post-Born correction applied to the contributions of the background perturbers, and

$$\boldsymbol{\Gamma}_{ds}(\boldsymbol{\theta}) \equiv \sum_{l>d} \boldsymbol{\Gamma}_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})], \tag{1.242}$$

are the distortions to an infinitesimal source in plane s as viewed from d, evaluated along the post-Born corrected path. The details of the derivation of this expression are given in appendix A.2.2.

In terms of the amplification matrix, this can be written as

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{A}}_{os}(\boldsymbol{\theta}) - \boldsymbol{\mathcal{A}}_{ds}(\boldsymbol{\theta})\left[\mathbf{1} - \boldsymbol{\mathcal{A}}_{ods}(\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta})\right]\boldsymbol{\mathcal{A}}_{od}(\boldsymbol{\theta}), \tag{1.243}$$

where we have written

$$\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta}) = \boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}, \tag{1.244}$$

and where these matrices are defined analogously to the shear matrices, with

$$\boldsymbol{\mathcal{A}}_{ij} = \mathbf{1} - \boldsymbol{\Gamma}_{ij}. \tag{1.245}$$

Equation 1.243 is simply the amplification matrix of a strong lens in the presence of purely tidal line-of-sight perturbations, formally equivalent in the result in 1.238. However, this result is nonetheless more general, because the perturbations $\boldsymbol{\alpha}_{od}$, $\boldsymbol{\Gamma}_{od}$, $\boldsymbol{\Gamma}_{os}$ and $\boldsymbol{\Gamma}_{ds}$ are allowed to vary across the image plane.

Convergence, shear and rotation

Unlike the partial shear matrices, the complete shear matrix $\boldsymbol{\Gamma}$ is not generally symmetric, as the product of two symmetric matrices (as in equation 1.238) is not generally symmetric. It will then be decomposed as

$$\boldsymbol{\Gamma} = \begin{bmatrix} \kappa + \mathrm{Re}(\gamma) & \mathrm{Im}(\gamma) - \omega \\ \mathrm{Im}(\gamma) + \omega & \kappa - \mathrm{Re}(\gamma) \end{bmatrix}. \tag{1.246}$$

As before, κ encodes the trace, $\mathrm{Re}(\gamma)$ the trace-free diagonal elements, $\mathrm{Im}(\gamma)$ the symmetric off-diagonal elements, but there is now an additional term ω, which relates to the rotation of infinitesimal images with respect to their sources. Given this decomposition, and working in the complex formalism, we have that

$$\frac{\partial \underline{\alpha}}{\partial \underline{\theta}} = \kappa + \mathrm{i}\omega, \tag{1.247}$$

$$\frac{\partial \underline{\alpha}}{\partial \underline{\theta}^*} = \gamma. \tag{1.248}$$

It is then possible to show that

$$\kappa(\boldsymbol{\theta}) + \mathrm{i}\omega(\boldsymbol{\theta}) = (1 - \kappa_{od} - \kappa_{ds})\kappa_{ods}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}) - (\gamma^*_{od}\gamma_{ods} + \gamma_{ds}\gamma^*_{ods}) + \kappa_{os}, \tag{1.249}$$

$$\gamma(\boldsymbol{\theta}) = (1 - \kappa_{od} - \kappa_{ds})\gamma_{ods}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}) - (\gamma_{od} + \gamma_{ds})\kappa_{ods} + \gamma_{os}, \tag{1.250}$$

A full derivation of eqs. (1.248) and (1.250) is given in appendix A.2.3. Once again, note that the convergences and shears present in the above equations are generally not constant across the image plane. If no argument is specified, evaluation at $\boldsymbol{\theta}$ is implied.

Magnification

The image-plane magnification $\mu(\boldsymbol{\theta})$ is the ratio of the angular sizes of infinitesimal images in the image plane $\boldsymbol{\theta}^2$ to their angular size in the source plane $\boldsymbol{\beta}^2$ due to lensing, or in other words, $\mu(\boldsymbol{\theta}) \equiv \mathrm{d}^2\boldsymbol{\theta}/\mathrm{d}^2\boldsymbol{\beta} = 1/\det \boldsymbol{\mathcal{A}}$. Taking the determinant of $\boldsymbol{\mathcal{A}}$, its inverse can be expressed as

$$\mu^{-1}(\boldsymbol{\theta}) = |1 - \kappa - \mathrm{i}\omega|^2 - |\gamma|^2. \tag{1.251}$$

In the DL approximation, this can be written as

$$\mu^{-1}(\boldsymbol{\theta}) = [1 - 2(\kappa_{od} + \kappa_{ds})]\mu^{-1}_{ods}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}) + 2(1 - \kappa_{ods})(\kappa_{od} + \kappa_{ds} - \kappa_{os}) + 2\mathrm{Re}[\gamma^*_{ods}(\gamma_{od} + \gamma_{ds} - \gamma_{os})], \tag{1.252}$$

where $\mu^{-1}_{ods} \equiv (1 - \kappa_{ods})^2 - |\gamma_{ods}|^2$ is the inverse magnification of the dominant lens only.

1.3.3 Line-of-sight effects in the tidal regime

The DL approximation eliminates the need for recursive calculations, and quantities of interest (such as the lens equation, eq. (1.236), the amplification matrix, eq. (1.243), or the magnification, eq. (1.252)) are functions of the image position $\boldsymbol{\theta}$ only. However, the summation terms in eq. (1.236) nonetheless require the functional forms of each of the $\boldsymbol{\alpha}_{old}$, $\boldsymbol{\alpha}_{odl}$ and $\boldsymbol{\alpha}_{ols}$ terms to be known for $l \neq d$, which adds a significant degree of complication to the single plane lens equation in any practical context.

This complexity can be greatly alleviated by assuming that all the non-dominant lenses ($l \neq d$) are *tidal*, or in other words, that their effects on a light beam are constant across the region of interest at the scale of the beam. In order for this assumption to hold, the non-dominant deflectors must either themselves be approximately smooth and constant in density over this region (producing an external convergence), or consist of clumps of matter far from the beam (producing an external shear) [136, 58, 59, 88]. In this subsection, we once again follow [59].

Defining the tidal regime

A tidal lens is one for which the spacetime curvature it produces is effectively constant in the region through which the light beam passes, thus removing the dependency of the matter quadrupole matrix $\boldsymbol{\Sigma}_l$ and the partial shear matrix $\boldsymbol{\theta}_{ilj} = \boldsymbol{\Sigma}_l / \Sigma^{\rm crit}_{ilj}$ on the position $\boldsymbol{x}_l$ of the beam in the lth plane. From eqs. (1.223) and (1.224), we have

$$4\pi G \boldsymbol{\Sigma}_l = \frac{\mathrm{d}\hat{\boldsymbol{\alpha}}_l}{\mathrm{d}\boldsymbol{x}_l}, \tag{1.253}$$

and so the tidal approximation equates to the statement that the deflection angle $\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l)$ is linear in $\boldsymbol{x}_l$. From this, it follows that for any couple of positions $\boldsymbol{x}_l$ and $\boldsymbol{x}'_l$,

$$\hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}_l) = \hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}'_l) + 4\pi G \boldsymbol{\Sigma}_l \cdot (\boldsymbol{x}_l - \boldsymbol{x}'_l), \tag{1.254}$$

hence any associated partial displacement angle (eq. (1.220)) can be found by multiplying the above expression throughout by D_{lj}/D_{ij}, and by referring to eq. (1.224),

$$\boldsymbol{\alpha}_{ilj}(\boldsymbol{\beta}_{il}) = \boldsymbol{\alpha}_{ilj}(\boldsymbol{\beta}'_{il}) + \frac{4\pi G D_{lj}}{D_{ij}} \boldsymbol{\Sigma}_l \cdot (D_{il}\boldsymbol{\beta}_{il} - D_{il}\boldsymbol{\beta}'_{il}), \tag{1.255}$$

$$= \boldsymbol{\alpha}_{ilj}(\boldsymbol{\beta}'_{il}) + \boldsymbol{\Gamma}_{ilj} \cdot (\boldsymbol{\beta}_{il} - \boldsymbol{\beta}'_{il}). \tag{1.256}$$

The projected potential $\hat{\psi}_l(\boldsymbol{x}_l)$ can be Taylor expanded at the reference point $\boldsymbol{x}'_l$. The first derivative is the deflection angle, and the second is given by

$$4\pi G \boldsymbol{\Sigma}_l = \frac{\mathrm{d}^2 \hat{\psi}_l}{\mathrm{d}\boldsymbol{x}_l^2}, \tag{1.257}$$

from which it is easily seen that higher-order derivatives vanish, because $\boldsymbol{\Sigma}_l$ is constant. We therefore arrive at

$$\hat{\psi}_l(\boldsymbol{x}_l) = \hat{\psi}_l(\boldsymbol{x}'_l) + \hat{\boldsymbol{\alpha}}_l(\boldsymbol{x}'_l) \cdot (\boldsymbol{x}_l - \boldsymbol{x}'_l) + \frac{1}{2}(\boldsymbol{x}_l - \boldsymbol{x}'_l) \cdot 4\pi G \boldsymbol{\Sigma}_l \cdot (\boldsymbol{x}_l - \boldsymbol{x}'_l). \tag{1.258}$$

The lens equation

In the tidal regime, the displacement angle reduces to

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = (\mathbf{1} - \boldsymbol{\Gamma}_{ds}) \cdot \boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0})] + \boldsymbol{\alpha}_{os}(\mathbf{0}) + \boldsymbol{\Gamma}_{os}\boldsymbol{\theta}, \tag{1.259}$$

where $\boldsymbol{\alpha}_{os}(0)$ is defined as the displacement of an image due to all the lenses except the dominant one,

$$\boldsymbol{\alpha}_{os}(\mathbf{0}) \equiv \sum_{l \neq d} \boldsymbol{\alpha}_{ols}(\mathbf{0}). \tag{1.260}$$

The details of this derivation are given in appendix A.3.1.

The shear matrix

The shear matrix $\boldsymbol{\Gamma}(\boldsymbol{\theta})$ in the tidal regime is found by taking the derivative of equation 1.259 wrt $\boldsymbol{\theta}$. Recall that all the shear matrices associated with the non-dominant lenses are independent of $\boldsymbol{\theta}$, and so the partial derivatives of these terms (and terms evaluated at $\mathbf{0}$) are zero.

$$\boldsymbol{\Gamma}(\boldsymbol{\theta}) = (\mathbf{1} - \boldsymbol{\Gamma}_{ds})\boldsymbol{\Gamma}_{ods}[\boldsymbol{\theta} - \boldsymbol{\Gamma}_{od}\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0})](\mathbf{1} - \boldsymbol{\Gamma}_{od}) + \boldsymbol{\Gamma}_{os}. \tag{1.261}$$

The result is almost identical to eq. (1.238), except for the replacement of the argument of $\boldsymbol{\Gamma}_{ods}$.

1.3.4 Lensing degeneracies and the observability of line-of-sight effects

Homogeneous displacements

Further simplifications to the lens equation and shear matrix are possible upon a recognition of certain key degeneracies in eqs. (1.259) and (1.261). The $\boldsymbol{\alpha}_{os}(\mathbf{0})$ term, in other words, the displacement caused by all but the dominant lens to a source located on the optical axis, shifts the apparent positions of all images, without having any effect on their apparent shape (eq. (1.261)) or luminosity (eq. (1.252)). The

unlensed position of the source $\boldsymbol{\beta}$ cannot be known except by tracing back the light ray using the lens equation, and so the global shift $\boldsymbol{\alpha}_{os}(\mathbf{0})$ is fully degenerate with $\boldsymbol{\beta}$. It can be eliminated by redefining $\boldsymbol{\beta} \to \boldsymbol{\beta} - \boldsymbol{\alpha}_{os}(\mathbf{0})$.

Similarly, the $\boldsymbol{\alpha}_{od}(\mathbf{0})$ term which appears in eqs. (1.259) and (1.261) describes a global shift to images within the dominant plane by the foreground lenses, including light from the main lens itself, without affecting shapes or luminosities. It is thus similarly unobservable and fully degenerate with the true position of the main lens, and can be eliminated by redefining the origin of the dominant lens such that positions within plane d are given by $\boldsymbol{x}_d \mapsto \boldsymbol{x}'_d \equiv \boldsymbol{x}_d - D_{od}\boldsymbol{\alpha}_{od}(\mathbf{0})$.

Following these redefinitions, the lens equation becomes

$$\boldsymbol{\beta} = \boldsymbol{\mathcal{A}}_{os}\boldsymbol{\theta} - \boldsymbol{\mathcal{A}}_{ds}\boldsymbol{\alpha}_{ods}(\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}), \tag{1.262}$$

and the shear matrix is

$$\boldsymbol{\Gamma}(\boldsymbol{\theta}) = (\mathbf{1} - \boldsymbol{\Gamma}_{ds})\boldsymbol{\Gamma}_{ods}[\boldsymbol{\theta} - \boldsymbol{\Gamma}_{od}\boldsymbol{\theta}](\mathbf{1} - \boldsymbol{\Gamma}_{od}) + \boldsymbol{\Gamma}_{os}. \tag{1.263}$$

Source-position transforms and the minimal lens model

From eqs. (1.262) and (1.263), it is possible to identify a further degeneracy between the shear matrices. As discussed above and in section 1.2.4, the source position $\boldsymbol{\beta}'$ is not directly observable, and thus any equation which relates the observed image positions $\boldsymbol{\theta}$ to some transformation of $\boldsymbol{\beta}'$ is no less applicable to lensing observables. Multiplying the lens equation 1.262 throughout by $\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}_{ds}^{-1}$ gives

$$\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\beta} = \boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os}\boldsymbol{\theta} - \boldsymbol{\mathcal{A}}_{od}\boldsymbol{\alpha}_{ods}(\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}), \tag{1.264}$$

or equivalently,

$$\tilde{\boldsymbol{\beta}} = \boldsymbol{\mathcal{A}}_{\mathrm{LOS}}\boldsymbol{\theta} - \frac{\mathrm{d}\psi_{\mathrm{eff}}}{\mathrm{d}\boldsymbol{\theta}}, \tag{1.265}$$

where

$$\tilde{\boldsymbol{\beta}} = \boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\beta} \approx (\mathbf{1} - \boldsymbol{\Gamma}_{od} + \boldsymbol{\Gamma}_{ds})\boldsymbol{\beta}, \tag{1.266}$$

$$\psi_{\mathrm{eff}}(\boldsymbol{\theta}) \equiv \psi_{ods}[\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}], \tag{1.267}$$

$$\boldsymbol{\mathcal{A}}_{\mathrm{LOS}} \equiv \boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os} \approx 1 - \boldsymbol{\Gamma}_{od} - \boldsymbol{\Gamma}_{os} + \boldsymbol{\Gamma}_{ds}. \tag{1.268}$$

In [59], eq. (1.265) is referred to as the "minimal lens model", in that it fully encodes the relevant degeneracies of lensing with tidal perturbations along the line of sight. We see that the amplification matrices $\boldsymbol{\mathcal{A}}_{od}$, $\boldsymbol{\mathcal{A}}_{os}$ and $\boldsymbol{\mathcal{A}}_{ds}$ all contribute to $\boldsymbol{\mathcal{A}}_{\mathrm{LOS}}$, and thus they are not measurable independently of one another. This degeneracy is further illustrated in mock data in [77]. While the argument of the lensing potential ψ_{ods} needs to be evaluated at a position which is modified by the foreground perturbers, which is therefore influenced by $\boldsymbol{\mathcal{A}}_{od}$ in a way which is not degenerate with $\boldsymbol{\mathcal{A}}_{os}$ and $\boldsymbol{\mathcal{A}}_{ds}$, this is highly model dependent, and a more general potential ψ_{eff} would be able to accommodate these modifications. Thus $\boldsymbol{\mathcal{A}}_{od}$ is degenerate with the main lens, as is explored below.

The foreground shear-ellipticity degeneracy

Because the foreground amplification matrix $\boldsymbol{\mathcal{A}}_{od}$ enters the lens equation within the argument of the main lens term, its effects are tied up in the effects of the main lens itself. In particular, the effects of foreground shear are, at least to linear order, degenerate with the effects of the main lens ellipticity. This point is made in [59], in which the degeneracy is demonstrated explicitly for a main lens with an SIE profile. Working at linear order in e, the lensing potential of such a lens is given by

$$\psi_{\mathrm{SIE}}(\boldsymbol{\theta}) = \theta_{\mathrm{E}}\theta\left[1 + \frac{1}{3}\mathrm{Re}\left(e\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right]. \tag{1.269}$$

When line-of-sight effects are included, the argument of the lensing potential becomes

$$(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta} = \underline{\theta} - \gamma_{od}\underline{\theta}^*, \tag{1.270}$$

$$= \theta\left[\mathrm{e}^{\mathrm{i}\varphi} - \gamma_{od}\mathrm{e}^{-\mathrm{i}\varphi}\right], \tag{1.271}$$

$$= \theta\left[(1 - \gamma_{od})\cos(\varphi) + (1 + \gamma_{od})\mathrm{i}\sin(\varphi)\right]. \tag{1.272}$$

The magnitude of this is, to first order in γ_{od},

$$|(\mathbf{1}-\mathbf{\Gamma}_{od})\boldsymbol{\theta}| = \theta\left(1-\mathrm{Re}(\gamma_{od}\mathrm{e}^{-2\mathrm{i}\varphi})\right), \tag{1.273}$$

the polar angle φ becomes

$$\varphi' = \varphi + 2\gamma_{od}\sin(\varphi)\cos(\varphi), \tag{1.274}$$

and therefore,

$$\mathrm{e}^{-2\mathrm{i}\varphi'} = \mathrm{e}^{-2\mathrm{i}\varphi}\left[1-4\mathrm{i}\gamma_{od}\sin(\varphi)\cos(\varphi)\right]. \tag{1.275}$$

Substituting these in place of $\boldsymbol{\theta}$ in 1.269, we immediately see that the modifications to φ are $\mathcal{O}(e\gamma_{od})$, and thus can be ignored at first order. The resulting potential takes the form

$$\psi_{\mathrm{SIE}}[(1-\mathbf{\Gamma}_{od})\boldsymbol{\theta}] = \theta_{\mathrm{E}}\theta\left\{1+\frac{1}{3}\mathrm{Re}\left[(e-3\gamma_{od})\mathrm{e}^{-2\mathrm{i}\varphi}\right]\right\}. \tag{1.276}$$

Therefore, to first order, there is no way to distinguish the effects of the ellipticity e from those of three times the foreground shear γ_{od}.

In the following, we repeat this calculation for a main lens described by the more general elliptical power law lens profile described in section 1.1.5. We will assume without loss of generality that the elliptical and optical axes are aligned. From the definitions in section 1.1.5, we have that

$$\theta_1 = \frac{\zeta}{f}\cos(\varphi_{\mathrm{ell}}), \tag{1.277}$$

$$\theta_2 = \zeta\sin(\varphi_{\mathrm{ell}}). \tag{1.278}$$

It is useful to express the axis ratio f in terms of the ellipticity e, via

$$e = \frac{1-f}{1+f} \quad \Longrightarrow \quad f = \frac{1-e}{1+e}. \tag{1.279}$$

For small e, we can approximate

$$f = 1-2e. \tag{1.280}$$

Importantly, it should be noted that the relationship between f and e given in [153] is slightly different, and would give, to linear order, $f = 1-e$. The definition of e used in eq. (1.279) is chosen for consistency with the notation and hence the result presented in [59].

Following this, the polar angle φ can be related to φ_{ell} by expanding eq. (1.159) at first order in e, to obtain

$$\varphi_{\mathrm{ell}} = \varphi + e\sin(2\varphi). \tag{1.281}$$

We can then rewrite

$$\mathrm{e}^{\mathrm{i}\varphi_{\mathrm{ell}}} = \left[1+\mathrm{i}e\sin(2\varphi)\right]\mathrm{e}^{\mathrm{i}\varphi}, \qquad \mathrm{e}^{2\mathrm{i}\varphi_{\mathrm{ell}}} = \left[1+2\mathrm{i}e\sin(2\varphi)\right]\mathrm{e}^{2\mathrm{i}\varphi}. \tag{1.282}$$

We also have that

$$\zeta = \theta\left\{1-e\left[1+\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right]\right\}, \tag{1.283}$$

and so

$$\zeta^{2-\gamma_{\mathrm{PL}}} = \theta^{2-\gamma_{\mathrm{PL}}}\left\{1-e(2-\gamma_{\mathrm{PL}})\left[1+\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right]\right\}. \tag{1.284}$$

Now, for $|z|<1$, ${}_2F_1(a;b;c;z)$ can be approximated by

$${}_2F_1\left(a;b;c;z\right) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!} = 1+\frac{ab}{c}z+\frac{a(a+1)b(b+1)}{c(c+1)}\frac{z^2}{2}+..., \tag{1.285}$$

and therefore

$${}_2F_1\left(1;\frac{\gamma_{\mathrm{PL}}-1}{2};\frac{5-\gamma_{\mathrm{PL}}}{2};-\frac{1-f}{1+f}\mathrm{e}^{\mathrm{i}2\varphi_{\mathrm{ell}}}\right) = 1-e\frac{\gamma_{\mathrm{PL}}-1}{5-\gamma_{\mathrm{PL}}}\mathrm{e}^{2\mathrm{i}\varphi}. \tag{1.286}$$

The displacement angle is therefore given by

$$\underline{\alpha}(\zeta,\varphi_{\mathrm{ell}}) = (\theta'_E)^{\gamma_{\mathrm{PL}}-1}\theta^{2-\gamma_{\mathrm{PL}}}\mathrm{e}^{\mathrm{i}\varphi}\left\{1+e\left[\gamma_{\mathrm{PL}}-1+\mathrm{i}\sin(2\varphi)-(2-\gamma_{\mathrm{PL}})\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)-\frac{\gamma_{\mathrm{PL}}-1}{5-\gamma_{\mathrm{PL}}}\mathrm{e}^{2\mathrm{i}\varphi}\right]\right\}, \tag{1.287}$$

and its conjugate by

$$\underline{\alpha}^* = (\theta'_E)^{\gamma_{\rm PL}-1}\theta^{2-\gamma_{\rm PL}}\mathrm{e}^{-\mathrm{i}\varphi}\left\{1+e\left[\gamma_{\rm PL}-1-\mathrm{i}\sin(2\varphi)-(2-\gamma_{\rm PL})\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)-\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\mathrm{e}^{-2\mathrm{i}\varphi}\right]\right\}. \tag{1.288}$$

The products of these terms with $z^* = \theta\mathrm{e}^{-\mathrm{i}\varphi}$ and $z = \theta\mathrm{e}^{\mathrm{i}\varphi}$ respectively are

$$z^*\underline{\alpha} = (\theta'_E)^{\gamma_{\rm PL}-1}\theta^{3-\gamma_{\rm PL}}\left\{1+e\left[\gamma_{\rm PL}-1+\mathrm{i}\sin(2\varphi)-(2-\gamma_{\rm PL})\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)-\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\mathrm{e}^{2\mathrm{i}\varphi}\right]\right\}, \tag{1.289}$$

$$z\underline{\alpha}^* = (\theta'_E)^{\gamma_{\rm PL}-1}\theta^{3-\gamma_{\rm PL}}\left\{1+e\left[\gamma_{\rm PL}-1-\mathrm{i}\sin(2\varphi)-(2-\gamma_{\rm PL})\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)-\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\mathrm{e}^{-2\mathrm{i}\varphi}\right]\right\}, \tag{1.290}$$

and so

$$\psi_{\rm EPL} = \frac{(\theta'_E)^{\gamma_{\rm PL}-1}}{3-\gamma_{\rm PL}}\theta^{3-\gamma_{\rm PL}}\left[1+\frac{e}{2}\left(\gamma_{\rm PL}-1\right)-\frac{e}{2}\left(2-\gamma_{\rm PL}+\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\right)\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right]. \tag{1.291}$$

Absorbing the $1+e(\gamma_{\rm PL}-1)$ term into the definition of a new Einstein radius,

$$\theta_{\rm E} = \left[1+e(\gamma_{\rm PL}-1)\right]^{\frac{1}{\gamma_{\rm PL}-1}}\theta'_{\rm E}, \tag{1.292}$$

we obtain

$$\psi_{\rm EPL} = \frac{\theta_{\rm E}^{\gamma_{\rm PL}-1}\theta^{3-\gamma_{\rm PL}}}{3-\gamma_{\rm PL}}\left[1-e\left(2-\gamma_{\rm PL}+\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\right)\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right]. \tag{1.293}$$

To evaluate this at a position modified by the foreground shears, we must once again replace θ using 1.273. The modified polar angle will again result in a $\mathcal{O}(e\varphi)$ correction to $\psi_{\rm EPL}$, and can thus be ignored. Replacing θ, we obtain

$$\psi_{\rm EPL}[(1-\boldsymbol{\Gamma}_{od})\boldsymbol{\theta}] = \frac{\theta_{\rm E}^{\gamma_{\rm PL}-1}\theta^{3-\gamma_{\rm PL}}}{3-\gamma_{\rm PL}}\left\{1-\left[(3-\gamma_{\rm PL})\mathrm{Re}\left(\gamma_{od}\right)+e\left(2-\gamma_{\rm PL}+\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\right)\right]\mathrm{Re}\left(\mathrm{e}^{-2\mathrm{i}\varphi}\right)\right\}. \tag{1.294}$$

From this, it can easily be seen that substituting $\gamma_{\rm PL} = 2$ gives the expression for $\psi_{\rm SIE}[(\mathbf{1}-\boldsymbol{\Gamma}_{od})\boldsymbol{\theta}]$ given in eq. (1.276). To linear order in e and γ_{od}, an elliptical power law lens with ellipticity $e_{\rm eff}$ is indistinguishable from a lens with ellipticity e and in the presence of a foreground shear with

$$\gamma_i^{od} = \frac{1}{3-\gamma_{\rm PL}}\left(2-\gamma_{\rm PL}+\frac{\gamma_{\rm PL}-1}{5-\gamma_{\rm PL}}\right)\left(e_i^{\rm eff}-e_i\right). \tag{1.295}$$

In the language of [77], the isopotential contours of the effective potential $\psi_{\rm eff}$ are the images of the isopotential contours of the main lens potential ψ, distorted by the perturbations encoded in $\boldsymbol{\mathcal{A}}_{od}$. Because the ellipticity of the SIE and EPL lens models appear in exact form in the convergence, rather than the potential, this degeneracy is only true to first order for these models. However, a model in which the ellipticity was implemented directly in the lensing potential would render this degeneracy exact, and thus differentiating between e and γ_{od} from lensing observables is hopelessly model-dependent.

The mass-sheet degeneracy and the line-of-sight convergence

One famous consequence of the source position tranformation is the mass-sheet degeneracy, as discussed for single plane lensing in 1.2.4, which essentially renders measurements of the line-of-sight convergences κ_{os}, κ_{ds} and κ_{od} impossible without assumptions on the profile of the main lens, or measurements of observables such as the main lens velocity dispersion to provide constraints on the mass lens' mass. To illustrate this, suppose each amplification matrix was transformed according to

$$\boldsymbol{\mathcal{A}}_{os}^{\lambda} = (1-\lambda_{os})\boldsymbol{\mathcal{A}}_{os}, \tag{1.296}$$

$$\boldsymbol{\mathcal{A}}_{od}^{\lambda} = (1-\lambda_{od})\boldsymbol{\mathcal{A}}_{od}, \tag{1.297}$$

$$\boldsymbol{\mathcal{A}}_{ds}^{\lambda} = (1-\lambda_{ds})\boldsymbol{\mathcal{A}}_{ds}. \tag{1.298}$$

Parameterised as normal in terms of convergence and shear, the resulting matrices would take the form

$$\boldsymbol{\mathcal{A}}_{ij}^{\lambda} = \begin{pmatrix} 1-\lambda_{ij}-(1-\lambda_{ij})\kappa_{ij}-(1-\lambda_{ij})\gamma_{ij}^1 & -(1-\lambda_{ij})\gamma_{ij}^2 \\ -(1-\lambda_{ij})\gamma_{ij}^2 & 1-\lambda_{ij}-(1-\lambda_{ij})\kappa_{ij}+(1-\lambda_{ij})\gamma_{ij}^1 \end{pmatrix}. \tag{1.299}$$

From this result, we can read off

$$\kappa_{ij} = (1-\lambda_{ij})\kappa_{ij} + \lambda_{ij}, \tag{1.300}$$
$$\gamma_{ij} = (1-\lambda_{ij})\gamma_{ij}, \tag{1.301}$$

and

$$\boldsymbol{\Gamma}^{\lambda}_{ij} = (1-\lambda_{ij})\boldsymbol{\Gamma}_{ij} + \lambda_{ij}\mathbb{1}. \tag{1.302}$$

Now, consider the minimal lens model, 1.265. Under the above transformation, this becomes

$$\tilde{\boldsymbol{\beta}}^{\lambda} = \boldsymbol{\mathcal{A}}^{\lambda}_{\rm LOS}\boldsymbol{\theta} - \frac{\mathrm{d}\psi^{\lambda}_{\rm eff}}{\mathrm{d}\boldsymbol{\theta}}, \tag{1.303}$$
$$\frac{1-\lambda_{od}}{1-\lambda_{ds}}\tilde{\boldsymbol{\beta}} = \frac{(1-\lambda_{od})(1-\lambda_{os})}{1-\lambda_{ds}}\boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta} - \frac{\mathrm{d}\psi_{\rm eff}[(1-\lambda_{od})\boldsymbol{\theta}]}{\mathrm{d}\boldsymbol{\theta}}. \tag{1.304}$$

Multiplying throughout by $(1-\lambda_{ds})(1-\lambda_{od})^{-1}(1-\lambda_{os})^{-1}$,

$$\frac{1}{1-\lambda_{os}}\tilde{\boldsymbol{\beta}} = \boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta} - \frac{1-\lambda_{ds}}{(1-\lambda_{od})(1-\lambda_{os})}\frac{\mathrm{d}\psi_{\rm eff}[(1-\lambda_{od})\boldsymbol{\theta}]}{\mathrm{d}\boldsymbol{\theta}}, \tag{1.305}$$

or equivalently,

$$\tilde{\boldsymbol{\beta}}' = \boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta} - \frac{\mathrm{d}\psi'_{\rm eff}}{\mathrm{d}\boldsymbol{\theta}}, \tag{1.306}$$

where

$$\tilde{\boldsymbol{\beta}}' \equiv \frac{1}{1-\lambda_{os}}\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\mathcal{A}}^{-1}_{ds}\boldsymbol{\beta}, \tag{1.307}$$
$$\psi'_{\rm eff}(\boldsymbol{\theta}) \equiv \frac{1-\lambda_{ds}}{(1-\lambda_{od})(1-\lambda_{os})}\psi_{\rm eff}[(1-\lambda_{od})\boldsymbol{\theta}]. \tag{1.308}$$

As the source position is unobservable, it is clear that $\tilde{\boldsymbol{\beta}}'$ is degenerate with $\tilde{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}$. The $(1-\lambda_{ds})(1-\lambda_{od})^{-1}(1-\lambda_{os})^{-1}$ term which multiplies $\psi_{\rm eff}$ can be fully absorbed into the lens model itself by uniformly rescaling its potential by the same factor. Finally, the $(1-\lambda_{od})$ term within the argument of $\psi_{\rm eff}$ is degenerate with a more or less steeply sloped profile. From this and eq. (1.301), it is clear that, without additional information about the lens profile, the line-of-sight convergences are degenerate with properties of the source and the main lens. The line-of-sight shears are also scaled by the unobservable $(1-\lambda^{\rm ms})$ factors, and only the reduced shear terms are unaffected by this degeneracy.

Sections 1.2.4 and 1.3.4 show that the mass-sheet degeneracy is present both in the convergence of the dominant lens and in that of the line-of-sight. The effects manifest identically, and time delays and kinematics are affected by the product of the two [96, 134, 138, 22]. Dominant lens convergence κ is not readily distinguishable from line-of-sight convergence κ_{ij}, and these in turn are not distinguishable from their mass-sheet transformed counterparts κ^{λ} and κ^{λ}_{ij}. Without additional constraints, any restrictions on the λ terms are fully dependent on the specific choice of lens model.

Parameterised line-of-sight effects on lensing

Equations (1.262) and (1.263) add 9 parameters to the lens model, in the form of 3 convergences $(\kappa_{od}, \kappa_{os}, \kappa_{ds})$ and 3 complex shears $(\gamma_{od}, \gamma_{os}, \gamma_{ds})$, which make up the 3 shear matrices $\boldsymbol{\Gamma}_{od}, \boldsymbol{\Gamma}_{os}, \boldsymbol{\Gamma}_{ds}$. However, from the discussion above, it is clear that only 2 of these parameters are readily measurable from strong lensing observables alone. The effects of $\boldsymbol{\Gamma}_{os}$ and $\boldsymbol{\Gamma}_{ds}$ are tied up with each other and with $\boldsymbol{\Gamma}_{od}$, and can only be measured in combination as $\boldsymbol{\Gamma}_{\rm LOS} \approx \boldsymbol{\Gamma}_{od} + \boldsymbol{\Gamma}_{os} - \boldsymbol{\Gamma}_{ds}$. The convergences are degenerate with the scaling and slope of the main lens profile, and the foreground shear γ_{od} is degenerate with the ellipticity of the main lens. Therefore, $\kappa_{\rm LOS}$, κ_{od} and γ_{od} can all be absorbed into a sufficiently complex lens model, and without independent constraints of this model, are not readily measurable. Thus, in practice, the minimal lens model reduces the number of relevant parameters to be included in a lens model to 2: the components of the complex shear $\gamma_{\rm LOS}$.

This is generally how lenses are modelled, with a parametric main lens model supplemented with a 2-component external shear, treated as a nuisance parameter meant to account for both line-of-sight effects and complexities within the main lens itself. This treatment is not without justification, and fitting real lensing images with overly simplified lens models can render any measurements of the external shear

meaningless. This was illustrated in [31], in which using simple elliptical models with external shear to fit realistically complex mock datasets with no external shear resulted in spurious measurements of external shears of up to 8%. Nonetheless, in [77], it is shown that accurate and precise measurements of $\gamma_{\rm LOS}$ are possible from complex mock images, if the main lens model is sufficiently rich. This point, and the role played by the foreground shear, is investigated in detail in chapter 3.

In the lens equation and associated calculations, the $\mathbf{\Gamma}_{\rm LOS}$ term appears to occupy an equivalent role to $\mathbf{\Gamma}_{os}$, were $\mathbf{\Gamma}_{od}$ and $\mathbf{\Gamma}_{ds}$ set to zero. This, naively, is how the $\mathbf{\Gamma}_{\rm ext}$ correction used in fits could be interpreted, and were this accurate, it would be identified as the same external convergence and shear present in weak lensing observations. Were this the case, the contributions of the secondary perturbers would peak midway between the observer and source. However, the presence of the distance ratios in eq. (1.218) mean that a perturber located between the observer and the lens will contribute more to $\mathbf{\Gamma}_{\rm LOS}$ than an identical perturber between the lens and the source. The full expression for $\gamma_{\rm LOS}$ takes the form

$$\gamma_{\rm LOS}(\boldsymbol{\theta}) = -\frac{3}{2}\Omega_{m0}H_0^2\int_0^{\chi_s}\mathrm{d}\chi(1+z)W_{\rm LOS}(\chi)\int_{\mathbb{R}^2}\frac{\mathrm{d}^2\boldsymbol{x}}{\pi x^2}\mathrm{e}^{2\mathrm{i}\xi}\delta[\eta_0-\chi,\chi,f_K(\chi)\boldsymbol{\theta}+\boldsymbol{x}], \tag{1.309}$$

with the weighting function

$$W_{\rm LOS}(\chi)\equiv\begin{cases}\frac{f_K(\chi_s-\chi)f_K(\chi)}{f_K(\chi_s)}+\frac{f_K(\chi_d-\chi)f_K(\chi)}{f_K(\chi_d)} & \chi\leq\chi_d,\\ \frac{f_K(\chi_s-\chi)f_K(\chi)}{f_K(\chi_s)}-\frac{f_K(\chi_s-\chi)f_K(\chi-\chi_d)}{f_K(\chi_s-\chi_d)} & \chi\geq\chi_d.\end{cases} \tag{1.310}$$

$\delta(\eta,\chi,\boldsymbol{x})$ is the matter density contrast at conformal time η, comoving radius χ and transverse position $\boldsymbol{x}=x(\cos\xi,\sin\xi)$ [86, 56, 58]. For comparison, the standard cosmic shear measured in weak lensing has a weighting function

$$W_{\rm st}(\chi)=\frac{f_K(\chi_s-\chi)f_K(\chi)}{f_K(\chi_s)}, \tag{1.311}$$

thus the standard cosmic shear cannot be directly equated to $\gamma_{\rm LOS}$.

The parameterisation of the minimal lens model given in [59] accounts fully for the degeneracies between the contributions of different perturbers, at least for simple models. As pointed out in the same paper, the parameterisation proposed in [23] is likely too ambitious, and fails to account for the linear contribution of γ_{od} to the displacement angle. Thus the weighting function in eq. (2.4) of [23] does not fully incorporate the differences in contributions between the foreground and background lenses. This could go some way towards explaining some of the discrepancies seen between the strong lensing and weak lensing shear estimates in [101], in which this formalism was applied.

Constraining the mass-sheet degeneracy

As discussed in sections 1.2.4 and 1.3.4, the biggest challenge to measurements of the Hubble constant via time delay cosmography is the mass-sheet degeneracy. Several techniques for constraining the convergence have been proposed and implemented. Constraints on the line-of-sight convergence are possible via estimates of the large scale structure in the vicinity of the lens. In [114, 52, 8, 172], mass profiles were fitted to known groups near the lenses in question, while [149] obtained galaxy counts in the lens field from [53], and measured the external convergence in mock fields taken from the Millenium Simulation [146] with the same properties as the real field. This approach was further refined by [67]. In [23], the relationship between the stellar mass of galaxies to the underlying mass distribution of a region of space was combined with measurements of the external shear from the strong lensing image, resulting in improved constraints over those made in the absence of the external shear measurements. Finally, [55, 118, 50, 157] measured the external shear in the vicinity of the lens via weak lensing and integrated these measurements to estimate the external convergence.

Although important steps towards constraining the line-of-sight convergence, estimates from weak lensing shear measurements in the proximity of the lens introduce a bias into the measurement, given the different weighting functions of $\kappa_{\rm ext}$ in weak lensing and $\kappa_{\rm LOS}$ in strong lensing, as discussed in section 1.3.4. This problem also affects estimates from galaxy counts or mass fitting nearby groups, as the contribution of foreground mass will be underestimated relative to that in the background of the lens if the convergence is not properly parameterised. This motivated the work of [109], in which the tidal multi-plane formalism introduced in [108] to properly capture these non-linear effects was shown to produce more reliable results from mock data than those which did not account for the effects of foreground perturbers.

The most important and frequently used method for constraining the mass-sheet degeneracy makes use of velocity dispersion data. Some recent examples of this include [132, 20, 37, 177, 111, 38, 144]. However, in each of these studies, the convergence was parameterised as $\kappa_{\rm ext} = \kappa_{\rm os}$ only, and the κ_{od} and κ_{ds} terms are omitted. [151] points out that this can lead to non-negligible biases in the inferred convergence, as, in general, weak lensing effects manifest differently in kinematics and imaging. This observation is also made in [64], who argue that lensing degeneracies can appear in the process of constraining the external convergence, which can bias the recovery of H_0.

The importance of the foreground convergence can be demonstrated as follows. Ignoring the effects of shear for now, and following the approach of [151], but absorbing κ_{ds} and κ_{os} into the $\kappa_{\rm LOS}$ term, the displacement angle under the DL approximation and in the tidal regime (making the source position transformation described in section 1.3.4) can be expressed as

$$\boldsymbol{\alpha}'(\boldsymbol{\theta}) = \frac{1-\kappa_{od}}{1-\kappa_{\rm LOS}}\boldsymbol{\alpha}\left[(1-\kappa_{od})\boldsymbol{\theta}\right]. \tag{1.312}$$

A spherically-symmetric power law lens has a deflection angle which is given by

$$\boldsymbol{\alpha}_{\rm EPL} = \left(\frac{\theta}{\tilde{\theta}_{\rm E}}\right)^{1-\gamma_{\rm PL}}\boldsymbol{\theta}. \tag{1.313}$$

Equation (1.312) can therefore be expressed as

$$\boldsymbol{\alpha}'(\boldsymbol{\theta}) = \frac{1-\kappa_{od}}{1-\kappa_{\rm LOS}}\left(\frac{|(1-\kappa_{od})\boldsymbol{\theta}|}{\tilde{\theta}_{\rm E}}\right)^{1-\gamma_{\rm PL}}(1-\kappa_{od})\boldsymbol{\theta}, \tag{1.314}$$

$$= \frac{(1-\kappa_{od})^{3-\gamma_{\rm PL}}}{1-\kappa_{\rm LOS}}\left(\frac{\theta}{\tilde{\theta}_{\rm E}}\right)^{1-\gamma_{\rm PL}}\boldsymbol{\theta}, \tag{1.315}$$

$$:= \left(\frac{\theta}{\theta_{\rm E}}\right)^{1-\gamma_{\rm PL}}\boldsymbol{\theta}. \tag{1.316}$$

The last equation is how this power law is often modelled, with the choice to choose apply an unmeasurable mass-sheet transform $\lambda_{ij} = 1/(1-\kappa_{ij})$, which eliminates the external convergences from the model. The power law parameter $\tilde{\theta}_{\rm E}$ and the observed Einstein angle $\theta_{\rm E}$ are therefore related via

$$\theta_{\rm E} = \left[\frac{(1-\kappa_{od})^{3-\gamma_{\rm PL}}}{1-\kappa_{\rm LOS}}\right]^{\frac{1}{\gamma_{\rm PL}-1}}\tilde{\theta}_{\rm E}. \tag{1.317}$$

The observable velocity dispersion for the PL profile, where $\tilde{\boldsymbol{\theta}} = (1-\kappa_{od})\boldsymbol{\theta}$ is the angle modified by foreground lensing, is (see, for example, [98])

$$\sigma^2(\boldsymbol{\theta}) = 2G\Sigma_{\rm crit}D_{od}\frac{\sqrt{\pi}\Gamma(\frac{\gamma_{\rm PL}}{2})}{\Gamma(\frac{\gamma_{\rm PL}-1}{2})}\tilde{\theta}_{\rm E}^{\gamma_{\rm PL}-1}\tilde{\theta}^{2-\gamma_{\rm PL}}, \tag{1.318}$$

$$= \left[\frac{1-\kappa_{\rm LOS}}{(1-\kappa_{od})^{3-\gamma_{\rm PL}}}\right](1-\kappa_{od})^{2-\gamma_{\rm PL}}\frac{D_{os}}{D_{ds}}J(\theta,\theta_{\rm E},\gamma_{\rm PL}), \tag{1.319}$$

$$= \frac{1-\kappa_{\rm LOS}}{1-\kappa_{od}}\frac{D_{os}}{D_{ds}}J(\theta,\theta_{\rm E},\gamma_{\rm PL}). \tag{1.320}$$

The angular diameter distance terms are introduced by $\Sigma_{\rm crit}$, and their ratio D_{os}/D_{ds} ensures that the Hubble constant factor cancels out. The function $J(\theta,\theta_{\rm E},\gamma_{\rm PL})$ depends on parameters of the main lens, but not on the external convergences or shears, and nor on cosmological parameters. In the original publication of [151], the weak lensing correction to the angle at which the velocity dispersion is measured was not included, but following a discussion with the authors, the updated version on the archive reflects this correction.

A typical approach to constrain the convergence via the velocity dispersion would be to obtain measurements of $\sigma^2(\boldsymbol{\theta})$, and compare these to predictions of $(D_{os}/D_{ds})J(\theta,\theta_{\rm E},\gamma_{\rm PL})$ from imaging observables (excluding any convergence from the model). The factor by which these terms differed would then be identified as $(1-\kappa_{\rm ext})$ (which we will denote as $(1-\kappa_{\rm ext}^{\sigma^2})$), and is then used to correct measurements of Hubble constant via

$$H_0^{\rm inferred} = (1-\kappa_{\rm ext}^{\sigma^2})H_0^{\rm model} = \frac{1-\kappa_{\rm LOS}}{1-\kappa_{od}}H_0, \tag{1.321}$$

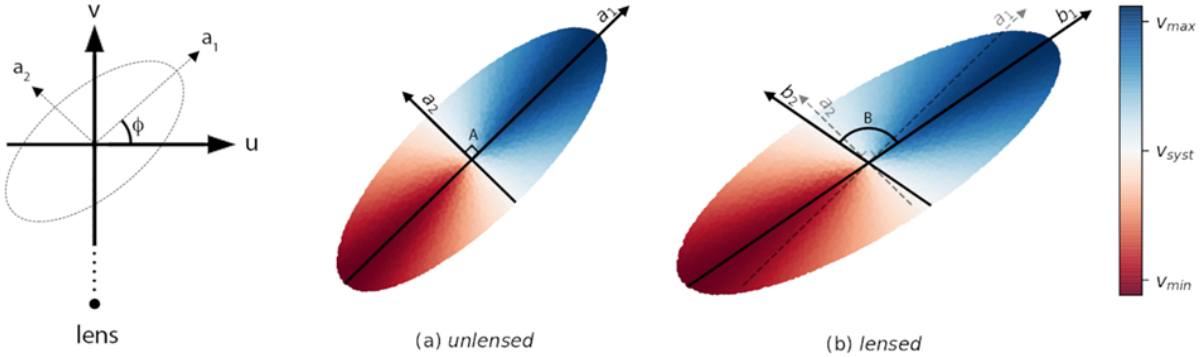

Figure 1.10: *Weak lensing distortions on velocity field, taken from [69]. The lensing axis is denoted by v and u, while the source has major and minor axes a_1 and a_2 respectively, at an angle ϕ with the lens axis. If a lens (with $\kappa = 0.3$ and $\gamma = 0.3$ in this example) is placed below the source, the orthogonality of the axes (now b_1 and b_2) defined according to the stellar velocities is broken.*

where H_0^{model} is the value of H_0 measured from time delay data with a model without convergence, and H_0^{inferred} is the value inferred after the correction is applied. However, the $\kappa_{\text{ext}}^{\sigma^2}$ correction factor is not the same as that which appears in the calculations governing imaging observables. The external convergence acts as a mass-sheet transformation analogous to the internal mass-sheet degeneracy (see section 1.3.4), and so, as shown in section 1.2.4, the true correction needed is

$$H_0 = (1 - \kappa_{\text{LOS}}) H_0^{\text{model}}. \tag{1.322}$$

Therefore, the inferred value of H_0 will differ from the true value by a factor of

$$\frac{H_0^{\text{inferred}}}{H_0} = \frac{1}{1 - \kappa_{od}} \approx 1 + \kappa_{od}. \tag{1.323}$$

This example is of course a simplified version of the real procedure used to constrain κ_{ext} via the velocity dispersion, in which uncertainties from the inference of parameters such as γ_{PL} must be taken into account, as well as the role of the point spread and luminosity weighting functions, and projection effects must also be considered. Nonetheless, it demonstrates that the convergence terms affecting the velocity dispersion and lensing observables are not straightforwardly factored out. While [151] recommends modelling with κ_{os}, κ_{ds} and κ_{od} as three separate parameters, the degeneracy between these terms means that they, as in section 1.3.4, be absorbed into the terms κ_{LOS} and κ_{od}. However, ignoring the effects of κ_{od} can lead to a fractional bias on the order of that foreground shear.

Could the foreground shear be constrained with velocity dispersion data?

In section 1.3.4, we pointed out that the ellipticity introduced to an image by the foreground shear is almost impossible to distinguish from the intrinsic ellipticity of the lens galaxy. However, if the velocity field of the lens galaxy is assumed to be well approximated by pure circular rotation (e.g. [113]), then distortions introduced by the foreground shear will result in deviations from axissymmetry in the lens velocity field compared with the observed light distribution. Velocity gradients will therefore no longer be perpendicular to their displacement from the centre of the galaxy, as illustrated in fig. 1.10. This has the potential to introduce further biases when constraining the mass profile, but may also constitute a measurable signal. This phenomenon was first proposed as an opportunity for weak lensing shear measurements in [24], and then later in [115, 42, 43]. It has been used to obtain shear measurements from a sample of weak lens galaxies in [69], but, to our best knowledge, has not yet been proposed as a constraint on the foreground shear in strong lensing images.

The observed velocity field of a galaxy is modified by the foreground according to

$$v_{\text{obs}}(\boldsymbol{\theta}) = v(\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}). \tag{1.324}$$

The effect of κ_{od} is simply to magnify the velocity field, without introducing measurable distortions. [69] point out that, while the effects of γ_1 (if chosen to align with the principle axis of the elliptical source)

preserve axissymmetry, those of γ_2 do not, and would thus escape degeneracy with the ellipticity. In several measurements from individual galaxies, this same study found that the shear between galaxy and observer could be contrained to within $\sigma_\gamma \sim 0.01$, with an estimated dynamical shape noise of $\sigma_{\rm dsn} \sim 0.03$ (capturing deviations from pure circular rotation), and so constraints within 0.05 on the foreground shear seem achievable with modern velocity dispersion data. This would represent a much tighter constraint on γ_{od} than was possible from the majority of lenses in appendix B.1, despite these measurements being made from a somewhat idealised mock dataset, and, crucially, despite having made strong assumptions about the underlying mass distribution. While the fact that only γ_2^{od} breaks axissymmetry means that this technique is somewhat limited when applied to a single galaxy, this problem can be immediately resolved by applying the same technique to any nearby galaxy experiencing the same foreground shear, for which the ellipticity would no longer align with this chosen shear axis, and thus both components could be extracted via a joint analysis. It therefore seems that precision weak lensing could serve as a powerful tool when constraining the foreground shear, which in turn has the potential to improve constraints on both κ_{od} and $\gamma_{\rm LOS}$.

1.4 Line-of-sight effects on time delays

A key area of focus for this project is on the effects of the line-of-sight convergence and shear on the time delays between strong lensing images. As in the case of the lens equation, degeneracies abound, and the focus of this section will therefore be to arrive at a general expression for time delays in the presence of tidal secondary deflectors which accounts for the degeneracies present therein. We begin by presenting the expressions for these time delays in the contexts of multi-plane lensing (section 1.4.1) and the DL approximation (section 1.4.2), before discussing the expression which applies to the case where all but the main deflector are tidal (section 1.4.3).

1.4.1 Multi-plane lensing time delays

In the multi-plane formalism, the delay between two images A and B of the same source $\boldsymbol{\beta}$, characterised by the paths $\{\boldsymbol{x}_l^A\}$ and $\{\boldsymbol{x}_l^B\}$ through the N planes, reads $\Delta t_{AB} = T(\{\boldsymbol{x}_l^A\}) - T(\{\boldsymbol{x}_l^B\})$, with

$$T(\boldsymbol{x}_l) = \sum_{l=1}^{N} \left[\frac{1}{2} \tau_{l(l+1)} |\boldsymbol{\beta}_{o(l+1)} - \boldsymbol{\beta}_{ol}|^2 - (1+z_l)\hat{\psi}_l(\boldsymbol{x}_l) \right], \tag{1.325}$$

where the time delay distance τ_{ij} between planes i and j is defined as

$$\tau_{ij} \equiv (1+z_i)\frac{D_{oi}D_{oj}}{D_{ij}}, \tag{1.326}$$

and, as before, $\beta_{ol} = \boldsymbol{x}_l/D_{ol}$ [59]. $T(\{\boldsymbol{x}_l\})$ is the difference in the length of a null geodesic in the presence of the N lenses, and the length of the null geodesic if these lenses were absent, and only the background FLRW spacetime was present. As before, we can interpret this as the contributions of a geometric and gravitational term, although, as before, this distinction is not fundamental. The first terms in the sum arise that the path length between lens l and $l+1$ followed by the light differs slightly from the undeflected path. A longer path leads to a longer time delay. The second term related to the gravitational time delay. Since the potential $\hat{\psi}$ is negative, this term is also positive, and increases as the potential becomes stronger. The total time delay along a specific path is given by the sum of these two terms evaluated for each of the lens planes.

1.4.2 Time delays in the dominant lens approximation

Under the DL approximation, eq. (1.325) reduces to

$$\begin{aligned} T(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}\tau_{ds}|\boldsymbol{\theta} - \boldsymbol{\alpha}_{od} - \boldsymbol{\beta}|^2 - (1+z_d)\hat{\psi}_d[D_{od}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{od})] \\ - \sum_{l<d}(1+z_l)\hat{\psi}_l(D_{ol}\boldsymbol{\theta}) - \sum_{l>d}(1+z_l)\hat{\psi}_l[D_{ol}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl})] + \mathcal{O}(\epsilon^4). \end{aligned} \tag{1.327}$$

A full derivation of this can be found in [59]. The first term relates to the additional path length travelled by light which is deflected by an amount $\boldsymbol{\theta} - \boldsymbol{\alpha}_{od} - \boldsymbol{\beta}$ (in other words, the geometric time delay caused

by the main lens, evaluated at a position at which it is evaluated is post-Born, and takes the effects of the foreground perturbers into account). The second term is the gravitational time delay due to the main lens, evaluated at position $\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}$. The third and fourth terms are the gravitational time delays of the foreground and background lenses respectively. The foreground gravitational time delays may be evaluated at $\boldsymbol{\theta}$ (the Born approximation), and the background delays must be evaluated at $\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}$, or in other words, at a position which only takes the dominant lens into account. Notice that the geometric time delay term is only present for the dominant lens, to $\mathcal{O}(\epsilon^2)$.

1.4.3 Time delays in the tidal regime

An approximate solution

In the tidal regime and the DL approximation, the time delay given in eq. (1.327) reduces to the form

$$T(\boldsymbol{\theta}, \boldsymbol{\beta}) = T_d(\boldsymbol{\theta}, \boldsymbol{\beta}') + \delta T(\boldsymbol{\beta}'), \tag{1.328}$$

with

$$T_d(\boldsymbol{\theta}, \boldsymbol{\beta}') = \frac{1}{2}\tau_{ds}(\boldsymbol{\theta} - \boldsymbol{\beta}') \cdot (\mathbf{1} - \boldsymbol{\Gamma}_{os} - \boldsymbol{\Gamma}_{od} + \boldsymbol{\Gamma}_{ds})(\boldsymbol{\theta} - \boldsymbol{\beta}') - (1 + z_d)\hat{\psi}_d[D_{od}(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}], \tag{1.329}$$

$$\delta T(\boldsymbol{\beta}') = -\sum_{l \neq d}(1 + z_l)\hat{\psi}_l(D_{ol}\boldsymbol{\beta}'), \tag{1.330}$$

where $\boldsymbol{\beta}' = \boldsymbol{\beta} + \boldsymbol{\alpha}_{os}(\boldsymbol{\beta}')$ is the weakly-lensed source position from section 1.3.4. A derivation of this equation can once again be found in [59].

Written as above, we see that the total time delay $T(\boldsymbol{\theta}, \boldsymbol{\beta})$ can be separated into $T_d(\boldsymbol{\theta}, \boldsymbol{\beta}')$, which corresponds to the time delay relative to the weakly-lensed ray which would be observed at $\boldsymbol{\beta}'$, and $\delta T(\boldsymbol{\beta}')$, which is the time delay between that ray and a ray propagating through an FLRW background in the absence of all perturbations (which would be observed at $\boldsymbol{\beta}$. Notice that there are no geometric terms present in $\delta T(\boldsymbol{\beta}')$, as these are all $\mathcal{O}(\epsilon^4)$.

Since only the time delay *between* two images from the same source is measurable, $\delta T(\boldsymbol{\beta}')$ is fundamentally unobservable. Therefore, omitting this term, substituting the minimal lens model eq. (1.265) in place of $\boldsymbol{\theta} - \boldsymbol{\beta}'$ and converting $\hat{\psi}_d$ to ψ_{ods},

$$T(\boldsymbol{\theta}) = \frac{\tau_{ds}}{2}\boldsymbol{\alpha}'(\boldsymbol{\theta}) \cdot (\mathbf{1} - \boldsymbol{\Gamma}_{os} - \boldsymbol{\Gamma}_{od} + \boldsymbol{\Gamma}_{ds})\boldsymbol{\alpha}'(\boldsymbol{\theta}) - \frac{(1 + z_d)D_{od}D_{os}}{D_{ds}}\psi_{ods}[D_{od}(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}] \tag{1.331}$$

$$= \tau_{ds}\frac{1}{2}(\mathbf{1} - \boldsymbol{\Gamma}_{\text{LOS}})^{-1}\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) \cdot (\mathbf{1} - \boldsymbol{\Gamma}_{\text{LOS}})(\mathbf{1} - \boldsymbol{\Gamma}_{\text{LOS}})^{-1}\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) - \tau_{ds}\psi_{\text{eff}}(\boldsymbol{\theta}) \tag{1.332}$$

$$= \tau_{ds}\left[\frac{1}{2}\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) \cdot (\mathbf{1} - \boldsymbol{\Gamma}_{\text{LOS}})^{-1}\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) - \psi_{\text{eff}}(\boldsymbol{\theta})\right], \tag{1.333}$$

we arrive at

$$T(\boldsymbol{\theta}) = \tau_{ds}\left[\frac{1}{2}\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) \cdot (\mathbf{1} + \boldsymbol{\Gamma}_{\text{LOS}})\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) - \psi_{\text{eff}}(\boldsymbol{\theta})\right], \tag{1.334}$$

where

$$\psi_{ods}(\boldsymbol{\beta}_{od}) \equiv \frac{D_{ds}}{D_{od}D_{os}}\hat{\psi}_d(D_{od}\boldsymbol{\beta}_{od}), \tag{1.335}$$

and

$$\boldsymbol{\alpha}_{\text{eff}}(\boldsymbol{\theta}) = \frac{\mathrm{d}\psi_{\text{eff}}}{\mathrm{d}\boldsymbol{\theta}} = (\mathbf{1} - \boldsymbol{\Gamma}_{\text{LOS}})\boldsymbol{\alpha}'(\boldsymbol{\theta}), \tag{1.336}$$

with $\boldsymbol{\alpha}'(\boldsymbol{\theta})$ defined as before in eq. (1.312). These results confirm that, in the analysis of strong-lensing time delays, two degeneracies emerge - firstly, between $\boldsymbol{\Gamma}_{od}$ and the intrinsic properties of the main lens within ψ_{eff}, and secondly, between the external convergences and shears within $\boldsymbol{\Gamma}_{\text{LOS}}$.

A more general solution

Equation (1.334) gives an expression for the time delay experienced by a strongly-lensed signal in the presence of perturbative secondary deflectors in the tidal regime and DL approximation, which can be used to calculate the relative time delay between two such signals originating from the same source. A more general expression for this time delay is possible, in which the influence of the secondary deflectors

is assumed to be tidal but not necessarily small. In [58], the case where tidal effects can be arbitrarily large and the background spacetime is not assumed to be FLRW is considered. Equation (11) in this paper gives an expression for the time delay in these conditions as

$$T(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}')\cdot\omega_o\boldsymbol{\mathcal{D}}_{od}^T\boldsymbol{\mathcal{D}}_{ds}^{-1}\boldsymbol{\mathcal{D}}_{os}(\boldsymbol{\theta}-\boldsymbol{\beta}') - (1+z_d)\hat{\psi}[\omega_o\boldsymbol{\mathcal{D}}_{od}\boldsymbol{\theta}]. \tag{1.337}$$

Equation 30 in [58] gives an expression for this time delay when the background spacetime is assumed to be FLRW, which, written in the notation previously introduced, takes the form

$$T(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}\tau_{ds}(\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta})\cdot\boldsymbol{\mathcal{A}}_{od}^{\intercal}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os}(\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta}) - (1+z_d)\hat{\psi}[D_{od}\boldsymbol{\mathcal{A}}_{od}\boldsymbol{\theta}], \tag{1.338}$$

or, expressed in terms of the weakly-lensed source position $\boldsymbol{\beta}'$,

$$T(\boldsymbol{\theta},\boldsymbol{\beta}) = \tau_{ds}\left[\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}')\cdot\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\beta}') - \psi_{\rm eff}(\boldsymbol{\theta})\right], \tag{1.339}$$

with

$$\boldsymbol{\mathcal{A}}_{\rm LOS} \equiv \boldsymbol{\mathcal{A}}_{od}^{\intercal}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os}, \tag{1.340}$$

$$\boldsymbol{\beta}' = \boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta}, \tag{1.341}$$

and $\psi_{\rm eff}(\boldsymbol{\theta})$ as defined in eq. (1.268). If $\boldsymbol{\mathcal{A}}_{od}$, $\boldsymbol{\mathcal{A}}_{os}$ and $\boldsymbol{\mathcal{A}}_{ds}$ are assumed to be small, then this expression reduces to that given in eq. (1.334), as shown in appendix A.3.2. Note that the transpose of $\boldsymbol{\mathcal{A}}_{od}$ will equal $\boldsymbol{\mathcal{A}}_{od}$ if this matrix is assumed to be symmetric, which is often the case. However, if rotation terms are included in its parameterisation for extra flexibility in a model, than this will no longer be the case.

The Fermat Potential

The traditional Fermat Potential $\phi(\boldsymbol{\theta},\boldsymbol{\beta})$ serves a dual purpose. When multiplied by the time delay distance τ_{ds}, it gives the time delay of a light ray relative to the path it would follow in the absence of any gravitational fields. By Fermat's Principle in section 1.2.1, setting the derivative of the Fermat Potential with respect to $\boldsymbol{\theta}$ equal to zero allows the lens equation to be derived.

Equation (1.327), the time delay function in the DL approximation, was derived by ignoring any terms which depend only on the source position, and which are therefore unobservable. Its derivation also made use of the lens equation, meaning that it is only valid for combinations of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ which satisfy eq. (1.265). Thus, although similar in appearance to a Fermat Potential multiplied by the time delay distance, it cannot be interpreted in the same way. In the tidal regime, eq. (1.339) does not give the absolute time delay of a ray relative to the unlensed path, but instead allows relative time delays between multiple solutions of the lens equation corresponding to the same $\boldsymbol{\beta}$ to be calculated or modelled. However, its derivative does nonetheless lead to the lens equation, as can be shown as follows.

In the notation of [59], the Fermat potential $\phi(\boldsymbol{\theta},\boldsymbol{\beta})$ is such that $\partial\phi/\partial\boldsymbol{\theta} \propto \boldsymbol{L} \equiv \boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta} - \frac{\mathrm{d}\psi_{\rm eff}}{\mathrm{d}\boldsymbol{\theta}} - \tilde{\boldsymbol{\beta}}$, which can be set to 0 for true physical trajectories of the light ray. We consider the derivatives of T.

$$\frac{1}{\tau_{ds}}\frac{\partial T}{\partial\boldsymbol{\theta}} = \frac{1}{2}\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\beta}') + \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}')\cdot\frac{\partial\boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta}}{\partial\boldsymbol{\theta}} - \frac{\partial\psi_{\rm eff}}{\partial\boldsymbol{\theta}}. \tag{1.342}$$

Using

$$\boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\beta}' = \boldsymbol{\mathcal{A}}_{od}^{\intercal}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}, \tag{1.343}$$

and the fact that $\boldsymbol{\mathcal{A}}_{\rm LOS}$ is a symmetric matrix (when ignoring rotations), we arrive at

$$\frac{1}{\tau_{ds}}\frac{\partial T}{\partial\boldsymbol{\theta}} = \boldsymbol{\mathcal{A}}_{\rm LOS}\boldsymbol{\theta} - \frac{\partial\psi_{\rm eff}}{\partial\boldsymbol{\theta}} - \tilde{\boldsymbol{\beta}}. \tag{1.344}$$

We can therefore identify the Fermat potential in the tidal regime as

$$\phi(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}')\cdot\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\beta}') - \psi_{\rm eff}(\boldsymbol{\theta}). \tag{1.345}$$

Chapter 2

`lenstronomy` Implementation

As the next generation of telescopes begin scanning the sky, the number of candidate lensing images has increased exponentially [103, 171, 122, 79, 82, 80], and is expected to exceed 100 000 in the space of a few years [39]. Several open source strong lens modelling software packages have been developed in the last decade or so to meet the challenge of analysing this rapidly expanding dataset. Some of these are built to handle all of the key steps in strong lens modelling, such as `gravlens` [89], `lenstool` [92], `LENSED` [152], `PixeLens` [135], which is built specifically for multiply-imaged point sources, `Ensai` [72], which carries out parameter inference with convolutional neural networks, rather than the more typical, slower, maximum likelihood methods, and `PyAutoLens` [119], which incorporates a grid source reconstruction method, rather than relying only on traditionally parameterised source models. Others act more as support packages, carrying out specific lens modelling-related tasks, such as `glafic` [121], and `PySPT` [170], built primarily to quantify the systematic uncertainties associated with the degeneracy of the source-position transform.

As the field of gravitational lensing continues to expand, the theoretical framework to tackle each facet of strong lens analysis becomes increasingly specialised. While theoretical developments are a crucial first step to better understanding these systems, their most significant impact is only felt when they are applied to the analysis of real lenses, which, in turn, is only possible when these developments are carefully implemented in the software used in such analyses. The line-of-sight effects formalism discussed in chapter 1 are no exception. Software infrastructure is best placed to incorporate new theory when it is built with this in mind, written intuitively, made open-source, and set up to enable contributions beyond those of the original developers. In order to keep pace with the rapidly evolving field, the most reliable packages are those which allow extensive flexibility in modelling choices, and can efficiently handle anything from a detailed examination of a single lens system to the batch processing of many such systems. These standards are perhaps met best by `lenstronomy` [21], to which the line-of-sight modifications to the lens equation and Fermat Potential function described in chapter 1 have therefore been added.

In section 2.1, we give a brief overview of `lenstronomy`, and describe its core modules. In section 2.2, we introduce the `LOS` subpackage to handle line-of-sight corrections, discussing the previous work of Hogg et al. [77] to implement imaging-related calculations in section 2.2.1 and section 2.2.2. In section 2.2.3, we discuss the `fermat_potential` function added to this subpackage as part of this project, and in section 2.3, we describe the methods by which this implementation is tested.

2.1 `lenstronomy`

`lenstronomy` is an open source software package developed in `python` by Simon Birrer and Adam Amara, designed primarily for strong lens modelling, which has been used in a wide range of lensing-related applications (see e.g. [117, 46, 87, 63] for some recent examples). It supports a wide range of complexity when modelling sources and lenses, allows image reconstruction with the integrated effects of an arbitrary number of mass and light profiles at shared or different redshifts, and is designed to be stable and accurate as it is scaled from the detailed analysis of individual lenses to the automated processing of several hundred lens candidates detected by future surveys. Crucially, for the analysis presented in chapter 3, `lenstronomy` allows imaging, time-delay and kinematic data to be simultaneously considered when inferring parameters and constraining models.

`lenstronomy`'s structure is modular, with each model consisting of multiple classes and sub-packages designed to carry out independent tasks. Some of the most important of these are

- `LensModel` – handles all lensing-related computation. Defines the choice of one or more lensing mass profiles, and computes deflection angles and ray tracing, magnifications, time delays, shear matrices and other related quantities. Commonly used profiles defined within this module include the point source, singular isothermal ellipse and elliptical power law profiles described in section 1.1.5, and dozens more. These profiles can be superimposed in the same plane, as is done, for example, when adding external convergence and shear to a main lens profile, or placed at different redshifts, in which case calculations are handled recursively. The former approach is handled with the `SinglePlane` class, and the latter with `MultiPlane`.

- `LightModel` – describes the surface brightness distributions of galaxies (extended sources) used to model the source and lens light. As with `LensModel`, `LightModel` faciliates the superposition of any number of profiles within the same plane or at different redshifts.

- `PointSource` – handles the computation of quantities relating to point sources (such as quasars, supernovae etc), which requires different numerics to calculations involving extended sources. Facilitates the superposition of any number of point sources. The position of point sources can either be given in the lens plane, in which case `ray_shooting` maps these positions to their shared origin in the source plane, or given in the source plane, in which case `LensEquationSolver` maps this position to one or more in the lens plane.

- `Data` – handles the implementation of quantities relating to properties of the telescope and observation (as opposed to the actual lensed image which is observed). This module is split into two classes - `Data` handles features of the imaging data such as the coordinate-to-pixel transformation and its inverse, the noise properties of the image etc, while `PSF` governs the convolution of the point spread function.

- `ImSim` – combines the `Data`, `LensModel`, `SourceModel` and `PointSource` specifications to simulate an image with `ImageModel`, either to create a mock image, or for the reconstruction of a real image, to be compared with the original. Only the `Data` and `PSF` classes are required – for example, omitting a `PointSource` class will simply result in an image without point sources, and omitting `LensModel` will result in an image of the unlensed source

- `Sampling` – handles the comparison between data and models to optimise parameters (e.g. with Particle Swarm Optimisation, PSO) and determine posterior distributions (e.g. with Markov Chain Monte Carlo methods, implemented via `emcee`). The `Param` class handles the parameters being sampled, including setting specific constraints between parameters, choosing to fix certain parameters and so on, while the `Likelihood` class manages the specific likelihood function being implemented in the sampling. Finally, with the `Sampler` class, upper bounds, lower bounds, estimated spreads and starting points of parameters are specified, and using the specified likelihood function and sampling routine, the sampling itself is carried out.

These modules are supported by several others, which perform related tasks. `Analysis` contains various classes to assist in the computation of various lensing relating quantities (such as converting a light profile to a mass profile, or a more sophisticated handling of time delay data. `Cosmo` is used to define a cosmological model when simulating a lens, and handles all cosmology-related calculations. `GalKin` models stellar kinematics to constrain the mass profile of the main lens (to alleviate the mass-sheet degeneracy). `Plots` is used for plotting data and results. `SimulationAPI` performs various supporting tasks when carrying out simulations, such as configuring data settings to match those of the corresponding telescope. `Util` contains a wide range of supporting functions, such as unit conversions, adding noise to images etc. Finally, `Workflow` allows sampling to be carried out sequentially, for iterative improvement to the optimised results, and handles the parallelization of multiple tasks.

2.2 The `LineOfSight` subpackage

In its earlier releases, `lenstronomy` was designed to handle line-of-sight effects either with a full multi-plane treatment, or alternatively via the addition of the `convergence` and `shear` profiles to the main lens plane, the equivalent of adding external convergence and shear.

To encode more physically-relevant line-of-sight effects, the `LineOfSight` subpackage was incorporated into `lenstronomy`, featuring the `LOS` and `LOS_MINIMAL` profiles. When either `LOS` or `LOS_MINIMAL` are added to the list of lens profiles, the `LensModel` module implements the `SinglePlaneLOS` class, rather than

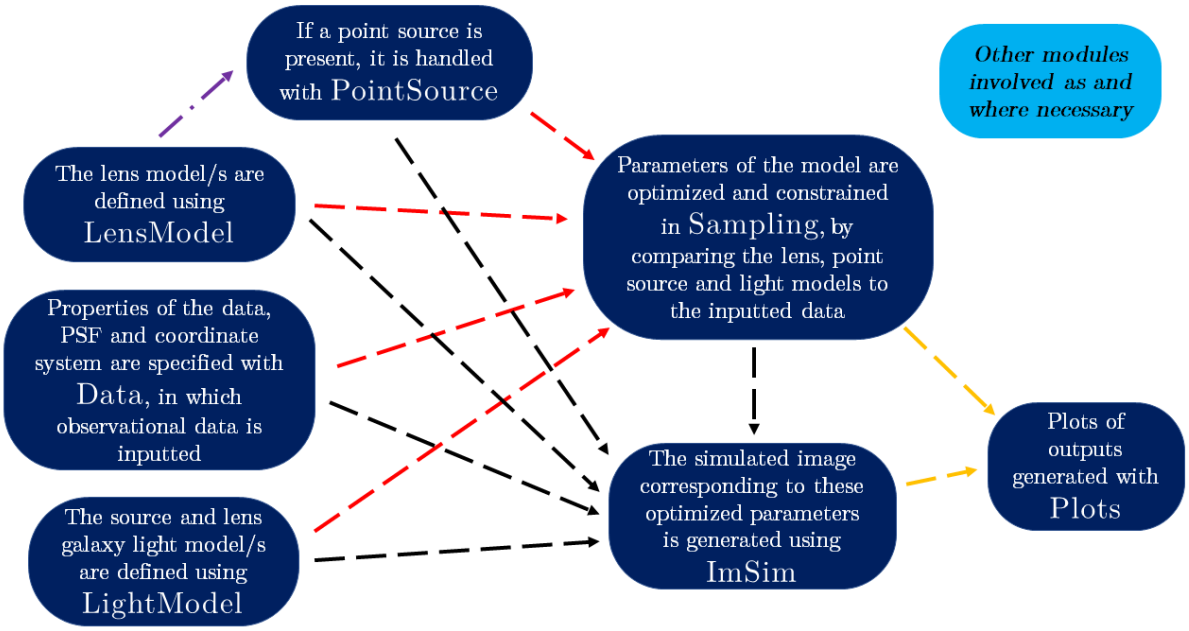

Figure 2.1: *Analysing an image with* `lenstronomy`

the `SinglePlane` class which would usually handle single plane lensing in `lenstronomy`. As these tidal line-of-sight effects are an implementation of the Dominant-Lens approximation, they are incompatible with multi-plane lensing, and so selecting `lenstronomy`'s `multiplane` option will trigger an error.

2.2.1 The `LOS` profile

The `LOS` profile is an implementation of tidal line-of-sight corrections to lensing effects in the form of the tidal matrices $\mathbf{\Gamma}_{od}$, $\mathbf{\Gamma}_{ds}$, and $\mathbf{\Gamma}_{os}$. As described in eq. (1.262), the resulting lens equation takes the form

$$\boldsymbol{\beta} = (\mathbf{1} - \mathbf{\Gamma}_{os})\boldsymbol{\theta} - (1 - \mathbf{\Gamma}_{ds})\boldsymbol{\alpha}_{ods}[(1 - \mathbf{\Gamma}_{od})\boldsymbol{\theta}]. \tag{2.1}$$

The tidal matrices $\mathbf{\Gamma}_{ab}$ are parameterised according to eq. (1.246), where the convergence κ_{ab}, complex shear $\gamma_{ab} = \gamma_1^{ab} + \mathrm{i}\gamma_2^{ab}$ and rotation ω_{ab} are the parameters inputted as keyword arguments when defining the lens model.

While the profile selection and keyword argument specifications are handled identically to that of an ordinary lens profile, `LOS` is not simply added as an additional profile on top of others specified, as would be done if `convergence` or `shear` were selected, but rather uses all lens keyword arguments and the other selected profiles to define the `SinglePlaneLOS` class. Within this class, a superposition of all other selected lens profiles is treated as the profile of the main lens ψ_{ods}. Due to the efficient use of interdependencies within `lenstronomy`, it was only necessary to explicitly a few key calculations within `SinglePlaneLOS`. All others are simply inherited from the `SinglePlane` class, as these are either independent of line-of-sight effects (such as the two or three dimensional integrated masses of the lens profile), or are calculated by calling one of the functions already modified within `SinglePlaneLOS`. For example, no changes to the `ray_shooting` function are necessary, as it calls the deflection angle function `alpha` which is modified in `SinglePlaneLOS`.

2.2.2 `LOS_MINIMAL`

If `LOS_MINIMAL` rather than `LOS` is included in the list of lens models, the minimally viable lens model described in eq. (1.265) is implemented, for which the lens equation takes the form

$$\tilde{\boldsymbol{\beta}} = (\mathbf{1} - \mathbf{\Gamma}_{\mathrm{LOS}})\boldsymbol{\theta} - (\mathbf{1} - \mathbf{\Gamma}_{od})\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \mathbf{\Gamma}_{od})\boldsymbol{\theta}]. \tag{2.2}$$

Thus `LOS_MINIMAL` reduces the number of parameter inputs to two convergences κ_{LOS} and κ_{od}, two complex shears γ_{LOS} and γ_{od}, and two rotations ω_{LOS} and ω_{od}. When `LOS_MINIMAL` is selected in the

list of lens profiles and these parameters specified, inputted parameters of $\mathbf{\Gamma}_{\rm LOS}$ are simply set equal to those of $\mathbf{\Gamma}_{os}$, while those of $\mathbf{\Gamma}_{od}$ were set equal to those of both $\mathbf{\Gamma}_{ds}$ and $\mathbf{\Gamma}_{od}$. From this point onwards, the subpackage will simply implement the `LOS` profile with these tidal matrices.

2.2.3 The `fermat_potential` function

Earlier versions of `SinglePlaneLOS`, written by Natalie Hogg and Pierre Fleury, implemented changes to the calculation of the deflection angle with `alpha` and the Hessian matrix of second derivatives with `hessian`. These enabled the calculation of all imaging-related quantities, but not the calculation of time delays. As part of the work of this project, the `fermat_potential` function was added to `SinglePlaneLOS`, which implements line-of-sight corrections to the Fermat potential of the main lens.

The Fermat potential function implemented is that given in eq. (1.345) as

$$\phi(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta}')\cdot\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\beta}') - \psi_{ods}[(\mathbf{1}-\mathbf{\Gamma}_{od})\boldsymbol{\theta}], \tag{2.3}$$

where, as before,

$$\boldsymbol{\mathcal{A}}_{\rm LOS} \equiv \boldsymbol{\mathcal{A}}_{od}^{\intercal}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os}, \tag{2.4}$$

$$\boldsymbol{\beta}' = \boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta}, \tag{2.5}$$

with $\boldsymbol{\beta}$ referring to the position of the source, and $\boldsymbol{\beta}'$ to the position of the source as it would be seen in the presence of weak perturbers only. As is the case elsewhere in `SinglePlaneLOS`, the main lens potential ψ_{ods} is simply determined by the other specified lens profile/s, but evaluated at the modified position $(\mathbf{1}-\mathbf{\Gamma}_{od})\boldsymbol{\theta}$.

`fermat_potential` takes a pair of image coordinates (`x_d` and `y_d`), the keyword arguments of the lens model, and (optionally) a pair of source coordinates (`x_source` and `y_source`) as inputs, and returns the Fermat potential as follows

1. the lens model keyword arguments are split into those of the main lens (`kwargs_main`) and those of the `LOS` profile (`kwargs_los`)
2. the amplification matrices $\boldsymbol{\mathcal{A}}_{od}$, $\boldsymbol{\mathcal{A}}_{os}$ and $\boldsymbol{\mathcal{A}}_{ds}$ are calculated from `kwargs_los`, according to eq. (1.75)
3. $\boldsymbol{\mathcal{A}}_{od}^{\intercal}$, $\boldsymbol{\mathcal{A}}_{os}^{-1}$ and $\boldsymbol{\mathcal{A}}_{os}^{-1}$ are calculated from these matrices, and then $\boldsymbol{\mathcal{A}}_{\rm LOS}$ is calculated as $\boldsymbol{\mathcal{A}}_{\rm LOS} = \boldsymbol{\mathcal{A}}_{od}^{\intercal}\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os}$
4. the modified image position $(\mathbf{1}-\mathbf{\Gamma}_{od})\boldsymbol{\theta}$ is calculated
5. $\psi_{\rm eff}(\boldsymbol{\theta})$ is calculated by inputting `kwargs_main`, `x_d` and `y_d` into the expression for the potential of the main lens profile
6. if the source position $\boldsymbol{\beta}$ was left unspecified, it is calculated using the `ray_shooting` function
7. $\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta}$ is calculated
8. $\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta})$ is calculated
9. the geometric term is calculated as $\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta})\boldsymbol{\mathcal{A}}_{\rm LOS}(\boldsymbol{\theta}-\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta})$
10. the Fermat potential is returned as this geometric term minus $\psi_{\rm eff}(\boldsymbol{\theta})$

While `fermat_potential` can be called directly to return a numerical value of the Fermat potential at a specified image position, the function's primary use is in the calculation of time delay and the calculation of time delay distances. As τ_{ds} involves only the choices of cosmology and main lens and source redshifts, no further changes to `lenstronomy` were required. The `arrival_time` function in the `lens_model` class simply inherits this Fermat potential, as do the `fermat_potential` and `time_delays` functions within the `TDCosmography` class, thus ensuring that line-of-sight effects are recognised and implemented wherever time delays are calculated or used in `lenstronomy`.

2.3 Testing `SinglePlaneLOS`

To ensure that the `SinglePlaneLOS` class was working as it should, a Jupyter notebook was written by Hogg and Fleury in which the outputs of functions within `SinglePlaneLOS` could be compared to those of the equivalent multi-plane setting. Supplementing the main lens with three tidal matrices $\mathbf{\Gamma}_{os}$, $\mathbf{\Gamma}_{od}$ and $\mathbf{\Gamma}_{ds}$ in the main lens plane, which is the configuration implemented by `SinglePlaneLOS`, is exactly equivalent to placing three shear planes $\mathbf{\Gamma}_f$, $\mathbf{\Gamma}_d$ and $\mathbf{\Gamma}_b$ between the observer and dominant lens, in the dominant lens plane, and between the dominant lens and the source respectively, at arbitrary redshifts $z_f < z_d < z_b$.

In this multi-plane setup, non-linear couplings between the planes introduce higher order convergence and rotation, which must be accounted for when comparing the `multiplane` and `SinglePlaneLOS` settings. To convert from the `multiplane` shear planes to line-of-sight equivalents, we have

$$\mathbf{\Gamma}_{od} = \mathbf{\Gamma}_{ofd} = \frac{D_{fd}D_{os}}{D_{od}D_{fs}}\mathbf{\Gamma}_f, \tag{2.6}$$

$$\mathbf{\Gamma}_{ds} = \mathbf{\Gamma}_{dbs} = \frac{D_{os}D_{db}}{D_{ob}D_{ds}}\mathbf{\Gamma}_b, \tag{2.7}$$

and

$$\mathbf{\Gamma}_{os} = \mathbf{\Gamma}_f + \mathbf{\Gamma}_d + \mathbf{\Gamma}_b - [\mathbf{\Gamma}_d\mathbf{\Gamma}_{od} + \mathbf{\Gamma}_b\mathbf{\Gamma}_{odb}(\mathbf{1} - \mathbf{\Gamma}_{od})], \tag{2.8}$$

where

$$\mathbf{\Gamma}_{ofb} = \frac{D_{os}D_{fb}}{D_{ob}D_{fs}}\mathbf{\Gamma}_f, \tag{2.9}$$

$$\mathbf{\Gamma}_{odb} = \frac{D_{os}D_{db}}{D_{ob}D_{ds}}\mathbf{\Gamma}_d. \tag{2.10}$$

Initial tests confirmed that, with arbitrary choices of redshifts and shears, the `alpha`, `ray_shooting` and `hessian` functions returned identical values to within 10^{-16}, 10^{-16} and 10^{-8} precision respectively.

Now, given that the Fermat potential function implemented in `lenstronomy` was derived by ignoring terms which were constant for a given source position $\boldsymbol{\beta}$ (and therefore unobservable), we would not expect arbitrary coordinates to return the same Fermat potential in the `multiplane` and `SinglePlaneLOS` cases, and these should only be compared in the context of relative time delays between multiple images of the same source. We therefore instead substitute arbitrary source plane coordinates into the `LensEquationSolver` function, and use the outputted solutions to the lens equation in the image plane as our image coordinates. When these comparisons are made, `multiplane` and `SinglePlaneLOS` agree within 10^{-11} precision.

A final, necessary point of comparison is to test whether the `LOS` and `LOS_MINIMAL` profiles agree when $\mathbf{\Gamma}_{\rm LOS} = \mathbf{\Gamma}_{os}$ and $\mathbf{\Gamma}_{ds} = \mathbf{\Gamma}_{od}$. In each of the tested cases, `ray_shooting`, `hessian` and `fermat_potential`, the agreement was exact. The `LOS` subpackage was therefore added to the most recent releases of `lenstronomy`, passing all necessary unit tests.

Chapter 3

Measuring the Line-of-Sight Shear

Having laid the theoretical foundations of line-of-sight effects on strong lens images and time delays, arriving at a minimal mathematical model for these effects, and having completed the implementation of these models into a competitive strong lensing software package, the stage is perfectly set to explore the measurability of these effects. Several key questions will dictate our approach to this, as we look to fill gaps in the existing literature, but the goal at the heart of these and this chapter will be to gain insight into the kinds of sources, lens systems, data parameters and approaches to fitting which best enable accurate and precise measurements of the line-of-sight shear. We begin in section 3.1 by introducing the use of mock datasets to test constraints on parameters of interest. In section 3.2, we generate a simplified but representative sample of strong lens images and their associated time delays, and then extract and compare measurements of the line-of-sight shears from these. Measurements of the foreground shear and ellipticity made from the same dataset are discussed in appendix B. In section 3.3, we turn our attention to the specific effect that the image noise and time delay uncertainties have on these measurements. In section 3.4, we investigate the strength of the degeneracy presented in section 1.3.4 and examine the role of the post-Born correction to the main lens in shear measurements. Finally, in section 3.5, we discuss the prospects for measuring the line-of-sight shear from different types of lens systems, and contextualise these results within the findings presented elsewhere in the literature.

3.1 MCMC Fitting and mock data

In its design to reconstruct the source light, lens light and lensing potential responsible for producing an observed strong lensing system, and comparing the resulting model image with the original, `lenstronomy` also very effectively facilitates the construction of mock imaging data. An immediate benefit of this process is that the effects of varying specific parameters or model choices on resulting images can immediately be discerned. Most importantly for our purposes, however, is that these mock images can in turn be analysed and the inputted profiles and parameters reconstructed. As these inputs are known, this allows for a straightforward and precise comparison between factors such as the quality of data, parameter values and profile choices on the ability of `lenstronomy` to accurately and precisely constrain key parameters.

In typical studies of strong lensing systems, parameterised models of the source, lens and line-of-sight are chosen, and the parameter values which best fit the given data are found using an optimiser. In `lenstronomy`, this is typically done with a particle swarm optimisation (PSO) routine. The posterior distributions around these parameters are then sampled in order to report uncertainties on these parameters, typically using a Markov chain Monte Carlo (MCMC) sampler. This process is often carried out several times, with progressively more complex models being employed as earlier, simpler iterations provide better starting points for these fits.

The specific question we are seeking to answer in the analysis which follows is whether the presence of point source images and time delay data in observations of strong lensing systems can improve constraints on the line-of-sight shear, in comparison to the case where only a lensed extended source is present, a task for which the MCMC sampler is well-suited. Lensing degeneracies manifest in potentially complex covariances between parameters, and to obtain robust estimates for the posterior distributions (which characterise the probabilities of the model parameter taking on different values), and thus to properly understand how well a certain parameter of a given model can be constrained from a dataset, an MCMC implementation rises head and shoulders above alternative approaches [76].

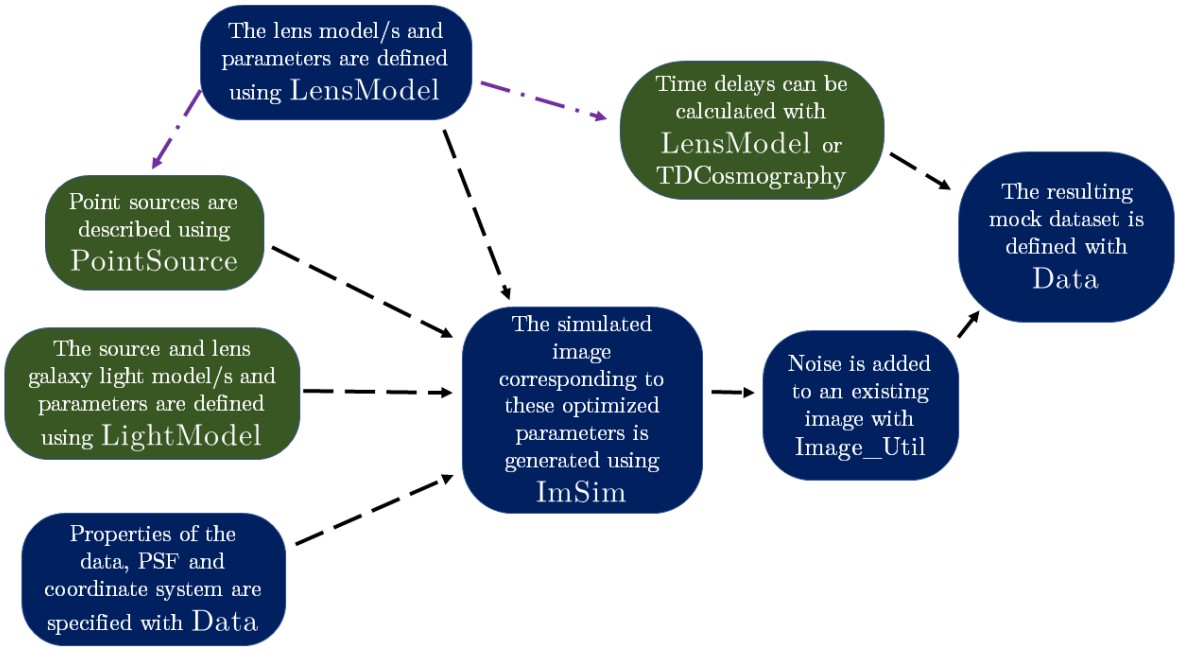

Figure 3.1: *Generating mock data with* `lenstronomy`. *Steps shown in blue are always carried out, while those in green are taken only when appropriate for the specific dataset in question.*

3.2 Measuring shear in a statistically representative sample of lenses

3.2.1 Generating the mock data

Early tests confirmed a significant degree of variability in the relative accuracy and precision of measurements of the line-of-sight shear from mock images. To comment effectively on the utility of different lensing configurations and datasets when trying to measure this shear, it is therefore necessary to consider a wide range of lens, source, and data parameters, and make statistical comparisons between results from different classes of images. For this comparison to be relevant to modern astronomy, the spread in parameters should approximate the real distributions expected in lensing surveys wherever possible.

With this goal in mind, 195 sets of parameterised lens models were generated to be analysed and compared. Each set contained an image featuring a lensed extended source only, a second featuring a lensed point source only, and a third featuring a lensed point and extended source. Within a set of three such images, all other parameters (relating to the position of the source, the cosmology, the data, the gravitational profile of the main lens and the line-of-sight effects) were kept exactly the same, and the third image was generated using parameters for the extended and point sources identical to those used when generating the first two images.

The resulting data-set was intended to represent a sample of strong gravitational lensing images which, while simplified, is nonetheless representative of the typical diversity in observational data. By simulating sets of identical lenses, with images within a set differing only in the class of source present, it was possible to make direct comparisons between point source and extended source images as candidates for measurements of the external shear, and images featuring point sources with or without time delay data.

Cosmology

For all images, a flat ΛCDM cosmology with no baryons was chosen and implemented using `astropy`, with the Hubble constant H_0 set to 70 km.s^{-1}.Mpc^{-1}, and Ω_{m0} (the density of non-relativistic matter in units of the critical density at $z = 0$) set to 0.3. With these choices, $\Omega_{\Lambda 0}$ is calculated as $1 - \Omega_{m0} = \Omega_{\Lambda 0}$.

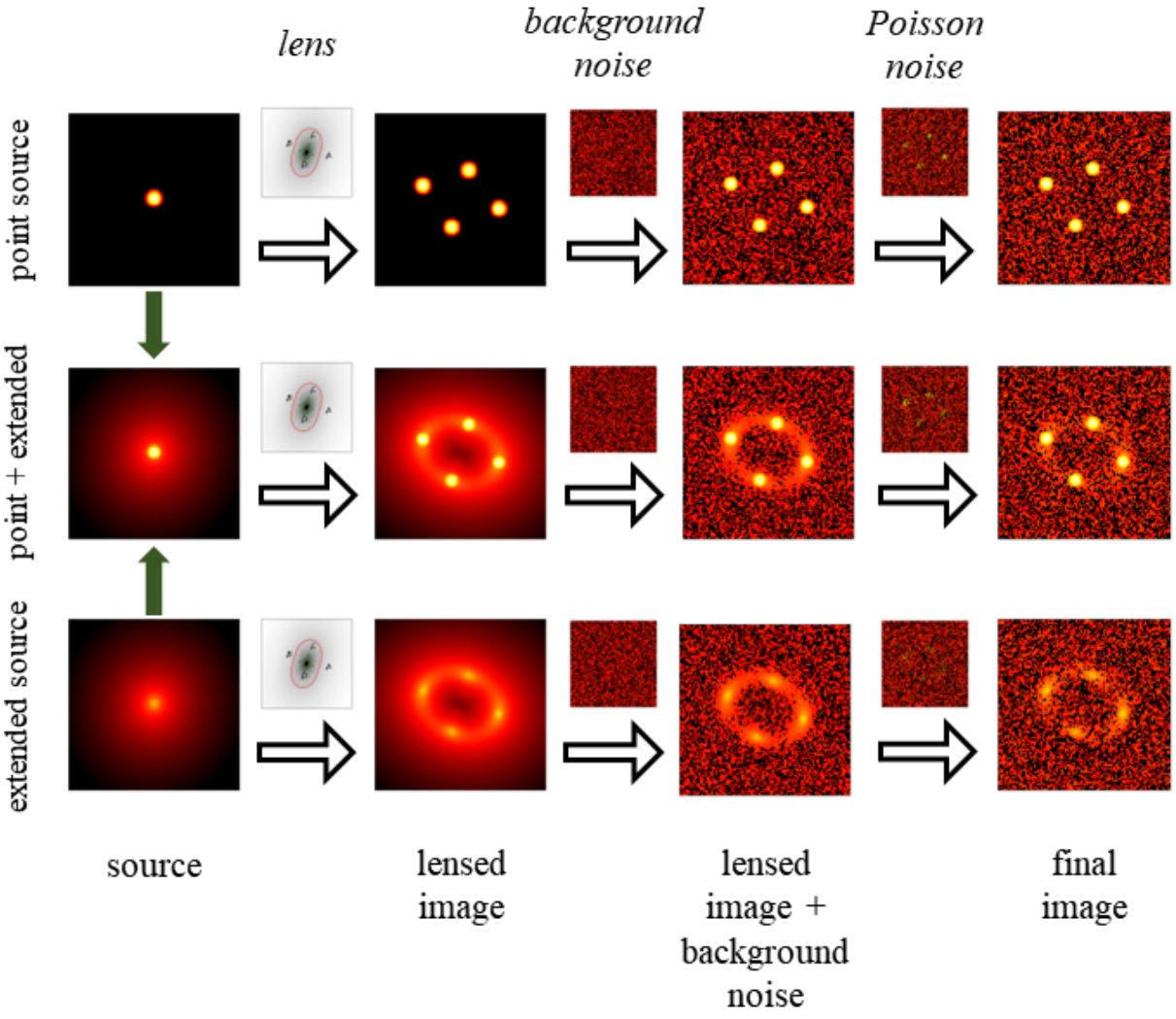

Figure 3.2: *The effects of lensing and adding noise with* `lenstronomy`. *The generation of background noise depends only on the dimensions of the image, while Poisson noise is sensitive to the counts in each pixel of the penultimate image, including the contribution of the background noise. The Poisson noise must therefore be generated after the background noise has already been added.*

Source parameters

The light from each extended source galaxy is modelled with a Sérsic ellipse profile, frequently used in both time-delay cosmography (e.g. [148, 22]) and in lensing simulations (e.g. [31]), for which the intensity at R is calculated according to

$$I(R) = I_0 \mathrm{e}^{-b_n\left(\frac{R}{R_{\mathrm{Sérsic}}}\right)^{\frac{1}{n}}}, \tag{3.1}$$

where I_0 is the surface brightness ("amplitude") at the half-light radius (the radius within which half of the total integrated surface brightness of the galaxy is contained), and $b_n \approx 1.99n - 0.327$. In `lenstronomy`, the ellipticity components e_1 and e_2 are typically given as arguments, rather than the axis ratio f. The parameter R is calculated from the inputted values of θ_1, θ_2, e_1 and e_2 by firstly converting e_1 and e_2 into an axis ratio q and an orientation ϕ_G via

$$\phi_G = \frac{1}{2}\arctan\left(\frac{e_2}{e_1}\right), \tag{3.2}$$

$$f = \frac{1-c}{1+c}, \tag{3.3}$$

where $c = \sqrt{e_1^2 + e_2^2}$. The modified position $\boldsymbol{x}$ is then calculated by a multiplication with the rotation matrix, according to

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos(\phi_G) & \sin(\phi_G) \\ -\sin(\phi_G) & \cos(\phi_G) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}. \tag{3.4}$$

R is then calculated according to

$$R = \sqrt{f^2 x_1^2 + x_2^2}. \tag{3.5}$$

The values of e_1 and e_2 were sampled at random from a uniform distribution bounded by -0.1 and 0.1.

In `lenstronomy`, flux counts are measured in arbitrary units, and thus when generating an image, it is only the relative contributions of different sources of flux (such as extended sources, point sources, background sky noise and Poisson noise) which are relevant, and a proportional scaling of these contributions has no effect on the resulting image. It is therefore convenient to standardise a parameter governing the flux contribution from one of these sources, and scale all the others according to this standard. We chose to set I_0 to a constant arbitrary value of 1000 for each extended source galaxy.

The Sérsic index n was chosen by sampling from a Gaussian distribution centered on 4, with a standard deviation 2. Values higher than 10 or lower than 0.5 are rejected, to avoid unrealistically atypical profiles [32, 180]. The resulting distribution of n is shown in fig. 3.7.

Now, assuming that potential source galaxies are uniformly distributed in comoving space, the number of galaxies within a solid angle Ω in the sky between comoving distance χ and $\chi + \mathrm{d}\chi$ will be proportional to the corresponding comoving volume, $\mathrm{d}V_c = \Omega\chi^2\mathrm{d}\chi$, as illustrated in fig. 3.3. Therefore, the probability that a source lies between χ and $\chi + \mathrm{d}\chi$ is distributed according to

$$p(\chi)\mathrm{d}\chi = \frac{\mathrm{d}V_\mathrm{c}}{V_\mathrm{tot}} = \frac{3\chi^2\mathrm{d}\chi}{\chi^3_\mathrm{max}} = \mathrm{d}\left(\frac{\chi^3}{\chi^3_\mathrm{max}}\right), \tag{3.6}$$

where the normalisation has been made in terms of the total volume $V_\mathrm{tot} = \Omega/4\pi \times (4/3)\pi\chi^3_\mathrm{max}$ in the solid angle Ω out to the maximum considered comoving distance, χ_max. This maximum is chosen to be $\chi_\mathrm{max} = \chi(z_\mathrm{max})$, with $z_\mathrm{max} = 5$, following the results of [39].

Now, the PDF of the quantity $u = \chi^3/\chi^3_\mathrm{max}$ is simply

$$p(u) = p(\chi)\frac{\mathrm{d}\chi}{\mathrm{d}u} = 1, \tag{3.7}$$

for $u \in [0, 1]$. Thus, to obtain the comoving distance to the source $\chi(z_s)$, we simply sample u from a uniform distribution in $[0, 1]$, from which we calculate

$$\chi(z_s) = u^{\frac{1}{3}}\chi(z_s^\mathrm{max}). \tag{3.8}$$

The redshift of the source is then chosen to be the redshift z_s at this comoving distance. If the resulting value of z_s is smaller than 0.1, a new value of u is chosen and the calculation repeated, to avoid lens-source combinations at unrealistically small redshifts.

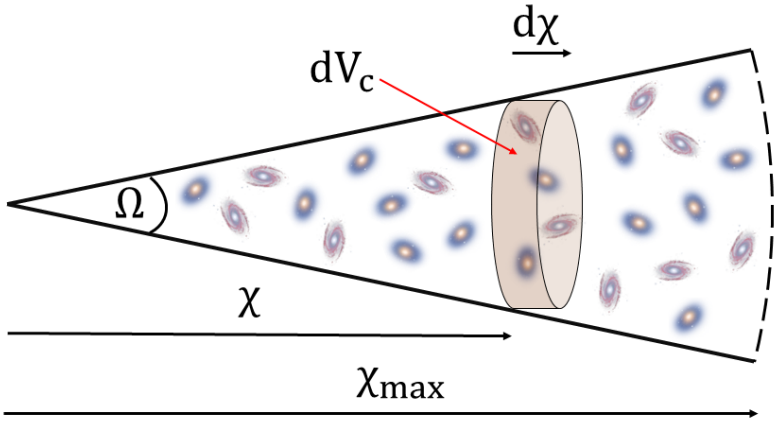

Figure 3.3: *The geometry of randomly sampling a galaxies within a solid angle* Ω*. Assuming that the density of galaxies in a certain volume of comoving space is independent of the redshift* z*, the number of galaxies at* z *and thus a comoving distance* $\chi(z)$ *is proportional to the volume of an infinitesimal comoving cylinder at that redshift and within that solid angle,* $\mathrm{d}V_\mathrm{c} = \Omega/4\pi \times 4\pi\chi^2\mathrm{d}\chi$*.*

It has been shown that, for typical galaxies, $R_{\mathrm{Sérsic}}$ is not independent of the redshift, and galaxies at higher redshift are smaller on average than their low-redshift counterparts. In [70], $R_{\mathrm{Sérsic}}$ was found to scale in inverse proportion to the Hubble parameter at that redshift for redshifts between 3 and 6, with a mean radius of 1.5 kpc at $z = 4$, with a standard deviation of the order 0.2. [139] found a similar result for galaxies between $z \sim 0.2$ and $z \sim 2$, with the average value of $R_{\mathrm{Sérsic}}$ reaching 6 kpc for star-forming galaxies and 4 kpc for quiescent galaxies at $z \sim 0.2$, and argue that this relationship arises from mergers causing galaxies to grow with time.

To approximate these results for our sample, the source radius is chosen by firstly determining a mean source radius at that redshift via the redshift and Hubble parameter, according to

$$\mu(R_{\mathrm{Sérsic}}) = 1500 \times \frac{H(4)}{H(z_s)}\mathrm{pc}, \tag{3.9}$$

and then sampling the source radius from a Gaussian distribution centered on $\mu(R_{\mathrm{Sérsic}})$, with standard deviation 30 % of this value. Values higher than 7000 pc or lower than 500 pc are rejected and the sampling repeated. The final Sérsic radius is found by converting this sampled value into units of arcseconds via the angular diameter distance D_{os} at z_s. This ensures a reasonable variability in galaxy sizes at a given redshift, but nonetheless respects the trends found in [70, 139]. The resulting distribution of $R_{\mathrm{Sérsic}}$ values with z_s is shown in fig. 3.7.

While the volume-limited approach to generating this dataset is effective in generating a realistic distribution of source brightnesses at different redshifts, it does not fully take into account the selection effects present in a more realistic flux-limited survey. Flux limits are particularly relevant in strong-lens surveys because strongly-lensed images are magnified, leading to the so-called magnification bias, whereby the chances of observing a lens is higher than would be expected from the brightness of its source. Thus, in real strong-lens surveys, quadruply imaged sources (for which the total magnification is higher) are overrepresented, and sources far fainter than would normally be observed can still be imaged if strongly lensed. While these effects can change the proportion of different types of lenses within a sample, they do not change the conclusions drawn from specific types of lens images which are presented later in this work, and the volume-limited approach is adopted as a reasonable simplification.

Main lens parameters

For the gravitational field of the lensing galaxy, elliptical power law (EPL) profiles were chosen, as described in section 1.1.5. While EPLs are very effective at reproducing imaging observables, and are used in several lens reconstructions (e.g. [148, 22, 27]) they are increasingly thought to be oversimplifications of the underlying mass distribution (see e.g. [31]). Nonetheless, they serve as an adequate starting point for the specific questions motivating this work, and are thus used (as in [83] and sections of [77]). The ellipticity components of the profile, e_1 and e_2, were each chosen by sampling a uniform distribution

bounded by -0.1 and 0.1. For the power law index $\gamma_{\rm PL}$, a uniform distribution between 1.5 and 2.5 was chosen, producing a spread of $\gamma_{\rm PL}$ values typical for observed lenses (e.g. [2]).

As discussed in section 1.1.5, the value of $\theta_{\rm E}$ in an EPL lens acts as a characteristic length scale for the strength of the deflection caused by the lens for a light beam reaching the lens at different angular positions. For general EPL lenses, the relationship between $\theta_{\rm E}$ and intrinsic properties of the lensing galaxies is complicated by the presence of $\gamma_{\rm PL}$ and e in the lens parameterisation. However, for $\gamma_{\rm PL} \sim 2$ and $e \ll 1$, certain key relationships are well approximated by the SIS model discussed in section 1.1.5.

To produce a statistically representative distribution of Einstein radii, we sample a value of σ_v (representing the velocity distribution of the main lens) from a Gaussian distribution with mean 210 km.s^{-1} and standard deviation 54 km.s^{-1}, rejecting values higher than 400 km.s^{-1} or lower than 100 km.s^{-1}, which approximates the spread of σ_v values seen in typical galaxies capable of producing observable lensing events [39].

z_d is calculated almost identically to that of the source, except that $z_{\rm max}$ is set to equal z_s, and u is sampled from a uniform distribution between 0.01 and 0.99, to avoid placing the lensing galaxy too close to the observer or source.

With z_d, z_s and the cosmology parameters defined above, the angular diameter distances D_{ds} and D_s are calculated, and then the Einstein radius of the lens is calculated according to eq. (1.143), with the value of σ_v determined above. If the resulting value is less than 0.3 arcseconds, it is rejected, the velocity dispersion and lens redshift are recalculated, and a new value of $\theta_{\rm E}$ determined. If the resulting value of $\theta_{\rm E}$ is sufficiently large, $\theta_{\rm E}$, $e = (e_1, e_2)$ and $\gamma_{\rm PL}$ are used to define the EPL profile of the main lens. The resulting distributions of $\theta_{\rm E}$, z_s and z_d are shown in fig. 3.7. This distribution of redshifts is very similar to that predicted in [39], and is consistent with the typical deflector and source redshifts of surveys such as SLACS [26], COSMOS [140], CLASS [116] and SQLS [81] (see also the review in [160]).

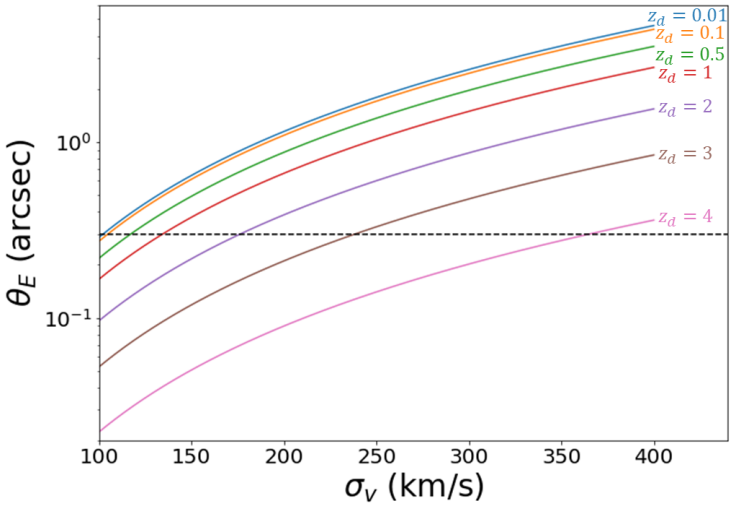

Figure 3.4: *The relationship between the velocity dispersion σ_v and the Einstein radius $\theta_{\rm E}$ for an SIS lens with $z_s = 5$, for different values of the deflector redshift z_d. The threshold value of $\theta_{\rm E} = 0.3$ is plotted as a horizontal black line.*

The dependency of $\theta_{\rm E}$ on z_d and σ_v is illustrated in fig. 3.4. From this, it is clear that the lowest permitted values of σ_v are only sufficient to meet the 0.3 arcsecond threshold for $\theta_{\rm E}$ if the main deflector is located very close to the observer, and that, for deflectors at very high redshifts, a very large velocity dispersion is needed to result in a sufficiently large $\theta_{\rm E}$.

Although `lenstronomy` faciliates the inclusion of the light profile of the lens itself, this was ignored for the purposes of this dataset. The handling of lens light when analysing real lensing images is done in several ways. In a typical study, Einstein rings and/or quasar images are often masked out by eye, leaving only the main lens light. This light is then fitted (e.g. with a Sérsic profile), and the reconstructed lens light model is subtracted from the image. The residuals are then reconstructed using a lens mass model and either a parametric or pixelised source (e.g. [48]). In other cases, in particular when z_s is large, the lens and source light might have sufficiently separated peak wavelengths to allow them to be almost entirely separated with multi-wavelength observations (e.g. [5]). While adding somewhat to the uncertainty budget, the uncertainty inherent in lens light subtraction is generally dwarfed by the challenge of reconstructing the lens mass distribution and source light, and is thus omitted for computational simplicity, as is done in e.g. [109, 31, 23] and parts of [77]. This is practically equivalent to assuming that the lens light has already been modelled and subtracted, or alternatively that Einstein rings or quasar images are observed in a wavelength band in which the lens light contribution is well below the background noise level.

It is also worth noting, that, given the significant contribution of the dark matter halo in typical lenses, there is not generally a strong relationship between the visible component of a lens and its mass

distribution, meaning that there is little to gain from imposing priors on the lensing potential based on its light profile (see [29] and references therein).

Line-of-sight effects

The full lens model consists of the EPL profile described above, supplemented with line-of-sight effects. From the discussion in section 1.3.4, the convergences κ_{os}, κ_{ds} and κ_{od} cannot be determined from lensing observables, and are therefore simply set to 0, along with the rotations ω_{os}, ω_{ds} and ω_{od}. These rotations are the result of couplings between successive lenses in a multi-plane setup, as consecutive shears act along non-aligned axes, but are generally $\mathcal{O}(\gamma^2)$ [77]. For the shears, while it is possible to calculate expected values γ_{os}, γ_{ds} values based on the comoving distances to the source (χ_s) and the lens (χ_d), this calculation is also highly dependent on the density contrast, which in turn can vary significantly along different lines of sight. Each of the 6 line-of-sight shear components is therefore simply chosen from a random uniform distribution between -0.1 and 0.1. $\gamma_{\rm LOS}$ is then calculated as

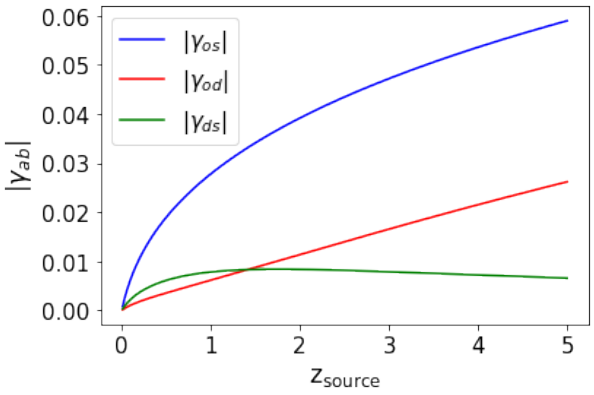

Figure 3.5: *The expected magnitudes of γ_{os}, γ_{od} and γ_{ds} for a lens located at redshift $z_d = 0.4z_s$, calculated according to eqs. (3.11) and (3.12).*

$$\gamma_{\rm LOS} = \gamma_{os} + \gamma_{od} - \gamma_{ds}. \tag{3.10}$$

The resulting spread of $\gamma_{\rm LOS}$ values is shown in fig. 3.7. This spread is larger than that inputted in [77], and commonly-cited estimates of shear in typical lens environments (such as in [69, 12, 41]), or the results of certain weak lensing studies (such as [11, 165, 173]). However, weak lensing surveys are not directly applicable to the shear strength in strong gravitational lensing, as lensing galaxies (typically early-type) are more likely to be found in over-dense regions of space [88], in which the typical shear is higher, and [78] argues that shear values larger than 0.1 would not be unlikely for quadruple-imaging systems. [174] further argue that the aforementioned estimates of shear in lens environments fail to account for the range of these environments, and find line-of-sight and environmental shears as high as 0.17 for observed lensing systems. Cosmological simulations demonstrate that the lensing efficiency of clusters can be significantly increased by line-of-sight structure [14, 110, 123], which would once again suggest that observed strong lenses are affected by above-average external shears. This is further supported by observations in [17], which revealed a pronounced overdensity of galaxies along the line-of-sight to strong lensing galaxy clusters. [101], combining weak and strong lensing measurements of shear, find external shears as high as 0.19.

To better understand the relationship between the shears used in this mock data and typical values of the cosmic shear, we make use of the result published in [28], in which it was found that the standard deviation of the κ_{os} for a *Planck*-2018 cosmology [6] was well-fitted by a function of the form

$$\sqrt{\langle \kappa_{os}^2 \rangle} = 0.0218 \left[(1 + 12.6)^{0.315} - 1 \right]. \tag{3.11}$$

The variances of the weak lensing convergence and shear are identical, and so eq. (3.11) can be used to calculate the standard deviation of $|\gamma_{os}|$, the quadratic mean of γ_{os} terms, which in turn can be used as a rough indicator of $|\gamma_{os}|$ along a typical line-of-sight. Then, under the *mean-field approximation*, in which the density contrast is taken to be proportional to the scale factor at the same comoving distance, γ_{os} can be related to γ_{od} and γ_{ds} via

$$\gamma_{ab} \approx \left(\frac{\chi_b - \chi_a}{\chi_s} \right)^2 \gamma_{os}, \tag{3.12}$$

where, as before, χ_i refers to the comoving distance at z_i [28]. Using eqs. (3.11) and (3.12), the expected values of the shear magnitudes are plotted in fig. 3.5 for a lens placed at $z_d = 0.4z_s$, as an illustration of the expected shear strengths for a typical line-of-sight, and fig. 3.6 compares the actual shear values used in generating the mock data with the expected values from eqs. (3.11) and (3.12). In several cases, the mock shear values are a few times larger than the expected values along random lines of sight, but, as discussed above, this is not unreasonable.

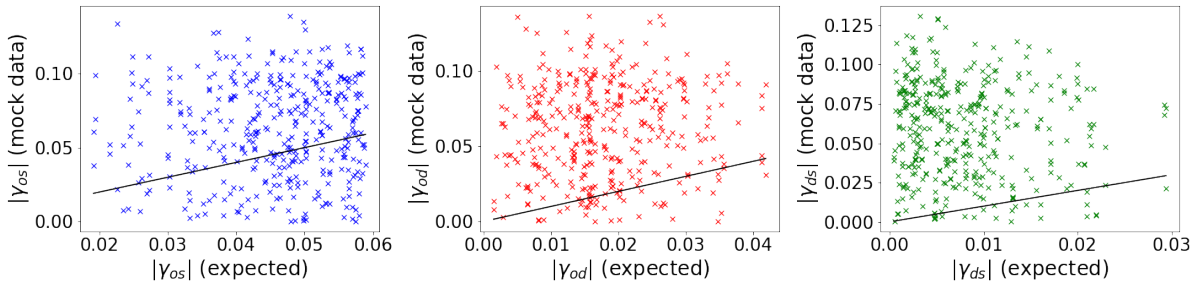

Figure 3.6: *The magnitudes of γ_{os}, γ_{od}, and γ_{ds} used in generating the mock data, plotted against the expected values of these shears calculated according to eqs. (3.11) and (3.12) at the relevant lens and source redshifts. Shown as black lines on each plot is the line along which the mock and expected data would be equal.*

While γ_{od}, γ_{ds} and γ_{os} were sampled from uniform distributions between -0.1 and 0.1, $\gamma_{\rm LOS}$ values can be seen from fig. 3.7 to be on the order of a few percent or less for most lenses, but reaching as high as $\sim 0.2 - 0.25$ in a few cases. This choice ensures that the mock dataset remains at least approximately representative of typical shear strengths, while still including some strong shear examples, the effects of which are of particular interest.

The source position and point source images

Once the EPL and LOS parameters had been set, lensing effects were modelled with both these profiles (see section 2.2), and the position of the source was then chosen. While the alignment of galaxies with large differences in redshift is predominantly random, we chose to sample the source position coordinates β_1 and β_2 from a uniform distribution out to 20 % of the $\theta_{\rm E}$ value for the lens in question. In general, sources and lenses which are better aligned will produce more complete Einstein rings or more multiply-imaged point sources, and this choice therefore imposes a selection bias towards better images, which are more likely to be chosen for measurements the external shear.

With this choice, `lenstronomy`'s `lensEquationSolver` function is used to determine the image positions corresponding to this source position. If multiple solutions are not found, and thus the choice of position and lens parameters does not result in strong lensing, then β_1 and β_2 are resampled and the calculation repeated. If, after 10 repeats, this alignment still fails to give rise to strong lensing, the model is rejected.

For the remaining lens configurations, the centre of the extended source is chosen to be these coordinates β_1 and β_2, and the magnification at the corresponding image positions is then calculated for each of the 2, 3 or 4 images. The actual observed magnification which is used when generating images is chosen by sampling from a normal distribution centered on the actual magnification, with standard deviation of 0.01, to account for the perturbative effects of phenomena such as microlensing.

In `lenstronomy`, point sources are handled separately to extended sources, and to simulate the effects of compact source such as quasars or supernovae, a point source is typically superimposed on an extended light distribution to model the host galaxy (e.g. [148, 176]), which is the approach adopted here. To ensure that the comparison between the information contained in point and extended source images is as direct as possible, the point source is placed at the centre of the extended source. The amplitude of this point source, measured in the same units as that of I_0 from eq. (3.1), is calculated by multiplying the total integrated flux of the extended Sérsic profile by a factor which is sampled from a uniform distribution between 50 and 500. The amplitude of the point source images is then set by multiplying this point source amplitude by the perturbed magnification at each of the image positions.

In the resulting dataset of 195 lens configurations, 127 featured doubly-lensed point sources, 7 that were triply-lensed, and 61 (around 22 %) that were quadruply-lensed. These proportions are not expected to reflect the real distribution of observed strong lensing events, given the restrictions on the ranges from which the source positions were sampled. However, it remains of interest to compare these to real data, as a means of identifying potential biases in the analysis. Around 25 % of strongly-lensed quasars are quadruply-imaged [106], a number to which the 31 % in our dataset is reasonably close. While [54] predict that triply-imaged quasar systems will be much rarer, this prediction was made using a non-elliptical main lens model plus external shear, a simplification which likely explains this difference. Nonetheless, while observed [88], triply-imaged systems are slightly rarer in real samples than in our mock dataset.

Data parameters and image generation

Having defined a cosmology, lens model (consisting of an EPL profile plus line-of-sight effects), and source light profile (either a Sérsic ellipse, a point source or a combination of the two), generating sets of mock images required only that data and measurement parameters be set. A Gaussian point spread function (PSF) was used, with the FWHM set to double the pixel size, which in turn was set to 0.05 arcseconds. The number of pixels along the width of the square pixel grid was chosen to be

$$n_{\text{pix}} = 4 \times \frac{\text{FWHM} + \theta_{\text{E}}}{\text{pixel size}} = 8(1 + 10\theta_{\text{E}}). \tag{3.13}$$

This equates to cropping an n_{pix} by n_{pix} region of a full WFC UVIS Hubble image centred on the centre of the lens. We found that the multiplicative factor of 4 ensured the minimum field of view which contained the entirety of the relevant image, thus maximising computational efficiency.

Having fixed $I_0 = 1000$ for every Sérsic source profile, the root mean squared background noise RMS_{bg} used when generating the image must be varied to take into account both the natural variability in source galaxy brightnesses and the loss in signal strength with increasing luminosity distances at larger redshifts. It should again be noted that fixing I_0 and varying the rms of the background noise is exactly equivalent to the more physically analogous choice to keep the background noise level in different images and varying I_0, but has the advantage of simplifying image plotting and later analysis.

A mock dataset meant to approximate one which might be used in real studies must take into account selection effects, whereby very low quality images with which fits are made almost impossible are rejected. Therefore, defining a notion of the "signal-to-noise ratio" of an image to be

$$\text{SNR} = \frac{I_0}{\text{RMS}_{\text{bg}}} = \frac{1000}{\text{RMS}_{\text{bg}}}, \tag{3.14}$$

images of sources at the maximum luminosity distance (with z_s close to 5) were given the minimum mean signal to noise ratio $\mu(\text{snr})$ which still enabled Einstein rings to be distinguished by eye (chosen to be 150). This value could then be scaled appropriately as the redshift decreased. In eq. (1.125), we saw that the received intensity (and thus the source amplitude in `lenstronomy`'s definition, and hence the background noise in our case) is simply equal to the intensity at the source divided by a factor of $(1+z_s)^4$. If it is further assumed that the intrinsic source amplitudes distribute themselves around the same mean, regardless of z_s, the mean observed signal-to-noise ratio at redshift z_s can therefore be calculated as

$$\mu(\text{snr}) = 150 \times \left(\frac{1 + z_{s,\text{max}}}{1 + z_s}\right)^4 = 150 \times \left(\frac{6}{1 + z_s}\right)^4. \tag{3.15}$$

The actual value of the SNR to be used when generating a specific set of images is then calculated by sampling a Gaussian distribution centred on $\mu(\text{snr})$, with a standard deviation of 50 %. To avoid skewing the final dataset with images which are either too low or unrealistically high in quality, SNR values below 100 or above 10^5 are rejected. The resulting distribution of snr values for the 195 lens sample is shown in fig. 3.7. Once a SNR value has been chosen, RMS_{bg} is calculated using eq. (3.14). To generate the actual Gaussian background noise contribution, the noise level in an individual pixel is found by sampling from a standard normal distribution and multiplying the result by RMS_{bg}, a process which is repeated for each of the pixels in the n_{pix} by n_{pix} pixel grid. This noise is added to the signal on a pixel by pixel basis to generate the image.

With the value of RMS_{bg} chosen, three images were created per set of lens, source, cosmology and data parameters, with the point source images excluded in the first, the extended source image excluded in the second, and both included in the third. The observant reader will notice that, with the SNR defined relative to the extended source, and with point source amplitudes being much larger than the integrated amplitudes of these extended sources, point source images will dominate over the background. This ensured that the composite extended + point source images could be compared directly to their single-source counterparts. The trade-off was that poor SNR point source images were not included in the data set. However, because much less information is contained in images of lensed point sources than those of extended galaxies, poor quality images featuring only point sources would not likely be used in attempts to measure the external shear. Furthermore, while point sources are essentially characterised by a single flux, extended sources feature a continuous range moving from their bright centres to fainter edges, and thus exploring the effect of variations in SNR is much more relevant for an extended source. As flux ratios are not considered here, the only relevant datapoints are image positions and relative time delays, neither of which have measurements which are strongly limited by the SNR for bright point

sources. Rather, the more informative comparison is to consider the SNR below which Einstein rings are less useful than clear images of multiply-imaged point sources, an intuition confirmed by the results presented later in this chapter.

The last step in finalising these images was to add Poisson noise, arising from the Poissonian uncertainty in the true signal as a discrete number of photons reaches the detector. To generate this, a mean value of the Poisson noise in a specific pixel is calculated according to

$$\sigma_{\text{Poisson}} = \sqrt{\frac{f_{\text{pix}}}{t_{\text{exp}}}}, \tag{3.16}$$

where f_{pix} is the number of counts per unit time in that pixel (after the addition of background noise), and t_{exp} is the exposure time. There is a fair amount of variations in exposure times for lensing observations in the literature (e.g. [148, 37]), but our choice was to simply set t_{exp} to 1600s for each image, which is a typical value for lensing observations with Hubble.

Figure 3.2 shows the simulation of a source, the effects of lensing, and the addition of background and Poisson noise to the image. The final result of these procedures was 195 sets of mock strong lensing images, meant to represent a simplified sample which is nonetheless representative in its distribution of parameters. A subset of this sample is shown in fig. 3.8.

While the significance of the decrease in intensity experienced by a light ray as it travels from source to observer is readily understood and simulated, this change in intensity is further complicated by the fact that, in any real survey, these light beams are not observed across the entire spectrum, but rather in only across a small range of wavelengths (or in a handful of such ranges). As a result of the cosmological redshift, the peak emission from the source can be shifted into or out of the wavelength band being observed. The so-called K-correction required to account for this requires knowledge of the expected spectra of the sources, and cannot be feasibly implemented in the mock data used for this work. Nonetheless, it would contribute an additional and potentially non-negligible source of error in a real dataset.

Time delays

To complete the dataset, the measured relative time delays between each point source image for each set of lensing events needed to be calculated. To this end, assuming an external convergence of 0, the true absolute and relative time delays in days were calculated for each image. The measured relative time delay was chosen by sampling a normal distribution centred on the true relative time delay, and with a standard deviation which itself was sampled from a uniform distribution between 0.3 % and 25 % of the true relative time delay. For the most part, this range is typical for the reported uncertainties of very well-measured relative time delays (e.g. in [148, 37, 154, 102, 111, 27]). It should be noted in these references and others that uncertainties of less than 1 % are rare and very difficult to obtain, but this best-case scenario is nonetheless included in our dataset. These high-precision measurements offer the best candidates for improved constraints with time delays, and thus are informative as to whether any advantages are, in theory, possible. Though these uncertainties can, in many cases, be larger, the width of the normal distribution is capped at 25 %, as we would not expect relative time delay measurements with very large uncertainties to offer any improvements when constraining shear measurements. This intuition is confirmed in section 3.3.

It should be noted that these time delay uncertainties are simply sampled from uniform distributions, depending only on the magnitude of the relative time delay, with no correlation with the quality of the image imposed. Although the quality of the image is of course a factor, there are multiple sources of uncertainty in time delay measurements – the intrinsic variability of the point source, microlensing phenomena etc, and these tend to dominate in importance over the signal-to-noise ratio of the image.

3.2.2 MCMC Fitting

Having generated 195 sets of 3 images, and having determined measured relative time delays for the point source images within each set, the MCMC method implemented in the `emcee` package [60] was used to measure parameters of interest. After a trial period, it was found that 4000 burn-in iterations and 4000 standard iterations was sufficient to guarantee convergence in most cases, and for those in which the chain did not converge, this remained the case even with much longer chains. These numbers were therefore used for each chain, with ten walkers per parameter. The main lens model was set to an EPL, and the source light model to a Sérsic ellipse when an extended source was present. Line-of-sight effects, however,

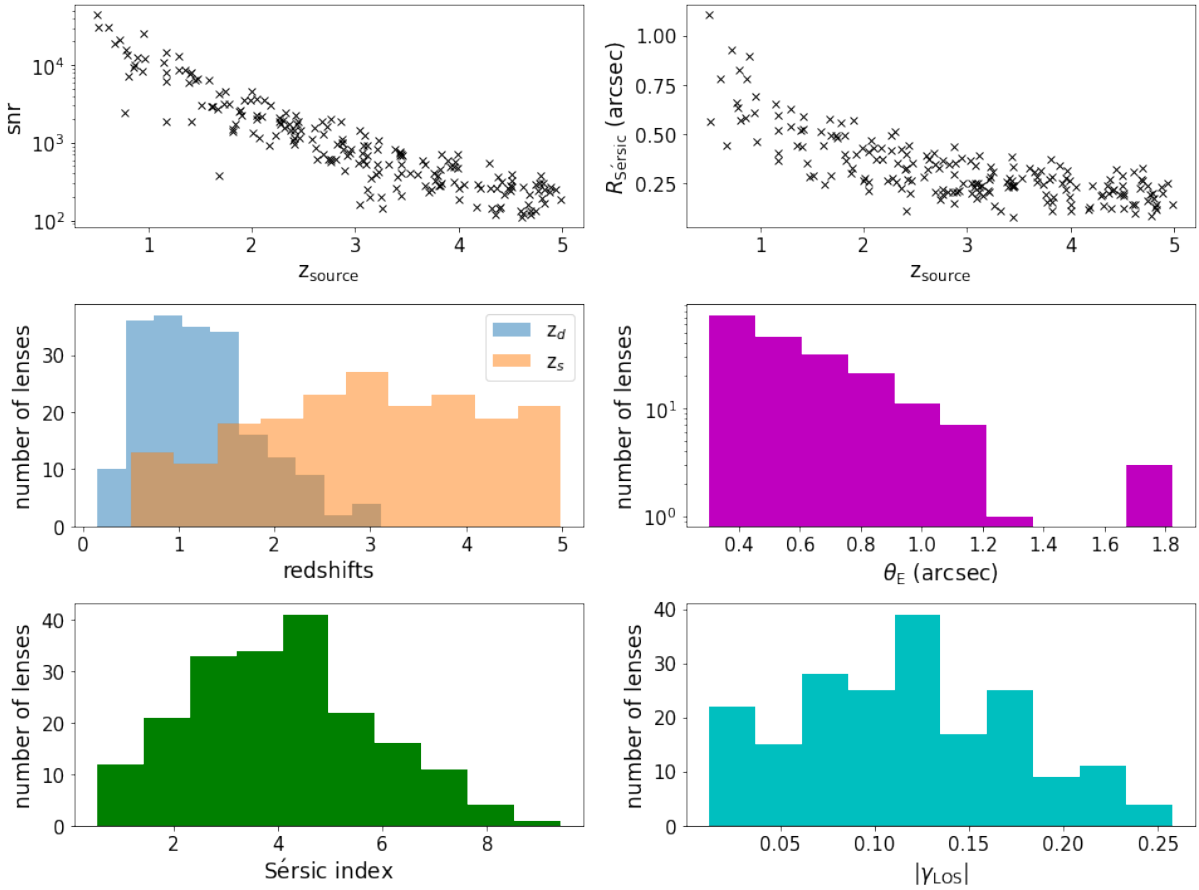

Figure 3.7: *The distributions of certain parameters used in generating the set of 195 mock lensing systems. In the top left figure, we see that source galaxies at higher redshifts will have weaker signals relative to the background sky. From the top right figure, we see that these more distant sources will also have, on average, slightly smaller Sérsic radii, although this dependence is less strong. From the central left image, we see that the lens redshifts peak around* $z_d = 1$*. This peak is pulled leftwards by the dependency of* $\theta_{\rm E}$ *on* z_d *(fig. 3.4) and rightwards by the increase of comoving volume with redshift (fig. 3.3). In the central right figure, we see, unsurprisingly, that smaller values of* $\theta_{\rm E}$ *are more frequent. However, for small enough values of* z_d *and large enough* σ_v*, values are found as high as* $\theta_{\rm E} \sim 1.8$*. In the bottom left figure, we see the the Sérsic indices follow the desired distribution, peaking near 4 but with a fair spread in values. Finally, in the bottom right figure, we see that* $\gamma_{\rm LOS}$ *values are predominantly small, but values as high as* $|\gamma_{\rm LOS}| \sim 0.25$ *are possible, in keeping with results presented in the literature.*

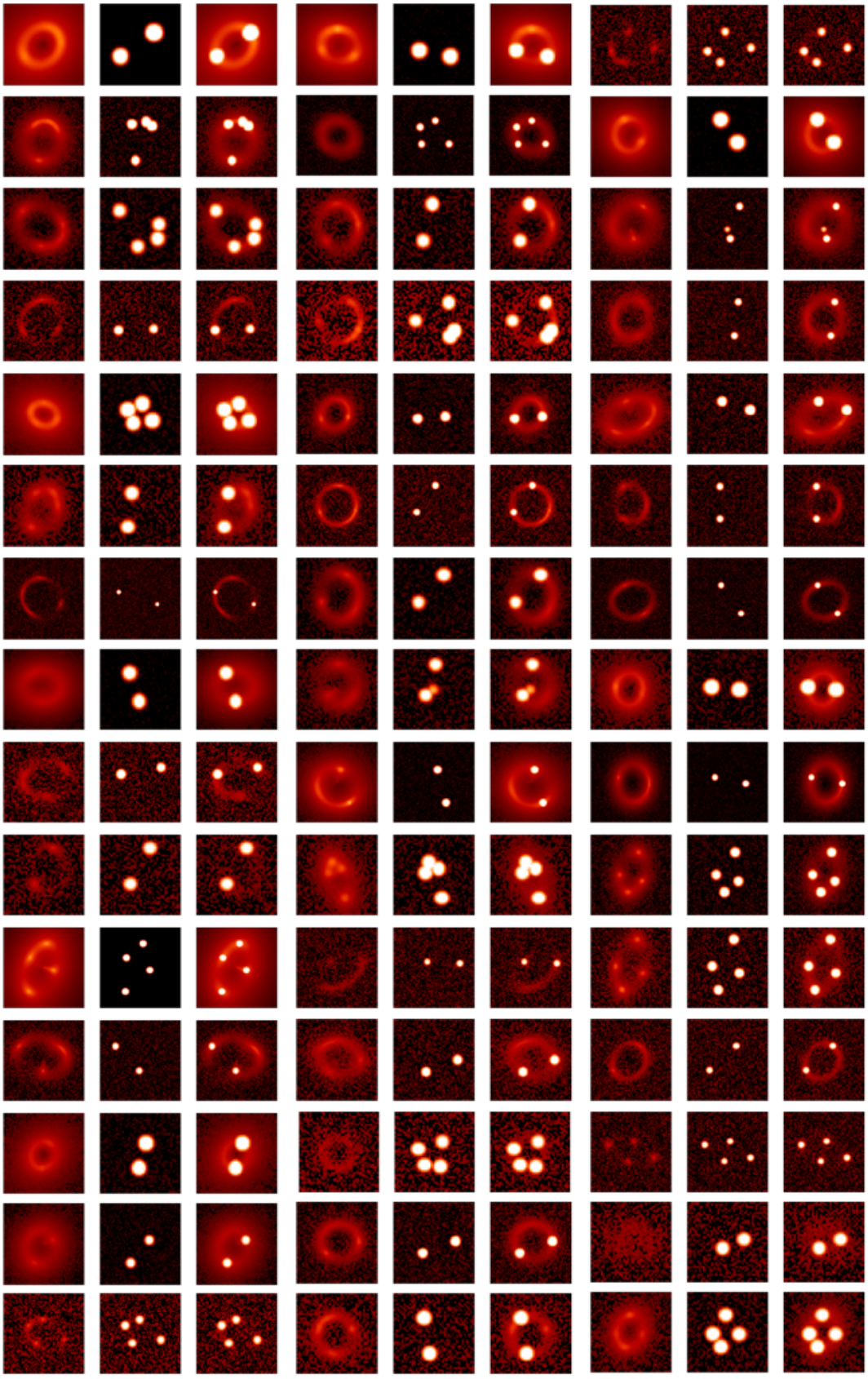

Figure 3.8: *A subset of the mock images used to investigate the value of point sources and time delay measurements on a statistically representative sample of lenses. Images are generated in sets of three: one featuring a lensed extended source, one a lensed point source, and the third a combination of the two.*

were modelled with the `LOS_MINIMAL` profile from section 2.2.2, for reasons discussed in section 1.3.4. The centre of the lens was fixed to $\boldsymbol{\theta} = \mathbf{0}$, κ_{od}, ω_{od} and $\kappa_{\rm LOS}$ were each fixed to 0, and $\omega_{\rm LOS}$ was left to vary as a free (nuisance) parameter. This latter choice is generally necessary to account for departures from the linear regime arising from strong shear effects and improve $\gamma_{\rm LOS}$ measurements [77]. When both point sources and extended sources were present, the position of their centres were coupled in the fit. All parameters were initialised with wide uniform priors, and an initial guess equal to the true value. As discussed in section 3.1, this approach is appropriate when seeking to comment on the shape and width of posterior distributions, and by ensuring sufficiently wide priors and enough burn-in iterations, we ensure that the parameter space is adequately explored and that the resulting posteriors are unbiased.

For each image of an extended source, only one MCMC was run. For each of the two images per lensing configuration which featured a point source, four runs were carried out. In the first, the measured time delays were omitted entirely. In the second, they were included in the dataset, with uncertainties given by the width of the normal distribution from which they were chosen (as described in section 3.2.1), and the time delay distance τ_{ds} was sampled as a parameter within the model, as would be the case if a measurement of H_0 was desired. In the third, rather than sampling τ_{ds}, it was fixed to the "true" value used in generating the mock image. In the fourth, τ_{ds} was instead fixed to an "incorrect" value, chosen by calculating τ_{ds} with the given values of z_s and z_d, but with a value of H_0 chosen between 60 and 100 km.s^{-1}.Mpc^{-1}. These variations are summarised and assigned labels in table 3.1. The result of these choices was 9 MCMC chains, from which best estimates and 1-σ uncertainties for each fitted parameter and 1 and 2 dimensional posterior distributions could be extracted.

Image types and fitting procedures			
image	**measured time delays**	**time delay distance**	**label**
extended source	n/a	n/a	E
point source	excluded	n/a	P
	included	sampled	PS
		fixed correctly	PF
		fixed arbitrarily	PA
point source + extended source	excluded	n/a	EP
	included	sampled	EPS
		fixed correctly	EPF
		fixed arbitrarily	EPA

Table 3.1: *The different types of fits considered when investigating measurements of $\gamma_{\rm LOS}$, γ_{od} and e. Each fit type is identified with a label, and corresponds to a unique combination of choices of source, whether time delay measurements are included, and, if included, whether and how τ_{ds} is fixed.*

3.2.3 Interpreting results

An example of the resulting output is shown in fig. 3.9. The dotted lines show the true value of the fitted parameter, the contour plots show the 2D posteriors, and the plots lying along the upper diagonal of the image show the 1D posteriors.

The 2D posteriors can be used to identify degeneracies - a roughly circular plot shows that the two parameters are non-degenerate, while a highly elliptical plot indicates a degeneracy between two parameters. For example, it is obvious from fig. 3.9 that γ_1^{od} and γ_2^{od}, the foreground shears, are highly degenerate with e_1 and e_2, the ellipticities of the dominant lens. Such degeneracies are not seen between other parameters.

In the analysis which follows, we will be primarily interested in features appearing in the 1D posteriors. In many cases, in particular for poorly-aligned or high SNR images, the MCMC parameter inference cannot constrain one or both components of $\gamma_{\rm LOS}$, γ_{od} or e. However, if, for example, e_1 cannot be constrained, it may nonetheless be valuable to consider constraints on $\gamma_{\rm LOS}$. Furthermore, an image which provides constraints on one of the shear components is still more informative than one in which neither can be constrained. We will let $p_{\mathcal{C}}$ be the percentage of parameter components which could be constrained, out of the 390 in total which we seek to measure for a given type of shear. For example,

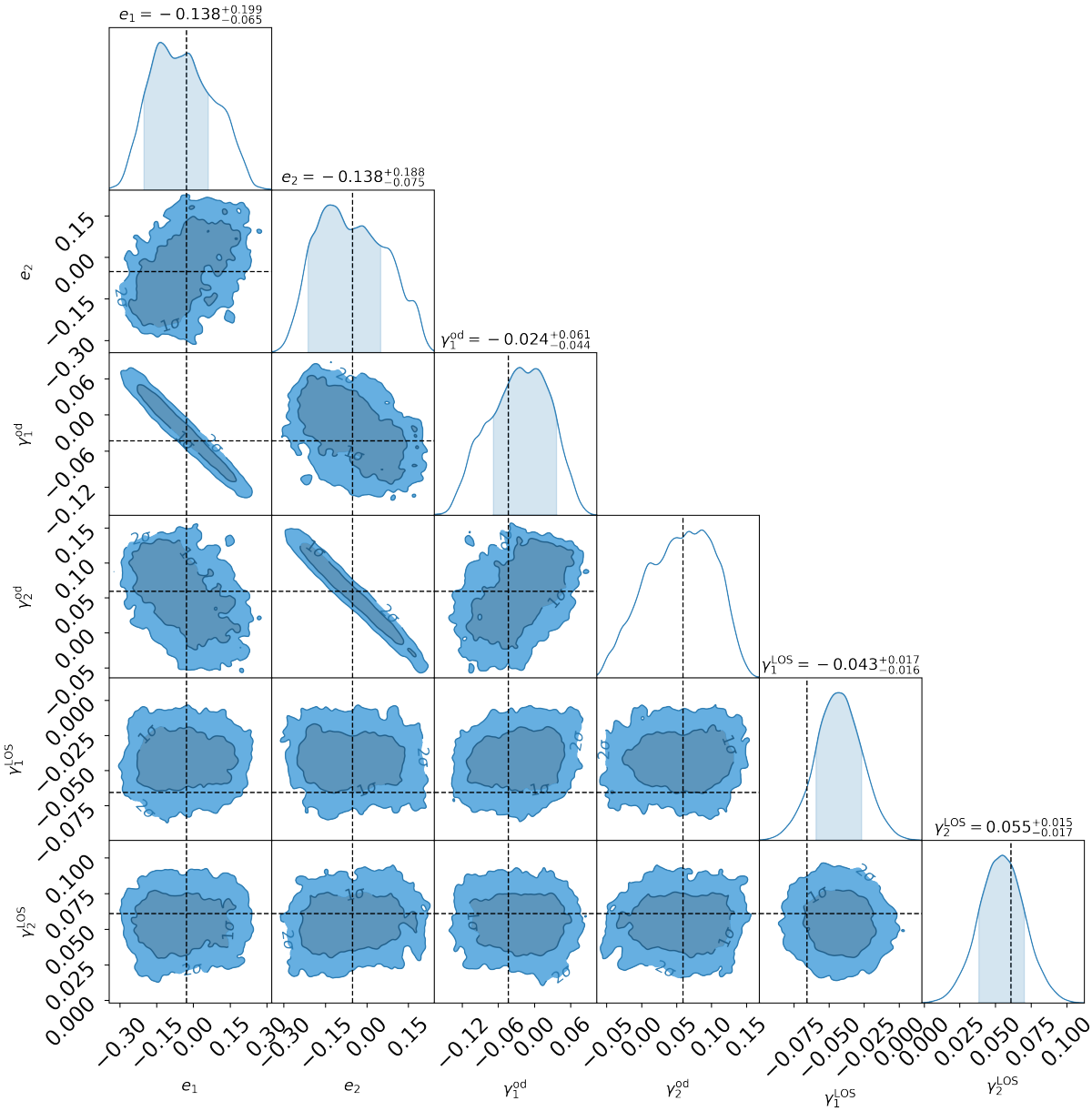

Figure 3.9: *One and two dimensional marginalised posterior distribution contour plots showing the constraints on the line-of-sight shears $\gamma_{\rm LOS}$ and γ_{od}, and well as on the main lens ellipticities, for an image featuring both an extended source and a point source, with relative time delay measurements included and the time delay distance sampled. This plot was produced using the* `chainconsumer` *package [75].*

a dataset of 195 images for which $\gamma_{\rm LOS}$ constraints were being compared, and in which $\gamma_1^{\rm LOS}$ could be constrained every time, but $\gamma_2^{\rm LOS}$ could never be constrained, would have $p_{\mathcal{C}} = 50\,\%$.

In those cases where a parameter could be constrained, the posterior distributions are not Gaussian in general, and the 1σ 68 % confidence domain is typically asymmetric. This can be easily seen in the 1D posteriors in fig. 3.9, for example. We will let $\sigma^{\pm}_{a,i}$ denote the upper $(+)$ and lower $(-)$ 1σ uncertainties for component $a \in 1, 2$ of a given parameter in image i for a given fit type, and define the symmetrised relative uncertainty on component a for image i to be

$$\sigma_{a,i} \equiv \frac{\sigma^{+}_{a,i} + \sigma^{-}_{a,i}}{2|\mathsf{b}^{\rm out}_{a,i}|}, \tag{3.17}$$

where $\mathsf{b}^{\rm out}_{a,i}$ refers to the peak of the posterior distribution for component a of parameter $\mathsf{b} \in \{\gamma_{\rm LOS}, \gamma_{od}, e\}$ in image i, or in other words, the best estimate from the MCMC parameter inference of component a the value of parameter b for i. Similarly, $\mathsf{b}^{\rm in}_{a,i}$ will refer to the "true" inputted value of component a of that parameter in i.

It will often be of interest to consider the proportion of images in which the $\sigma_{a,i}$ is within a certain percentage of $\mathsf{b}^{\rm out}_{a,i}$. We define $n_{\sigma<t}$ to be the number of times in which the value of $\sigma_{a,i}$ was less than the threshold percentage t across both components a and all images i, for a given parameter and fit type, and $p_{\sigma<t}$ to be this number as a percentage of $n_{\mathcal{C}}$.

Also of interest is the accuracy of the parameter recovery, which we will measure in terms of a quantity akin to the reduced chi-squared which takes into account the asymmetry of the 1σ errors, which we define by

$$\chi^2 \equiv \frac{1}{2n_{\rm c}} \sum_{a,i \in \mathcal{P}} \frac{1}{\tilde{\sigma}^2_{a,i}} \left(\mathsf{b}^{\rm out}_{a,i} - \mathsf{b}^{\rm in}_{a,i}\right)^2, \tag{3.18}$$

with the sum taken over all parameter components which were successfully constrained for the parameter and fit type under consideration, and with

$$\tilde{\sigma}_{a,i} \equiv \begin{cases} \sigma^{+}_{a,i} & \text{if } \mathsf{b}^{\rm out}_{a,i} \leq \mathsf{b}^{\rm in}_{a,i} \\ \sigma^{-}_{a,i} & \text{if } \mathsf{b}^{\rm out}_{a,i} > \mathsf{b}^{\rm in}_{a,i}. \end{cases} \tag{3.19}$$

We also consider the percentage $p_{<\sigma}$ of parameters which lay within the 1σ errors returned by the code, to identify cases in which the parameter inference over or under-estimates uncertainties. For well-defined error bars, we would expect $p_{<\sigma} \approx 68\,\%$.

In the example plotted in fig. 3.9, we see that the parameter $\gamma_2^{\rm od}$ could not be constrained at all in this particular fit, whereas all the other parameters could be constrained. We also see, for example, that the values of $\sigma_{a,i}$ on $\gamma_1^{\rm LOS}$ (~ 0.38) and $\gamma_2^{\rm LOS}$ (~ 0.29) are lower than those on e_1 and e_2 (~ 0.95 in each case), as well as on $\gamma_1^{\rm od}$ (~ 2.2). We also see that the true value of the parameter falls within the error bars for all of the constrained parameters except $\gamma_1^{\rm LOS}$.

3.2.4 Measuring the line-of-sight shear

Measuring the line-of-sight shear using extended sources

Before making any comparisons with other fit types, we consider the results of measuring $\gamma_{\rm LOS}$ from an extended source only. In general, for these simple models, $\gamma_{\rm LOS}$ is well recovered when an image of a strongly lensed extended source is fitted, as shown by the results in table 3.2. The two parameters could be constrained in 99 % of fits ($p_{\mathcal{C}} = 99\,\%$), and these constraints were within 20 % of the reported best guess at the parameter's true value in 71 % of the fits ($p_{\sigma<20\%} = 71\,\%$), and within 5 % in 42 % of the fits ($p_{\sigma<5\%} = 42\,\%$). Thus very tightly constrained results are not guaranteed, but reasonably good constraints are nonetheless possible in a majority of cases.

For the same dataset, $p_{<\sigma}$ is slightly lower than the expected value of 68 %, though not by much. Furthermore, the value of $p_{>2\sigma}$ indicates that there were more than twice as many outliers than expected, and the χ^2 parameter is greater than would be expected for well-defined error bars. These inaccuracies seem to result from an overestimation of constraints for more poorly aligned source-lens pairs, a point which is discussed further below.

Figure 3.10 shows the relationship between the precision and accuracy of the fits, and various parameters of the input model. It is clear that the signal-to-noise ratio has a strong influence on the tightness of the constraints, and this was found throughout this work to be the dominant determinant of how well line-of-sight shears could be measured. It is also clear that the Einstein radius has an influence, though

$\gamma_{\rm LOS}$ *fitting results when extended sources are present*						
		E	**EP**	**EPS**	**EPF**	**EPA**
constraints	$p_{\mathcal{C}}$	99 %	99 %	100 %	99 %	100 %
precision	$p_{\sigma<5\%}$	42 %	41 %	43 %	42 %	43 %
	$p_{\sigma<20\%}$	71 %	73 %	74 %	74 %	74 %
accuracy	$p_{<\sigma}$	64 %	66 %	66 %	66 %	66 %
	$p_{>2\sigma}$	11 %	11 %	11 %	12 %	11 %
	χ^2	4.26	4.24	4.45	4.76	4.63

Table 3.2: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source and extended source, for the various corresponding fit types described in table 3.1, for measurements of* $\gamma_{\rm LOS}$.

this is not quite as strong as in the case of the signal-to-noise ratio. It should be noted here that, while both the SNR and Einstein radii were sampled from functions of redshift, the SNR is loosely dependent on the source redshift (see fig. 3.7), while the Einstein radius depends primarily on the properties of the lens and the lens redshift, while having only a very weak dependence on the source redshift. Thus the improvements in constraints with $\theta_{\rm E}$ are not simply a result of the associated SNR values.

Another parameter which has been suggested to correlate with the ability to measure line of sight shear is the ratio of the Sérsic radius of the source to the distance between the centre of the source and the centre of the lens (see [77]). A source which is perfectly aligned with the lens will in general produce extended images carrying more information about the shear than a source at a larger distance, which are less likely to produce complete Einstein rings. This ratio quantifies the extent to which light from the extended source falls within this critical region, thus producing more complete arcs or rings. We do not see an obvious correlation between the constraints and this ratio. This could, however, simply be because lens events were selected to be quite closely aligned, in order to produce multiple point source images for time delay measurements, and thus there are not enough cases with a ratio less than 1 to make this plot fully illustrative.

More influential on the quality of parameter recovery was the position the centre of the extended source in relation to the caustic curve, as illustrated in fig. 3.14. This measure of the alignment does not depend solely on the source position, as the value of $\theta_{\rm E}$ of the main lens and the influence of the line-of-sight shears are also relevant. At a glance, it appears that, for a given SNR value, an extended source with a centre within the inner caustic (at which position a point source would be quadruply imaged) tends to provide tighter constraints on $\gamma_{\rm LOS}$ than those at positions which would lead only to triple or double imaging. This is not unexpected - although they do not necessarily improve the chances of a complete Einstein rings, sources centered within the inner caustic do tend to produce more featured rings than those placed between the inner and outer caustics, as can be seen in fig. 3.8. We would therefore expect these images to carry more detail about shear effects, and thus lead to more precise measurements. However, while it is true that the largest error bars belong to sources centered outside the inner caustic, a closer inspection of the data reveals that, on average, constraints become *less* tight when considering only sources at positions leading to quadruple imaging. Table 3.3 shows the fitting results when only these images are considered, in which we see that $p_{\sigma<5\,\%}$ and $p_{\sigma<20\,\%}$ drop from 42 % to 37 % and from 71 % to 65 % respectively. If only source positions which would lead to double imaging are considered, these parameters instead increase to 46 % and 73 % respectively. However, the errors in the latter case are clearly underestimated, with a χ^2 value of 5.9. When only sources centered within the inner caustic are considered, the values of $p_{<\sigma}$, $p_{>2\sigma}$ and χ^2 all imply better-defined error bars and more accurate results. It seems, therefore, that sources outside the inner caustic lead to an underestimation of error bars, introducing outliers into the overall dataset and degrading the χ^2 value across all images. When only well-aligned sources are considered, tight constraints are still possible, and these constraints are more reflective of the true uncertainties in the measurements. Above all, the clarity of the image above the background noise is the most important factor for tight constraints.

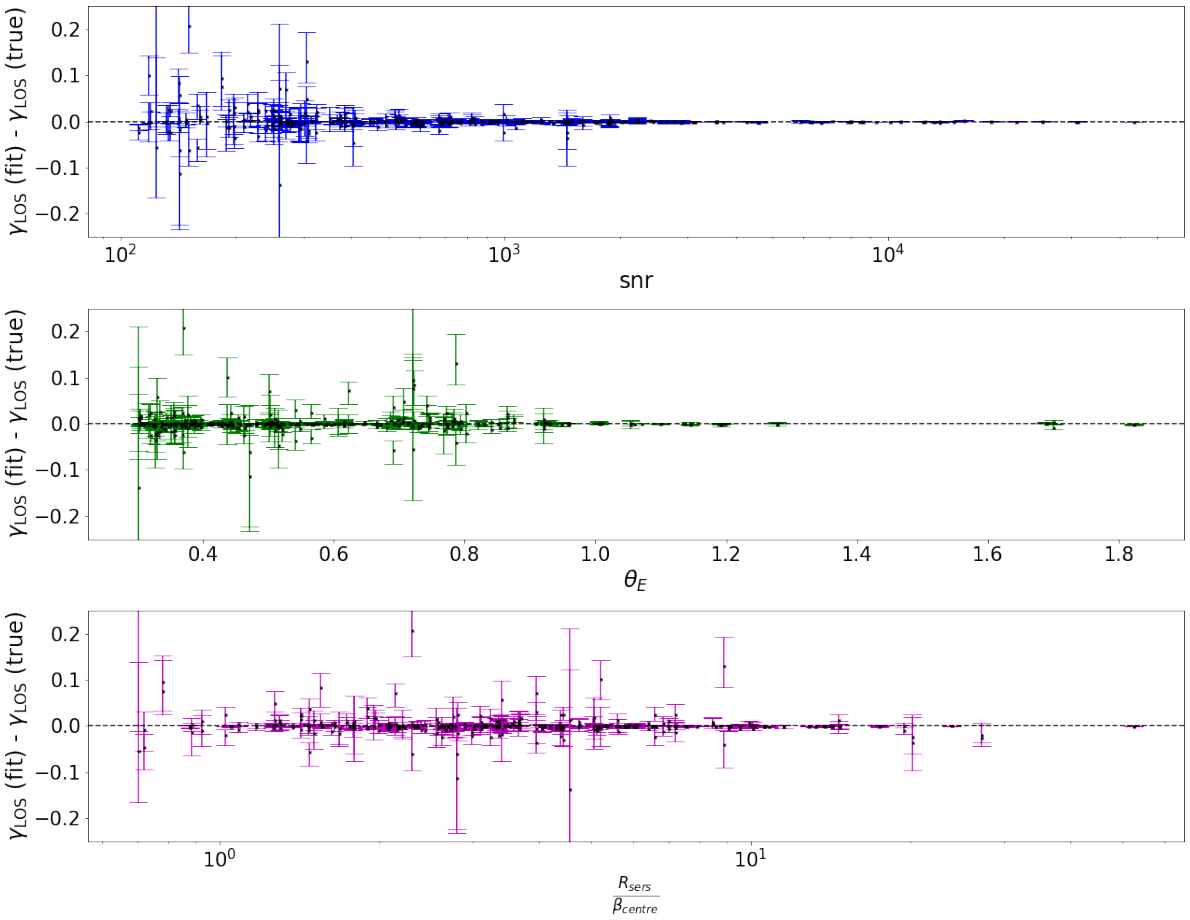

Figure 3.10: *The accuracy and precision of $\gamma_{\rm LOS}$ measurements from mock images of lensed extended sources, as a function of the signal-to-noise ratio (top plot), Einstein radius (centre plot), and the ratio of the Sérsic radius of the source to the distance between the centre of the source and the optical axis passing through observer position and the centre of the lens. The black crosses show the best estimate of the value of the parameters, and the error bars correspond to the 1-σ uncertainty on this value.*

$\gamma_{\rm LOS}$ *fitting results with extended sources (quadruply-imaged sources only)*						
		E	**EP**	**EPS**	**EPF**	**EPA**
constraints	$p_{\mathcal{C}}$	100 %	98 %	99 %	100 %	100 %
precision	$p_{\sigma<5\%}$	37 %	26 %	29 %	29 %	30 %
	$p_{\sigma<20\%}$	65 %	67 %	69 %	69 %	70 %
accuracy	$p_{<\sigma}$	69 %	74 %	76 %	75 %	74 %
	$p_{>2\sigma}$	7 %	2 %	4 %	5 %	5 %
	χ^2	1.13	0.89	0.97	1.02	1.26

Table 3.3: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source and extended source, for the various corresponding fit types described in table 3.1, for source-lens alignments leading to quadruple imaging, for measurements of* $\gamma_{\rm LOS}$.

The effects of adding point source images

The importance of extended sources to the measurements of the external shear has been investigated in studies such [77, 59, 101]. A question of interest to this work was the effect of adding point source images to the data with which the MCMC chains were run. This question can be asked in two ways - firstly, what are the effects of adding a point source to the centre of an extended source which is strongly lensed, and secondly, what are the effects of fitting a strongly lensed quasar in isolation, or in other words, a strongly lensed point source in the absence of an extended counterpart?

Unsurprisingly, in the case where a point source but no extended source is present, thus massively reducing the amount of data relevant to the fit, the ability of the MCMC chain to determine and constrain $\gamma_{\rm LOS}$ is greatly reduced, as can be seen in table 3.4. Here, in 37 % of cases, parameters could not be constrained at all (vs. 1 % for extended sources only). In 81 % of the cases in which parameters could be constrained, the uncertainty on the line-of-sight shears were more than double the uncertainty when an extended source was used, and tight constraints were essentially impossible. These effects can be clearly seen in fig. 3.13. The top plots show the precision and accuracy of LOS shear measurements using an image of an extended source only, while the second from top left shows the same data but for the case where only a lensed point source is present. It is immediately obvious that the constraints in the latter case are much poorer than in the former. Interestingly, $p_{<\sigma}$ is slightly lower than expected, suggesting uncertainties were underestimated in many cases. However, none of the reported values were outliers ($p_{>2\sigma}$). Furthermore, the value of χ^2 suggests that uncertainties are often overestimated. These points are best understood in the context of the hard bounds of ± 0.1 on the line-of-sight shears, which led to the distribution of $\gamma_{\rm LOS}$ values shown in fig. 3.7. From the slightly low value of $p_{<\sigma}$, we see that the scant information offered by point sources only is not generally enough to provide reliable constraints on $\gamma_{\rm LOS}$. However, because, in addition to being imperfectly defined, the uncertainties are very large, there simply are not many $\gamma_{\rm LOS}$ values extreme enough to meet the 2σ threshold to be classed as outliers. Furthermore, because the chains were initialised at the correct values of these parameters, it is possible that a chain which struggled to differentiate between several candidate solutions (and thus reported large uncertainties) would nonetheless be biased by the initial values. The combination of these effects is likely responsible for the low χ^2 value.

The situation improves somewhat when only quadruply-imaged point sources are considered, as is clear from both fig. 3.14 and table 3.5. For these configurations, $p_{\mathcal{C}}$ rises to 97 %, and, while constraints within 5 % remain almost impossible, constraints within 20 % do become more than 3 times more likely, and it can easily be seen in fig. 3.14 that quadruply lensed point sources lead to much tighter constraints than doubly lensed ones. Here, interestingly, $p_{<\sigma}$ is now significantly higher than expected, though there are still fewer outliers than expected. The χ^2 value increases to 0.45, which is still less than half the expected value, but nonetheless an improvement. The tendency to over-report uncertainties is seen consistently in all datasets featuring point sources only, as is clear from tables 3.4 and 3.5, and, in each case, the results are consistent with the explanation suggested in the previous paragraph.

Perhaps somewhat more interesting is the fact that the addition of point source images to an image of an extended source did not, on average, improve the precision or accuracy of the fits, as can be seen in

$\gamma_{\rm LOS}$ *fitting results when only point sources are present*						
		E	**P**	**PS**	**PF**	**PA**
constraints	$p_{\mathcal{C}}$	99 %	63 %	69 %	72 %	71 %
precision	$p_{\sigma<5\%}$	42 %	0 %	0 %	0 %	0 %
	$p_{\sigma<20\%}$	71 %	8 %	10 %	9 %	10 %
accuracy	$p_{<\sigma}$	64 %	58 %	68 %	68 %	64 %
	$p_{>2\sigma}$	11 %	0 %	1 %	1 %	0 %
	χ^2	4.26	0.28	0.25	0.34	0.34

Table 3.4: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source only, for the various corresponding fit types described in table 3.1, for measurements of* $\gamma_{\rm LOS}$.

table 3.2. The values of $p_{\mathcal{C}}$, $p_{\sigma<5\%}$ and $p_{\sigma<20\%}$ were almost identical. $p_{<\sigma}$ increases to 66 %, while the χ^2 value decreases fractionally, but these improvements to the accuracy are very slight. Once again examining fig. 3.13, we see this point illustrated as before. The second from top right plot shows the results of including both lensed point and extended sources in the fitted image, and we see that, overall, the precision and accuracy of measurements are not easily distinguishable from those when only an extended source is used. However, one feature worth noting is the improvement at low signal-to-noise ratios when point sources are included. See section 3.3 for a further exploration of this point.

In the case where only quadruply-imaged point source + extended source combinations were considered, these images actually performed slightly worse in comparison with extended sources only, with the relative fractions of parameters which could be constrained, or constrained to within 5%, both decreasing. From fig. 3.14, it appears that having a greater number of point source images present does offer improvements to the constraints at low SNR values, but from table 3.3, it is clear that this improvement is not seen in the dataset as a whole. Nonetheless, it is interesting to note that, for these quadruply-imaged systems, $p_{<\sigma}$ is greater than expected, there are fewer outliers, and the χ^2 value drops to 0.89, suggesting that uncertainties are somewhat overestimated in these cases.

In summary, therefore, quadruply-imaged quasars did generally provide enough data for line-of-sight shears to be constrained, but these constraints were a lot weaker than for an extended source. No advantage is offered by considering the lensing of an extended and point source together without time delay information versus a lensed extended source only for the mock data considered in this analysis.

These observations can be best understood in combination with fig. 3.13. When comparing the top plot (extended sources only) and the middle plot on the right (point + extended source), we see that the addition of point source images offers improvements for images with lower SNR values, but that these improvements disappear towards higher SNR values. When images feature large amounts of noise, the arcs or ring resulting from the lensing of the extended source are degraded, while the point source images are less affected. When the signal dominates over the noise, arcs and rings from extended sources are clear, while the point spread function becomes the dominant source of uncertainty. The result is that, for higher SNR images, rather than adding more information, the point spread function of point source images limits constraints on shears. This can be seen in the images shown in fig. 3.16, where only the SNR is changed between successive images, and will be further discussed in section 3.3. This also explains why quadruply-lensed point source images compared with the same lensing configuration in the absence of these point source images perform more poorly than images with two or three point source images, as the greater number of point sources in higher SNR images "pollutes" a larger proportion of the image. The degeneracy between the light from the point source and the host galaxy increases as the number of point source present increases.

Including time delays

The next question of interest was whether the inclusion of time delay data to images containing lensed point sources offered any improvements. There are two key pieces to this question: (i) does time delay data improve the measurements of LOS shears over an identical image featuring the same lensed point source but no time delay data? (ii) does the inclusion of time delay data mean that point source +

$\gamma_{\rm LOS}$ *fitting results with point sources (quadruply-imaged sources only)*						
		E	**P**	**PS**	**PF**	**PA**
constraints	$p_{\mathcal{C}}$	100 %	97 %	99 %	98 %	98 %
precision	$p_{\sigma<5\%}$	37 %	1 %	1 %	0 %	1 %
	$p_{\sigma<20\%}$	65 %	25 %	30 %	30 %	31 %
accuracy	$p_{<\sigma}$	69 %	84 %	95 %	91 %	80 %
	$p_{>2\sigma}$	7 %	1 %	4 %	4 %	1 %
	χ^2	1.13	0.45	0.40	0.56	0.59

Table 3.5: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source and extended source, for the various corresponding fit types described in table 3.1, for source-lens alignments leading to quadruple imaging, for measurements of* $\gamma_{\rm LOS}$.

extended source images become more effective than extended source only images?

We first consider the question of whether, given an image featuring point sources, the measurements of $\gamma_{\rm LOS}$ would be improved by the addition of time delay data. We will firstly compare the case where we have point sources, but no extended sources. Here, both the proportion of lenses in which time delay data allowed parameters to be constrained and the number of times the true parameter values lay within the reported error bars increased, but the constraints remained very poor, as can be seen in table 3.4 and fig. 3.13. The inclusion of time delays did mean tighter constraints more often than not, but these effects were small. Furthermore, the χ^2 value remained quite low, suggesting uncertainties are once again overestimated.

When only quadruply-imaged point sources were considered, line-of-sight shears could be constrained 99 % of the time with time delay data, versus 97 % of the time without. Once again, time delays typically lead to tighter constraints, but these effects were small.

Interestingly, the presence of time delay data made the fit even more likely to overestimate the width of the posterior distribution for quadruply-lensed quasars, with 95 % of the true values lying within the reported error bars (i.e. what we would expect for error bars representing 2σ rather than 1σ errors). Even more interestingly, however, is that we have $p_{>2\sigma} = 4\,\%$, meaning that the expected $\sim 27\,\%$ of values found between 1σ and 2σ from the reported value are greatly under-represented.

For images featuring both point sources and extended sources (table 3.2), the inclusion of time delays pushes the proportion of parameters which could be constrained to 100 %, and error bars remain well-defined. Time delay data meant that constraints became tighter about twice as often as the opposite, but in the majority of cases, there were no significant improvements for a given lens system. Again, referring to fig. 3.13 and comparing the second from top right and centre-right plots, it is clear that any improvements offered by including the time delays is marginal. As in the E and EP cases, the χ^2 value for EPS remains poor. When only quadruply-lensed images are considered, the fits improve significantly, though once again seeming to slightly overestimate the uncertainty of reported values. However, as for the dataset taken as a whole, the improvements over the EP case are small.

When the EPS results are compared to those from an extended source only ("E"), as when the same comparison was done without time delays, the addition of point sources and time delays to an image offers no substantial improvement at all to the ability of the fit to constrain parameters, or the tightness of these constraints. Therefore, as was the case with the inclusion of point sources, it seems that, for lensing events comparable to those generated in the mock dataset, using quasar images with time delays offers no real advantages over simply using lensed extended sources to measure the external shear.

The effects of fixing the time delay distance

In the comparisons in section 3.2.4, the time delay distance τ_{ds} was left as a free parameter, and sampled along with other parameters in the fit. While this would be the approach taken when using strong lensing time delays to infer the Hubble constant via τ_{ds}, we wanted to investigate whether fixing τ_{ds}, rather than sampling it, could improve shear measurements by reducing the degrees of freedom in the fit, and possibly alleviating potential degeneracies.

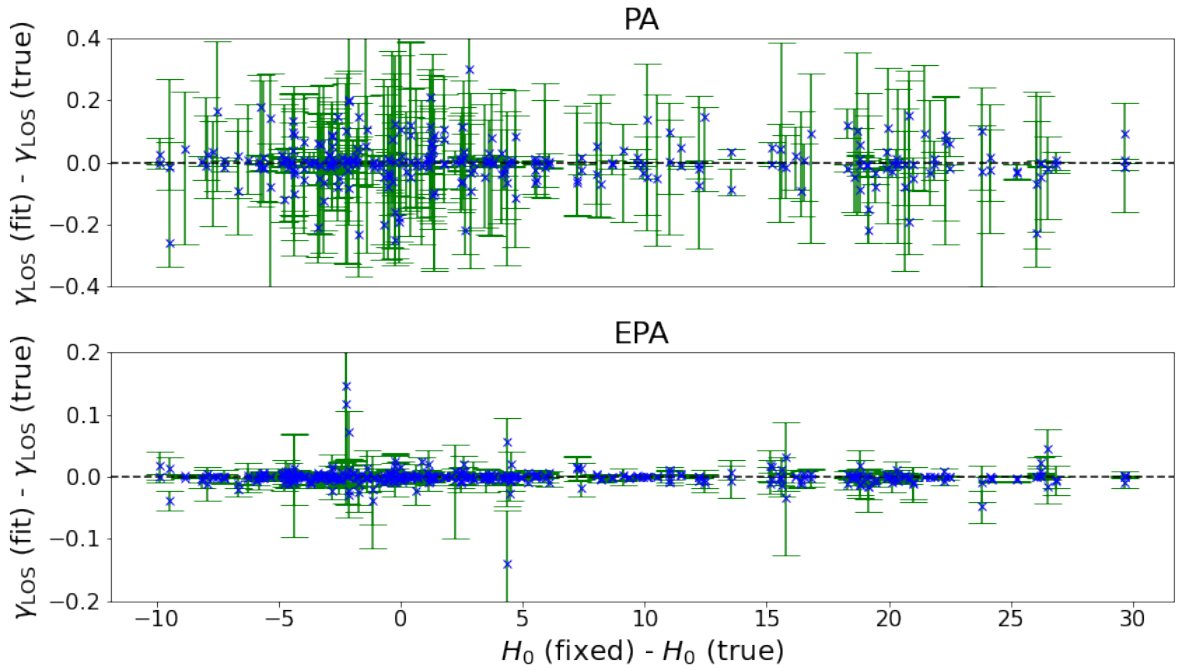

Figure 3.11: *The effects of fixing the time delay distance τ_{ds} to a value calculated with an incorrect Hubble constant on measurements of LOS shear. The difference between the best guess and the true value of the shear is shown in blue, with the reported uncertainties plotted as green error bars. The top plot shows the measurements using a lensed point source and time delay data only, while the bottom plot shows measurements made with the additional presence of a lensed extended source. On the x-axis, the difference between the value of the Hubble constant used in generating the mock data (70 km/s/Mpc) and the value used to fix τ_{ds} is plotted. Note the smaller y-range for the bottom plot, as the presence of an extended source generally leads to much tighter constraints. In both plots, a subset of the more poorly-constrained measurements have error bars which extend beyond the plotted y-axis limits, as these limits were chosen to make any subtle changes across different H_0 more easily visible.*

Using the same simulated mock lensing events, the fits were re-run with τ_{ds} fixed to the "true" value (with a Hubble constant of 70 km/s/Mpc, as used in generating the mock datasets). This was done for both the images featuring only point sources, and those featuring point sources and extended sources.

In both cases, the results were practically identical, with a fixed τ_{ds} leading to no real improvements in $p_{\mathcal{C}}$, $p_{<\sigma}$, $p_{>2\sigma}$, or the tightness of constraints, taken either as an average across the entire dataset, or in terms of changes between identical lenses from the EPS/PS to the EPF/PF cases, as can be seen in tables 3.2 and 3.4. The same patterns were seen when considering only quadruply-lensed point sources (tables 3.3 and 3.5), with no substantial improvements seen with a fixed τ_{ds}. This suggests that LOS shear is mostly uncorrelated with the overall normalisation of the time delays.

To further investigate this point, we redid each fit, but this time fixing τ_{ds} to an incorrect value, chosen by fixing the Hubble constant to values between 60 and 100 km/s/Mpc. Here, as before, the results were effectively identical, with the fixing of τ_{ds} to a value other than that used in generating the mock images doing nothing to improve or degrade the ability of the fits to constrain parameters or the tightness of these constraints, nor introducing any systematic biases into the reported values of LOS shears. This can be seen in fig. 3.11. The unimportance of fixing τ_{ds} is well illustrated by fig. 3.13, where the middle (τ_{ds} sampled), second from bottom (τ_{ds} fixed correctly) and bottom (τ_{ds} fixed randomly) rows show no perceivable difference whatsoever in accuracy or constraints of measured LOS shear parameters.

Illustrating the precision and accuracy of different datasets graphically

Figure 3.13 allows us to make some additional, general comments about the effectiveness of different datasets in measuring and constraining line-of-sight shear parameters. Firstly, it is clear that no dataset

improves significantly on the case where only an extended source is present, except perhaps for highly noisy images (see section 3.3 for more on this point). We also see that, for every image featuring an extended source, the signal-to-noise ratio has a very clear effect on the quality of the measurements, being a much more dominant effect than the presence or absence of point sources and time delay data. However, this is not at all the case when the extended source is absent, and only a point source is lensed. Here, signal-to-noise ratio appears entirely uncorrelated with the quality of the measurements, which is not unexpected. Since flux ratios are not used in these fits, as the information they carry is often greatly obscured by microlensing effects, the only quantity available with which to constrain parameters is the positions of the images (and the time delays, when those are used). The uncertainties in the positions are dominated by the point spread function, rather than the signal-to-noise ratio. Furthermore, since the images are generated with a background noise level calculated with respect to the brightness of the extended source, and since quasars are in general many times brighter than their host galaxies, this specific dataset does not feature images in which the point source images cannot be easily distinguished from the background noise. In cases where the images are faint enough for this background to be at a comparable signal strength, we would expect the SNR to become relevant to the quality of the fits. However, given how poorly even very clear point source images can constrain line-of-sight shears, those noise-dominated point source images are not of any real interest for shear measurements.

Comparing lensing events with different numbers of point source images

Figure 3.14 illustrates the value of having a greater number of point source images when measuring line-of-sight shears. The reported shear values and their uncertainties in these plots are colour-coded to indicate the number of point source images featured on the generated image used in the fit.

From the top plot, we see that, even in the case where no point sources are present, a lensing configuration which would quadruply lens a point source located at the centre of the extended source does go some way towards reducing outlying results, in particular for lower SNR images. However, as discussed previously, there was actually a smaller proportion of tighly constrained results from sources centered within the inner caustic, compared with sources located between the inner and outer caustics. If, by chance, source positions leading to double imaging were located at higher SNR and/or θ_{E} values than those leading to quadruple imaging, this could explain the trends in tightness of constraints. However, as can be seen in fig. 3.12, this does not seem to be the case. This result instead seemed to arise from an overestimation of constraints for more misaligned sources, as can be seen from improvements in accuracy in the E column from table 3.2 to table 3.3.

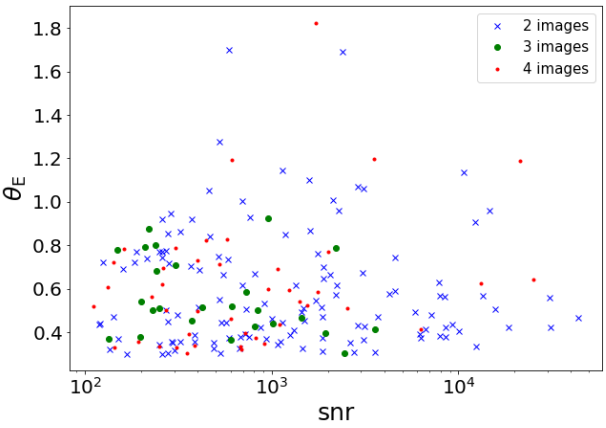

Figure 3.12: *The distributions of θ_{E} and SNR values used when generating the 195 sets of lensing images, divided by colour according to the number of solutions to the lens equation for the chosen source position (hence the number of images which a point source at that position would produce). On this plot, the mean coordinates are (3300,0.59) for 2 images, (813,0.56) for 3 and (2200,0.61) for 4.*

From the centre plot, we see that the effect of having a source within the inner caustic in removing outlying results is even stronger, and all of the worst-constrained results are those with only 2 point source images. Comparing tables 3.2 and 3.3, we see that the effect on $p_{\sigma<5\%}$ is even stronger, dropping from 41 % when considering the dataset as a whole to 26 % for sources within the inner caustic. However, the χ^2 value improves, and it seems that uncertainties are again underestimated in the double-imaging case. This would go some way towards explaining the trends in $p_{\sigma<5\%}$, as would the observations discussed in section 3.2.4, that adding point sources images to lensed extended sources can offer better constraints on poorer quality images, but can obscure details in higher quality images which makes very precise constraints more challenging, an effect exarcerbated by having more point source images present. These trends carry through to the cases where time delay data are also present.

From the bottom plot, we see that, just as would be expected, the effect of having a greater number of point source images is strongest when no extended source is present. Here, quadruply-imaged point sources allow for much tighter constraints than doubly-imaged point sources, although these constraints

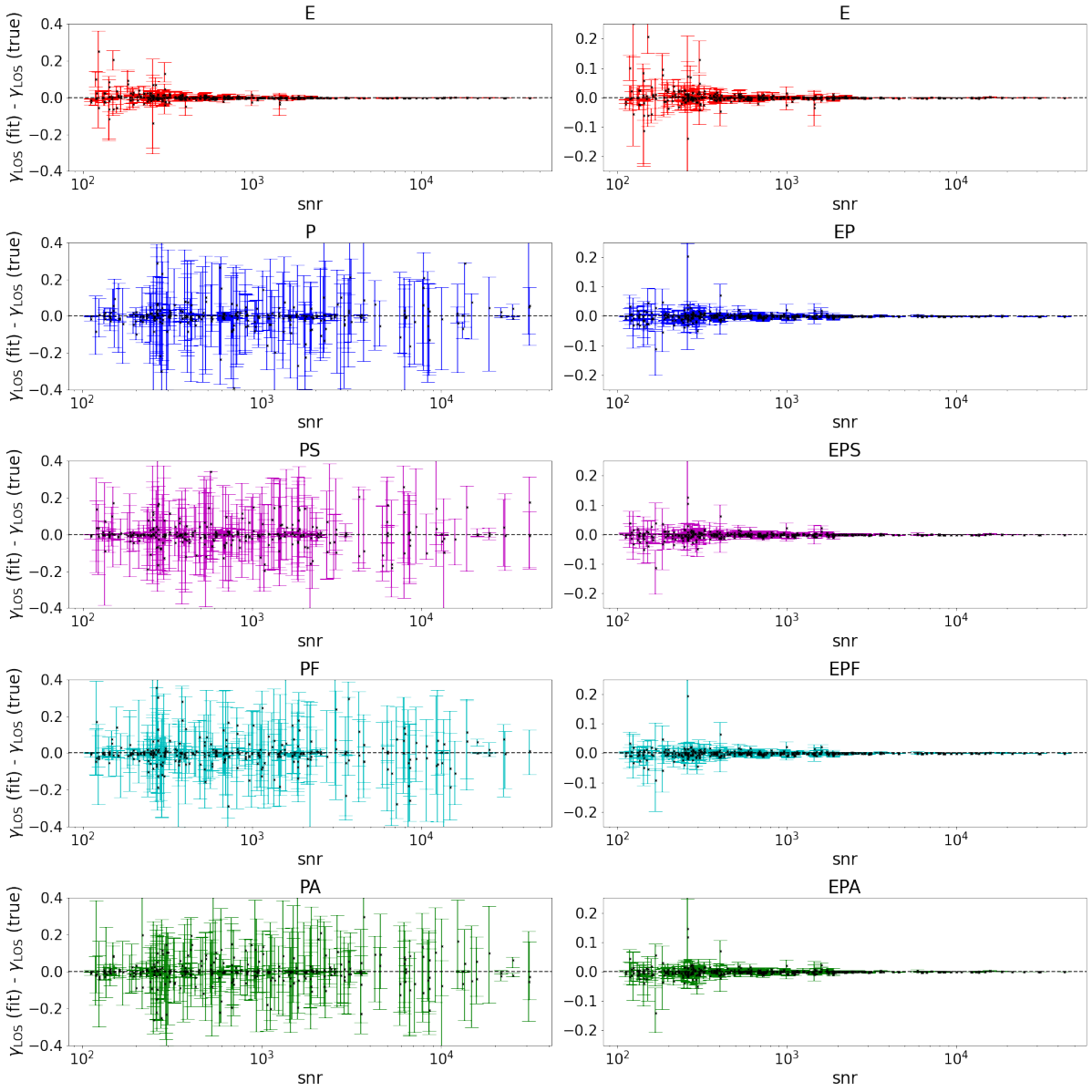

Figure 3.13: *Accuracy and precision of $\gamma_{\rm LOS}$ measurements for the various datasets and fitting procedures considered. Black crosses show the best estimate of the parameter's value (minus the value used to generate the mock data), and the error bars show the 1-sigma uncertainties on these values. Results measured from images with extended sources only are shown in the top row (in red), plotted twice for ease of comparison down the column. Below these plots, the right-hand column corresponds to images with both point and extended sources, and the left-hand column to those featuring point sources only. Going down successive rows, plots with dark blue error bars show measurements from data with no time delays, those with purple from data including time delays, with τ_{ds} sampled, those with light blue from data with time delays, with τ_{ds} fixed to the value used in generating the data, and those with light blue from data with time delays, with τ_{ds} fixed to a randomly-chosen incorrect value. As in fig. 3.11, the y-axis is restricted to better illustrate subtler differences in the constraints.*

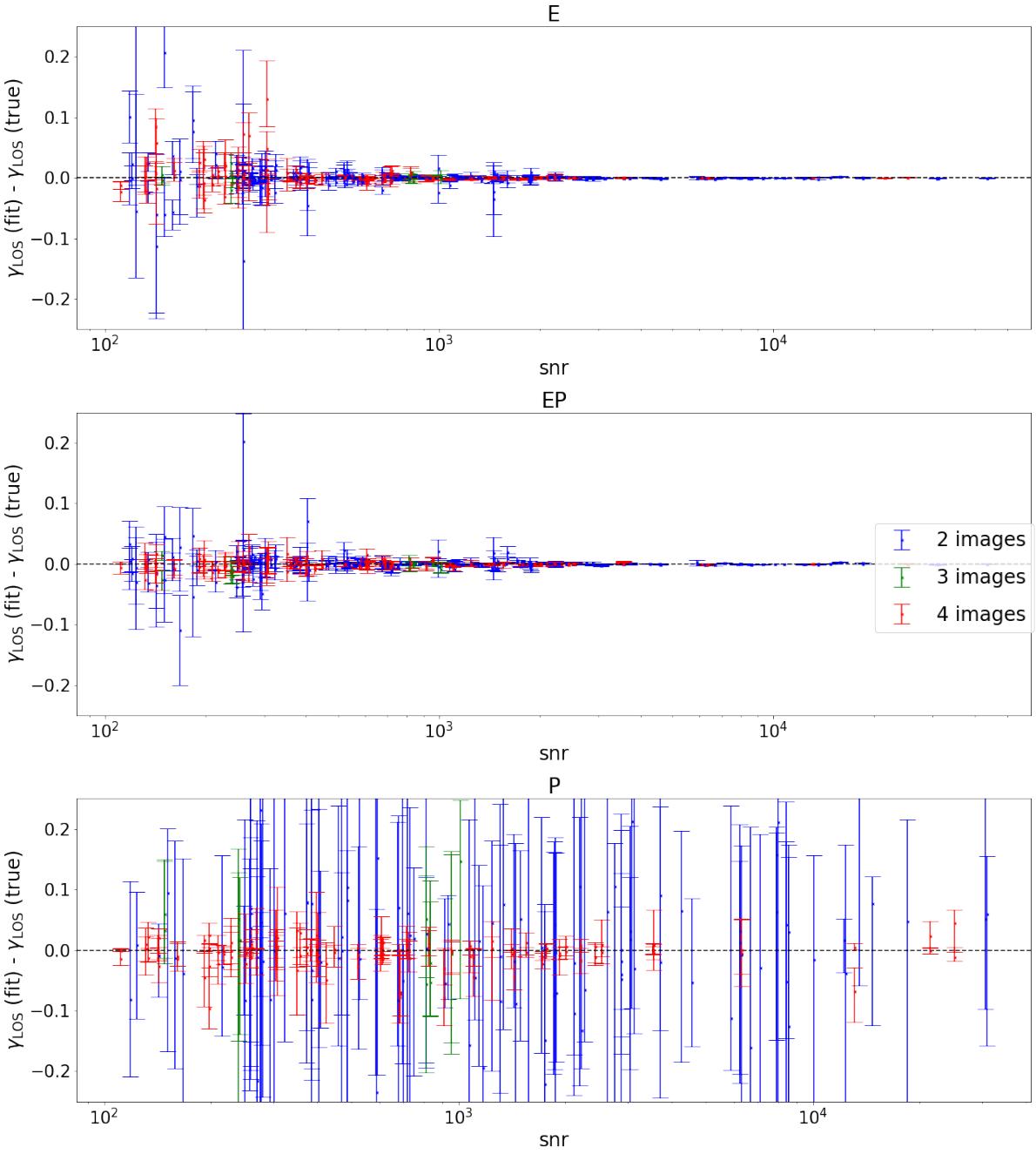

Figure 3.14: *The accuracy and precision of $\gamma_{\rm LOS}$ measurements for the E, EP and P fit types, as a function of snr, for sources placed at positions which result in a lens equation with either 2 (blue), 3 (green) or 4 (red) solutions.*

are still much poorer than for all higher-snr extended source or extended source + point source images. In this case, comparing the results in tables 3.4 and 3.5, it seems that these improvements are seen statistically as well as graphically, with $p_{\sigma<20\%}$ rising significantly when more point sources are present. Once again, these effects are seen consistently in the presence or absence of time delay data.

3.3 The significance of image noise and time delay measurement uncertainties

From the results presented in section 3.2, it is clear that, for simple mock images constructed with a typical range of source and lens parameters, assumed to have reasonably high values of signal-to-noise ratio and a typical spread of uncertainties in the relative time delays between point source images, simple, extended source images are those with which the line-of-sight shear can be measured most precisely. From plots such as those shown on figs. 3.13, B.1 and B.3, there is some indication that the precision of measurements of shears and ellipticities from poorer signal-to-noise ratio data are improved by the presence of point source images. Furthermore, as it is clear that the strongest determinant of the precision of a measurement for a given lensing configuration is the signal-to-noise ratio, the effects of the uncertainties in the relative time delay measurements are obfuscated, and it is therefore more challenging to comment on the importance of relative time delay measurements to constraining these shears. Therefore, to explore these relationships in more depth, we wanted to generate a series of identical images, which differed only in their signal-to-noise ratios and the uncertainties in their time delay data. By taking measurements of the line-of-sight shears from these images, we will be able to explore the effects of these variables on the tightness of constraints more effectively.

3.3.1 Generating the mock data and carrying out the fits

To generate the dataset needed for this particular line of inquiry, the cosmology and telescope parameters used were identical to those in section 3.2, and typical lens and source parameters were chosen to ensure that an extended source would produce a well-formed Einstein ring, and a point source would be lensed to give 4 images. These parameters include a source redshift of 1.5, lens redshift of 0.5, an elliptical power law main lens with ellipticity $(0.03, -0.05)$ and $\gamma_{\rm PL} = 2$, an Einstein radius of $\sim$ 0.7 found using a velocity dispersion of 210 km/s, and line-of-sight shears $\gamma_{\rm od} = (0.02, -0.07)$, $\gamma_{\rm os} = (-0.01, 0.03)$, and $\gamma_{\rm ds} = (0.05, -0.005)$, resulting in $\gamma_{\rm LOS} = (-0.04, -0.035)$.

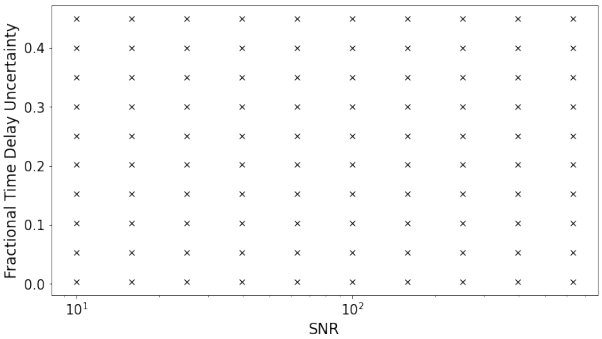

Figure 3.15: *The SNR and fractional uncertainty in the relative time delay values used when generating the 100 mock images in each category of image used in this section.*

For images containing an extended source, the extended source was given an ellipticity of $(-0.02, 0.01)$, a Sérsic index of 4, and a Sérsic radius of just below 4000 pc, which, at redshift 1.5, is equivalent to about 0.48 arcsec. This radius was chosen as a typical radius at that redshift, from the results presented in [70, 147]. For images containing a point source, this point source was given an amplitude of 200 times that of the integrated brightness of the extended source, located at Δ RA = 0.001 and Δ DEC = = -0.003 relative to the centre of the lens. The magnifications of the lensed point source images themselves were perturbed by a factor sampled from a normal distribution with a standard deviation of 0.01, to simulate the effects of phenomena such as microlensing. As in section 3.2, these lensing parameters were used to create three different classes of strong lens images - one in which only the extended source is present, one in which only the point source is present, and one in which both point source and extended source are present. For each of these classes, ten images were generated, differing only in the signal-to-noise ratios, which were varied between 10^1 and 10^3, distributed evenly in log space between the ten images. As before, the mean background surface brightness was calculated by dividing the surface brightness of the extended source at half light radius by the SNR value. Therefore, for each of the three classes of images, for a given SNR value the mean background was constant irrespective of the presence of the much brighter point source. The images for

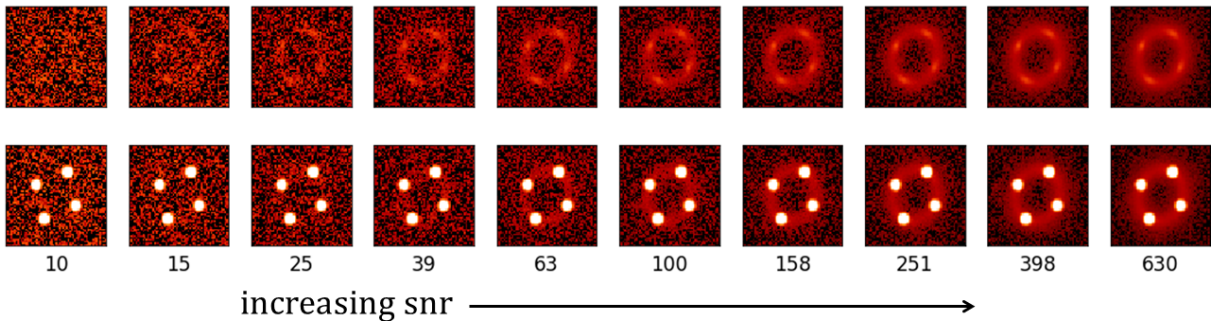

Figure 3.16: *The set of 10 images used to compare the effects of changing signal-to-noise ratios and the uncertainties in time delay data. The top row of plots show the images of lensed extended sources only, and the bottom row show the result of having a point source present at the centre of this extended source as well. Each image is identical, except for the SNR value used to generate the image from inputted lensing parameters, which are labelled. As before, these values are used to define the background count rate of the image relative to the counts at half-light radius of the extended source, before lensing.*

the cases where either only an extended source or both an extended and point source are present are shown in fig. 3.16.

The time delay of each of the four point source images were calculated, and the relative time delay measurements were generated by sampling from a normal distribution centred on the exact relative time delays, and with a standard deviation calculated as the relative time delay multiplied by some factor (the fractional uncertainty of the measurement). This fractional uncertainty was identical for each relative time delay measurement in a single image. This process was repeated ten times per signal-to-noise ratio value, such that, for each signal-to-noise ratio value, there were ten sets of relative time delay measurements, with uncertainties distributed evenly between 0.3 % and 50 %, which, as discussed in section 3.2, are values typical for the time delay measurements in surveys such as [111]. Given the dominant effect of the image itself over the time delay information, we once again did not consider time delay measurements less precise than those with an uncertainty of 50 %.

Using the steps described above, five sets of 100 mock datasets were created, corresponding to each of the possible combinations of signal-to-noise ratio values and relative time delay uncertainties, one set featuring images of a lensed extended source only, one set featuring images of a lensed point source only, one set featuring images of a lensed point source source and extended source, one set featuring images of a lensed point source and measured relative time delays, and one set featuring images of a lensed point source and extended source, and measured relative time delays. Figure 3.15 shows the SNR and fractional time delay uncertainties of each of the 100 mock images per set. As the first three cases lack time delay information, they should in fact be thought of as containing ten sets of ten identical images, each of the ten sets corresponding to a different SNR value. However, rather than simply using one image per SNR value in these cases, the MCMC sampling was repeated 10 times per SNR value, as this gave an indication of the spread in parameter constraints across multiple identical measurements, thus allowing for a more informed comparison between different lens configurations and fitting approaches.

Each of these mock images were then fitted as section 3.2, using the `LOS_MINIMAL` profile to describe the line-of-sight shears, running an MCMC chain with 10 000 steps. As before, this fit was run three times for each of the two images with time delays - one fit in which τ_{ds} was sampled, one fit in which it was fixed to the correct value, and one in which it was fixed with an incorrect value of H_0. From the results shown in the previous section, we would not expect any significant differences between these three fitting approaches. However, they are nonetheless included, firstly to validate this expectation, and, assuming that the expectation holds true, to get a measure of the variation across (effectively) identical measurements. The best guess at the correct value, and the associated 1-σ upper and lower bounds, were saved for $\gamma_{\rm LOS}$, $\gamma_{\rm od}$ and the lens ellipticity.

3.3.2 Line-of-sight shear measurements

Figure 3.17 shows the constraints on $\gamma_{\rm LOS}$ values in the form of a colour plot, shown as a function of the SNR and of the uncertainties in the measured relative time delays, for each of the data sets and fitting routines used. Lighter colours indicate tighter constraints, with $\sigma(\gamma_{\rm LOS})$ calculated as the quadratic mean of the symmetrised fractional 1σ uncertainties on the two $\gamma_{\rm LOS}$ parameters. When this value exceeds 0.2,

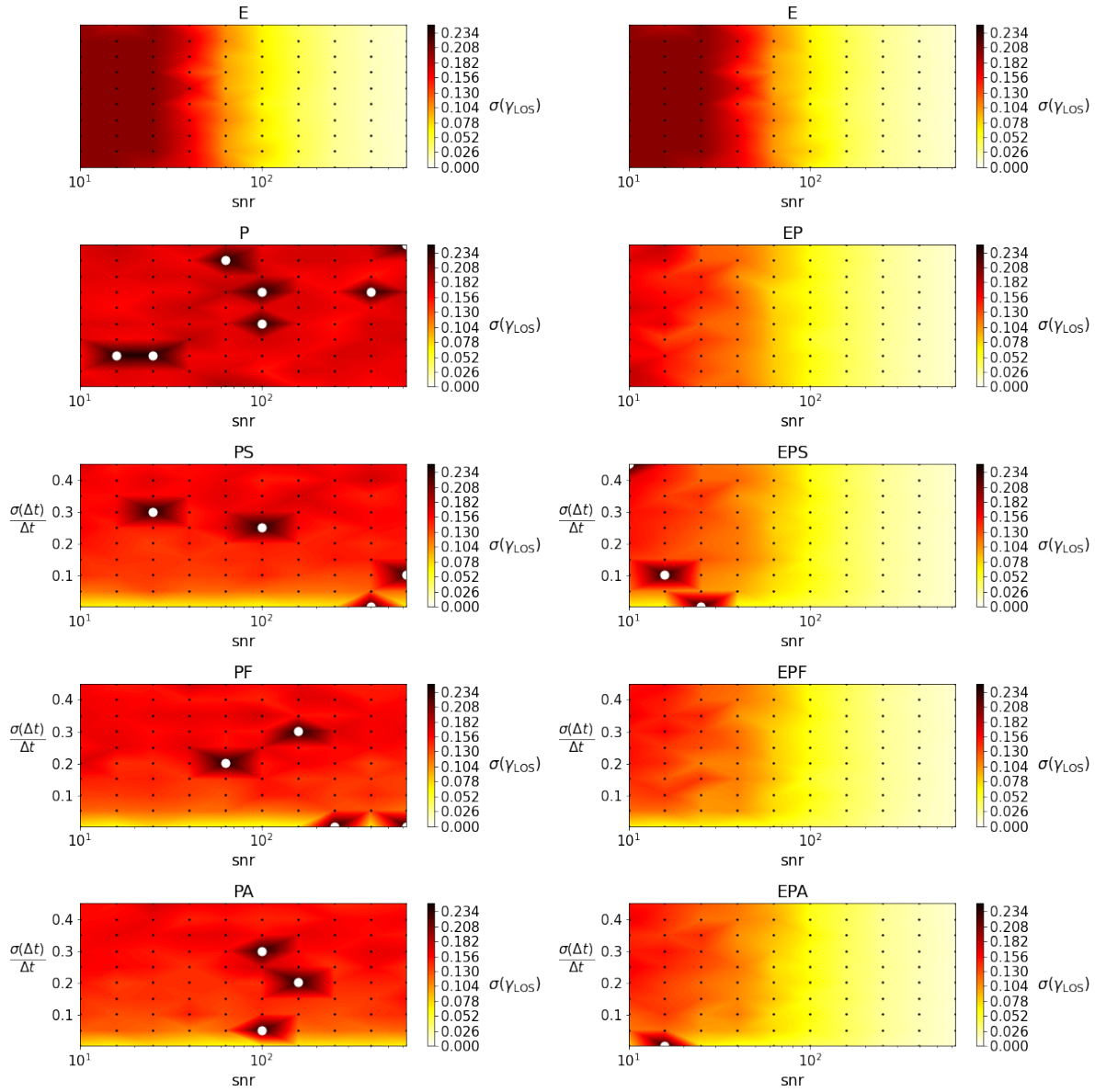

Figure 3.17: *Constraints on $\gamma_{\rm LOS}$, as a function of the SNR and relative time delay uncertainty, for each of the fit types under consideration. $\sigma(\gamma_{\rm LOS})$ is calculated as the quadratic mean of the fractional uncertainties on each of the components of $\gamma_{\rm LOS}$, and is plotted with darker shades representing poorer constraints. Any value which could not be constrained to within 0.2, is simply assigned a value of 0.2, for the continuity of the plot. If one of the $\gamma_{\rm LOS}$ parameters could not be constrained at all, it is assigned a value of 0.24, and a white circle is plotted. Note that, given the highly idealised mock data used in this section, the very tight constraints should not be taken as predictions for measurements from real datasets, but are rather intended to allow comparisons between the effect of SNR and the uncertainty in the time delays.*

the value is capped at 0.21, such that the plotted range allows for subtle differences at lower values of the relative uncertainty to still be clear. In the case where one or both of the $\gamma_{\rm LOS}$ parameters could not be constrained at all, a white point is plotted at that position on the plot, and the uncertainty is set to 0.24 (for the continuity of the colour plot).

Figure 3.18 shows the precision and accuracy of $\gamma_{\rm LOS}$ measurements as a classic error bar plot for each of the data sets used. On the left hand column, the plotted data corresponds to fits where the SNR was ~ 40, plotted as a function of the relative uncertainty in the measured time delay. In the cases where no time delay data are present, we would therefore expect the variation along the x-axis to give an indication of the variability of repeated identical measurements. On the right hand column, the plotted data corresponds to fits with the uncertainty in the relative time delay measurements was 0.3 %, i.e. the most precise time delay measurements considered in this section, plotted as a function of the SNR of the images used. This figure will allow us to examine subtleties in the effects of these variables which are not immediately apparent from fig. 3.17.

Extended Sources Only

The tightness of constraints when only an extended source is present is shown in the top plots on fig. 3.17 (plotted twice for ease of comparison down the column) and on fig. 3.18. As expected, the constraints show a clear dependency on the signal-to-noise ratio, improving steady as the level of noise decreases. We see some variability across the 10 fits of the same image, but without any significant outliers. In every case, even when the signal to noise ratio is very poor, both $\gamma_{\rm LOS}$ parameters could always be constrained.

Adding Point Sources

When a point source is present in the absence of an extended source, shown second from top left in figs. 3.17 and 3.18, three main features stand out – firstly, as discussed in the previous section, the SNR has no observable effect on the constraints. Secondly, we see that the constraints are, on average, poorer than the higher SNR extended source constraints, but better than the lower SNR extended source constraints. Quadruply-lensed point source images which dominate over the background level seem to give constraints comparable to complete Einstein rings with SNR values around 40, defined relative to the source light at half-light radius before lensing. This is illustrated clearly in fig. 3.18, where we see that, while the size of the error bars for $\gamma_{\rm LOS}$ measurements from an extended source only with SNR ~ 40 and from a point source only which dominates over the background level, the consistency of the measured parameters is much higher for the point source than for the extended source. In the former, the true value consistently lies near the centre of the error bars of the measured value, while in the latter, there is much more variability. This effect is seen when any of the low SNR extended source image measurements are compared with their point source image counterparts, and is consistent with the observation in section 3.2 that the mock images of quadruply-imaged point sources used in this analysis seem to lead to overestimated uncertainties of measured $\gamma_{\rm LOS}$ values. Finally, we see that, in 7 of the 100 cases, one or both of the parameters could not be constrained at all.

When both a point source and an extended source are present, as shown in the second from top right in figs. 3.17 and 3.18, we see a similar gradient as in the case of an extended source only, except that the constraints remain reasonable even when the extended source is barely distinguishable above the noise. For higher SNR values, constraints are just about identical to when the point sources are absent, but for lower signal to noise ratio values, the improvements over the extended source only case become more and more apparent. Thus, as inferred from the data presented in section 3.2, it seems that the presence of point sources is only advantageous when attempting to constrain $\gamma_{\rm LOS}$ from images in which the signal from the Einstein ring does not dominate over the background signal.

Adding Time Delays

When relative time delay measurements are included in the fitted mock data, the tightness of the constraints becomes a function of two variables, and this is plotted as a colour map in the bottom three rows of fig. 3.17.

The bottom three rows of the left-hand column in fig. 3.17 corresponds to the case where point sources and time delay measurements are present and extended sources absent. In each of these, we firstly notice that, once again, there are cases in which one or both of the $\gamma_{\rm LOS}$ parameters could not be constrained, and that this seems uncorrelated with the SNR or $\sigma(\Delta t)$, though this appears to occur less frequently

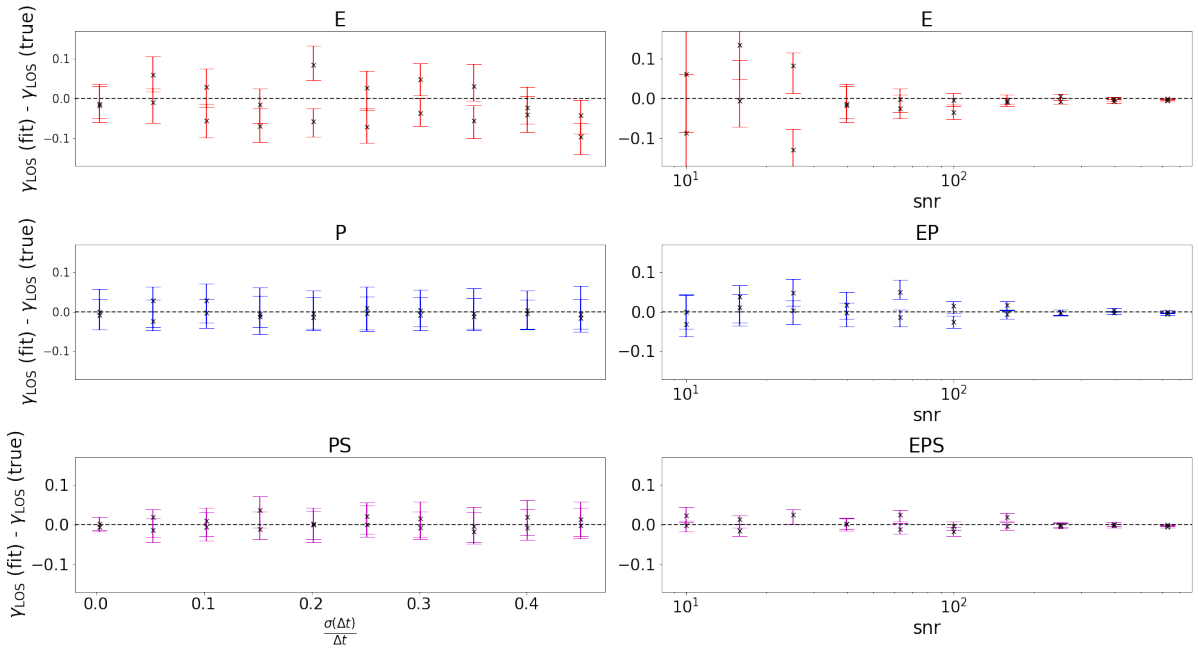

Figure 3.18: *Precision and accuracy of γ_{LOS} measurements for a selection of fit types from the dataset described in section 3.3. In the left column, measurements correspond to images with the lowest-considered SNR of 30, and are plotted as a function of the fractional measured time delay uncertainty (which will not affect the E and P cases). In the right column, measurements correspond to images with the lowest-considered fractional uncertainty in the measured time delays (0.3 %), and are plotted as a function of the snr.*

than in the case where no time delay measurements are included (in agreement with the results presented in section 3.2). Most notable, however, is that, when measured time delays with uncertainties smaller than around 10 % are included, there is an easily-noticeable improvement to the constraints on γ_{LOS} values. This is also seen in fig. 3.18, where we see the errors in γ_{LOS} being more than halved by the most precisely measured time delay data considered in this section.

The bottom three rows of the right-hand column in fig. 3.17 correspond to the case where point sources, extended sources and time delay measurements are all present. Here, we firstly see that there are low SNR cases where γ_{LOS} parameters could not be constrained, and this seems to be uncorrelated with the uncertainty on the time delays. Secondly, we see once again that the presence of precise time delay measurements improve the constraints on γ_{LOS} for images where the SNR is low. This is also seen in fig. 3.18.

3.4 Is the post-Born correction necessary when measuring shear?

In many traditional parameterisations of the line-of-sight shear, γ_{od} is omitted from the model, and the effects of line-of-sight perturbers are simply encoded in a γ_{ext} term which takes the place of γ_{LOS} in eq. (1.265) [148, 22, 176, 177]. In these examples, the external shear term is simply meant to account for the various perturbative effects of substructures and matter along the line-of-sight, and is treated as a nuisance parameter.

In this section, we want to consider whether this simplifying assumption affects the measurement of γ_{LOS}. By comparing models in which γ_{od} is included to those in which it is omitted, we are able to investigate whether any advantages are offered by the full minimal lens model over the traditional external shear parameterisation, and thus whether the degeneracy discussed in section 1.3.4 is enough to absorb the effects of the foreground shear into the main lens model, when this main lens is modelled with a simple EPL profile. This question is very similar to that posed in section 3 of [77]. However, here, we adopt a simpler EPL model for the main lens, as is done in many surveys, and additionally consider cases where point sources and time delays are present. Furthermore, we will also be able to investigate the effects of this assumption on attempted measurements of the main lens' ellipticity, as an illustration of the degeneracy discussed in section 1.3.4.

3.4.1 Line-of-Sight Shears

ignoring the post-Born correction in $\gamma_{\rm LOS}$ measurements					
		E		**EPS**	
		PB included	PB ignored	PB included	PB ignored
constraints	$p_{\mathcal{C}}$	99 %	94 %	100 %	94 %
precision	$p_{\sigma<5\%}$	42 %	36 %	43 %	40 %
	$p_{\sigma<20\%}$	71 %	66 %	74 %	70 %
accuracy	$p_{<\sigma}$	64 %	42 %	66 %	42 %
	$p_{>2\sigma}$	11 %	32 %	11 %	31 %
	χ^2	4.3	775	4.4	2888

Table 3.6: *The results of measuring $\gamma_{\rm LOS}$ using extended sources only (E) and extended sources, point sources and time delay measurements, with τ_{ds} sampled (EPS), comparing the case where the post-Born correction is included in the model, vs that in which it is ignored for measurements of $\gamma_{\rm LOS}$.*

Table 3.6 compares the results of measuring $\gamma_{\rm LOS}$ using a model with and without the γ_{od} term, for the E and EPS fit types. These were chosen following the results of sections 3.2 and 3.3 as the image types and approaches to fitting which would be the most likely to result in tightly-constrained and reliable estimates of $\gamma_{\rm LOS}$. From these results, what is immediately clear is that the biggest influence of including the foreground shear is on the accuracy of the results. While the ability of the code to constrain parameters, as well as the tightness of these constraints, is somewhat improved when γ_{od} is present, it is very clear from the $p_{<\sigma}$, $p_{>2\sigma}$ and χ^2 values that these bounds cannot be trusted in the absence of γ_{od}. This can also be seen in figs. 3.19 and 3.20, where these measurements are plotted as functions of the values of γ_{od} used in generating the models and the SNR respectively. In both plots, it is clear that a large number of outliers are present when the post-Born correction is omitted. Notably, fairly extreme outliers are present even when γ_{od} is as small as $|\gamma_{od}| \sim 0.05$. Furthermore, high quality images do little to alleviate this effect, with very extreme outliers seen even when the SNR is $\sim 10^4$.

These results agree qualitatively with those in [77], in which accuracy of line-of-sight shear measurements were seen when omitting the foreground shear, with no changes to the precision. The differences seen were more marginal than those present in this work, which likely finds its explanation in the fact that the lens model used in [77], when the foreground shear was omitted, was a combination of a Sérsic ellipse and an NFW halo. This richer model is better able to absorb the effects of the foreground shear than the EPL lens, and so, consistent with the conclusions in [77], it seems that omitting the foreground shear is only viable when the main lens model is sufficiently complex, and even relatively small values of γ_{od} are enough to introduce departures from the first-order degeneracy with e which make measurements of $\gamma_{\rm LOS}$ unreliable.

3.4.2 Ellipticities

In section 1.3.4, we saw that the ellipticity is, to first order, degenerate with the foreground shear. If γ_{od} is omitted from a model, we would expect the fit to compensate for this via the ellipticity. From the expression in eq. (1.294), we would expect the measured components e_i (fit), $i \in 1, 2$ to be related to the true ellipticity e_i via

$$e_i \text{ (fit)} = e_i + \tilde{\gamma}_{\rm PL}\gamma_i^{od}, \tag{3.20}$$

where

$$\tilde{\gamma}_{\rm PL} = (3 - \gamma_{\rm PL})\left(2 - \gamma_{\rm PL} + \frac{\gamma_{\rm PL} - 1}{5 - \gamma_{\rm PL}}\right)^{-1}. \tag{3.21}$$

Figures 3.21 and 3.22 show the results of measuring the ellipticity of the lens images generated in section 3.2 with the post-Born correction included in the model (i.e. sampling γ_{od}), and with the post-Born correction excluded (fixing γ_{od} to zero). The difference between e_i (measured) and e_i (true) is plotted on the y-axis (including errorbars), and $\tilde{\gamma}_{\rm PL}\gamma_{od}$ on the x-axis. The expected degeneracy is also plotted.

From these figures, we see that the measured ellipticities follow the line of the predicted degeneracy when the post-Born correction is excluded. This relationship is tightest when γ_{od}^i is small, but where γ_{od}^i

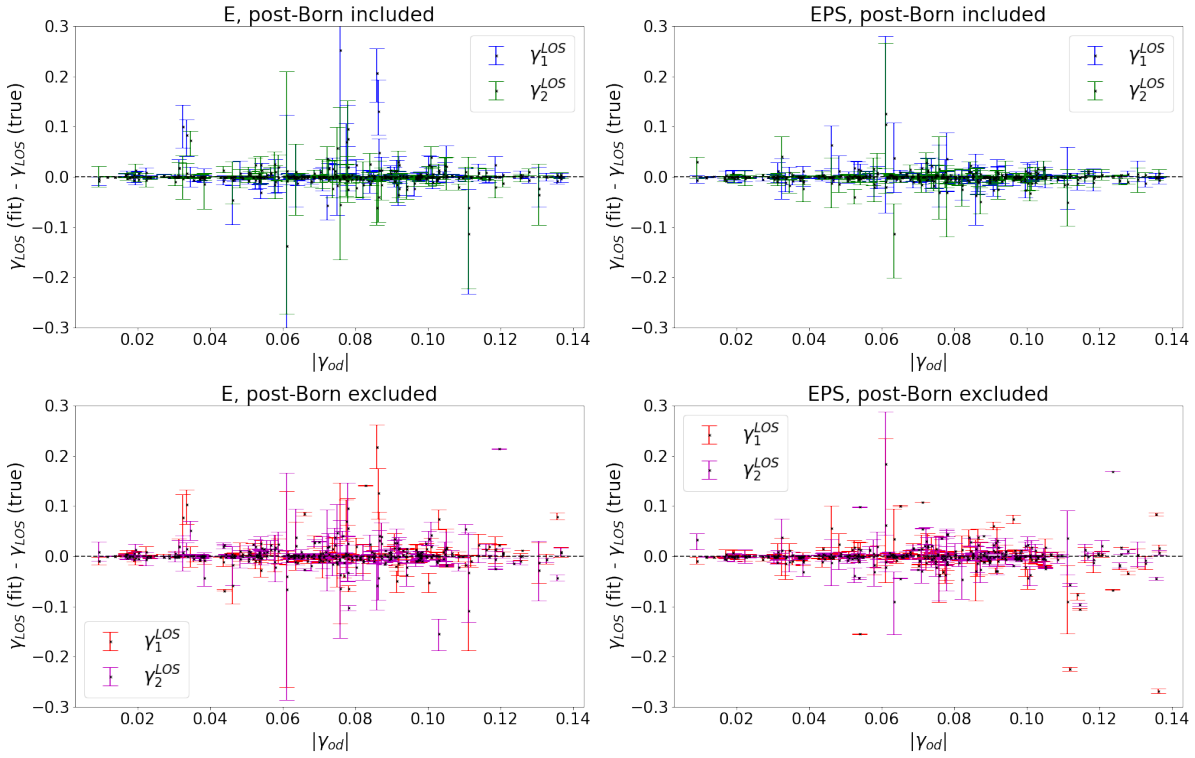

Figure 3.19: *Measurements of γ_{LOS} from images with a lensed extended source only (E), and images with a lensed point source and extended source, as well as time delay data, with τ_{ds} sampled (EPS), comparing the cases where γ_{od} is included (top) vs excluded (bottom) in the model, plotted as a function of the quadratic mean of the components of γ_{od} used when generating the image.*

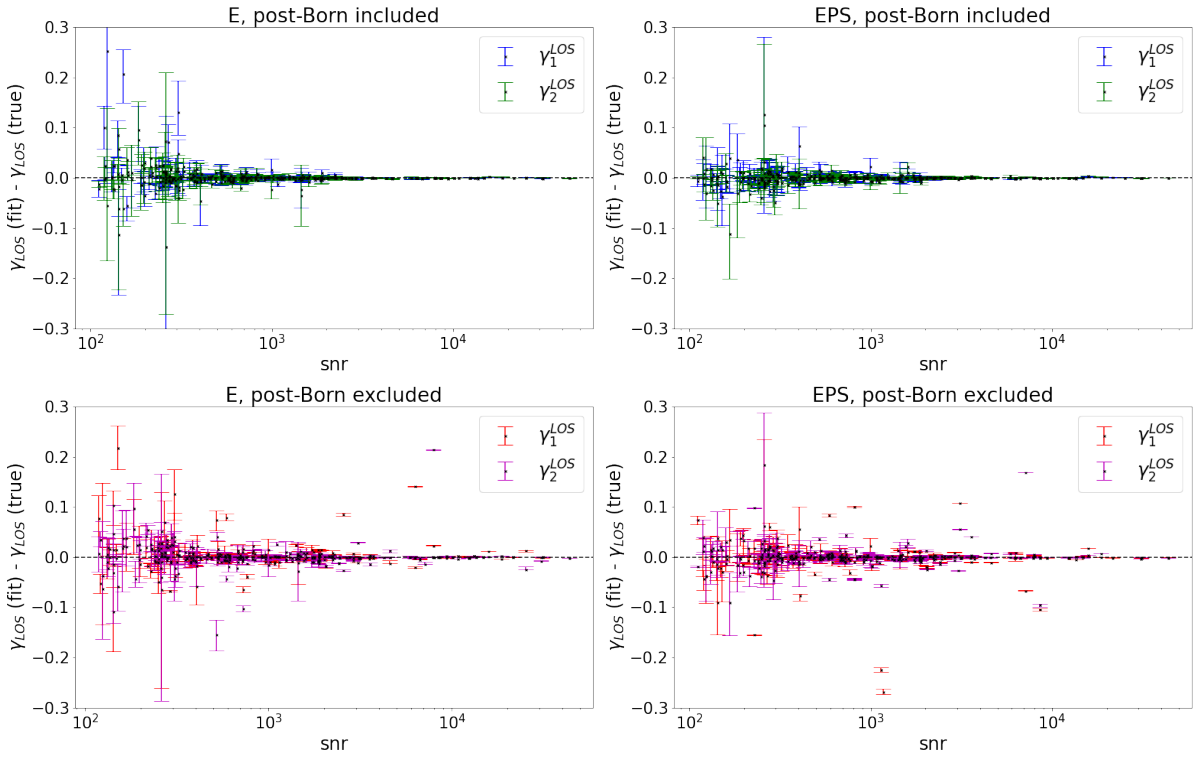

Figure 3.20: *Measurements of γ_{LOS} from images with a lensed extended source only (E), and images with a lensed point source and extended source, as well as time delay data, with τ_{ds} sampled (EPS), comparing the cases where γ_{od} is included (top) vs excluded (bottom) in the model, plotted as a function of the snr.*

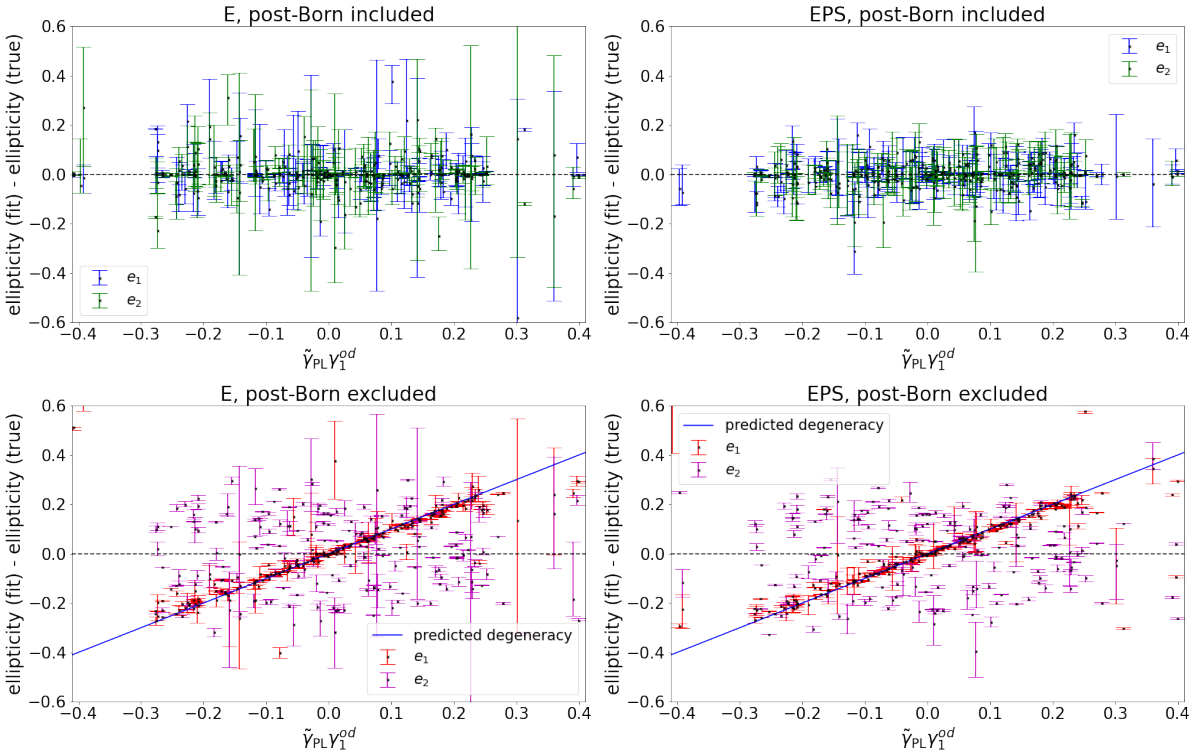

Figure 3.21: *The results of measuring the ellipticity from images with a lensed extended source only (E), and images with a lensed point source and extended source, as well as time delay data, with τ_{ds} sampled (EPS), comparing the cases where γ_{od} is included (top) vs excluded (bottom) in the model. In latter case, the predicted degeneracy between e_1 and γ_1^{od} is plotted as a blue line.*

is larger than around 0.15, the degeneracy is somewhat curtailed. This is what we would expect, given that the degeneracy is first-order only for an EPL lens model.

When the post-Born correction is included, we see that there are no systematic biases introduced when measuring the ellipticity, although the constraints on e_1 and e_2 are unsurprisingly weak. Both figures once again illustrate the advantages offered by images featuring lensed point sources seen in appendix B, in which the worst constrained and most outlying measurements are suppressed.

3.5 Discussion

From the results presented in this chapter, we can identify a number of patterns which emerge consistently from the data. It seems that the presence of an extended source in an image leads to dramatic improvement in obtaining tight constraints on the line-of-sight shears, provided that source can be clearly distinguished from the background noise. For these high signal-to-noise ratio images, it seems that the presence of a point source can in fact increase the uncertainty in $\gamma_{\rm LOS}$ measurements. This substantiates the approach in [23, 59, 77, 101], in which Einstein rings were taken to be the most promising sources of external shear measurements from strong lensing. However, when the signal-to-noise ratio is poorer, the presence of clearly visible point source images with the fainter extended source image leads to better constraints, and can mitigate inaccuracies in the reported results. Point sources images without a visible extended component, even when four images are present, provide only weak constraints on $\gamma_{\rm LOS}$, but these uncertainties seem often to be overestimated. This could, however, arise from the non-blind initialisation of the MCMC fitting.

In every case, the inclusion of time delay data improved the precision and accuracy of shear measurements compared to the cases where they were excluded from the same image, particularly for low SNR images, but this improvement was small, and not significant enough to offer improvements over the corresponding extended source only image, on average. No advantages are seen when fixing the time delay distance, rather than leaving it as a free parameter. It seems that the main advantage offered by time delays is as an additional test of the appropriateness of a model, mitigating outlying measurements and

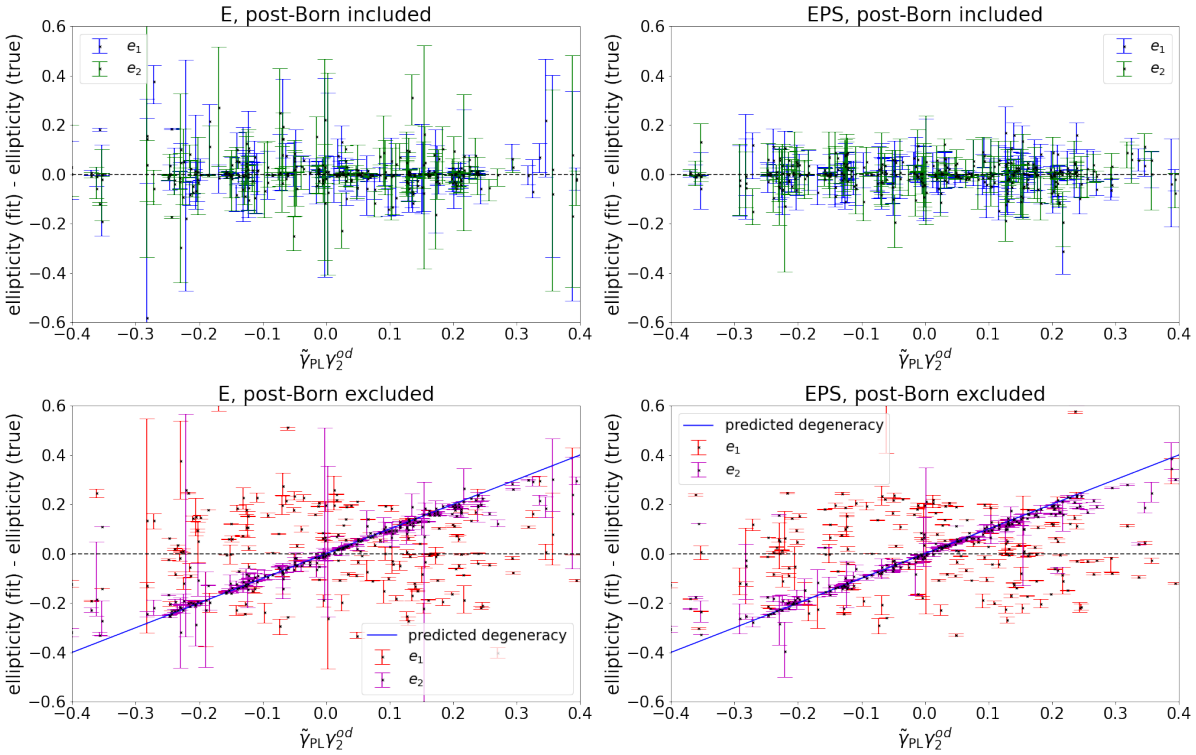

Figure 3.22: *The results of measuring the ellipticity from images with a lensed extended source only (E), and images with a lensed point source and extended source, as well as time delay data, with τ_{ds} sampled (EPS), comparing the cases where γ_{od} is included (top) vs excluded (bottom) in the model. In latter case, the predicted degeneracy between e_2 and γ_2^{od} is plotted as a blue line.*

adding a degree of confidence in measurements from images containing less information. These findings are consistent with those in [83] (whose results applied to reconstructing lenses, rather than measuring the shear specifically).

The jury remains out as to how confidently line-of-sight shear measurements can be linked to fundamental cosmology. Following the discussion in section 3.2.1, strong lensing images are often located in regions of the sky with overdensities of matter along the line-of-sight. While this means that these effects are easier to measure in strong lensing images, it also means that these measurements cannot as easily be generalised to the large-scale matter distribution along truly random lines of sight. The field remains somewhat divided between those advocating for the use of high quality lensing images to identify individual substructures, and those hoping to process of large samples of strong lenses for less fine-tuned but more stastistically meaningful measurements of the shear. For a study aiming to obtain very precise shear measurements from a small sample of sources, it seems that the best candidate images are those featuring well-resolved, high signal-to-noise ratio Einstein rings, in the absence of any point sources. However, for larger studies which may look to the more numerous population of strongly-lensed quasar systems to supplement findings from Einstein rings, it seems that time delay data should be included wherever possible. While their value in shear measurements alone does not seem to justify the time and cost required to obtain them, this data may, in many cases, be available from time delay cosmography studies, and in those cases, can provide a valuable additional check on results.

While much less frequent, lensed supernovae may provide a valuable opportunity for shear measurements. The last decade has seen the first discoveries of strongly-lensed supernovae [91, 129, 65]. In [91], for example, the host galaxy is also strongly lensed. Because of the comparative transience of supernovae, it may therefore be possible to directly compare shear measurements with and without the presence of the point source, thus potentially reaping the benefits of both extended sources in isolation and point sources with time delays for the same lens system. Additionally, when the lensed supernova is type Ia, as in [129], its absolute magnitude can be ascertained, meaning that the absolute, rather than relative, magnifications can be included as constraints, providing additional checks on the mass model.

Another topic of discussion within the literature has been the importance of the foreground shear.

Though largely ignored from models, it has been argued in [88] and [23] that the post-Born correction can significantly improve the quality of fits. [175] found that comparisons between early external shear measurements from simple lens + external shear parameterisations and actual lens environments found that this shear disagreed both in magnitude and direction with what would be expected. The conclusion drawn from this was the external shear was compensating for simplifications in the main lens model, but it may in fact have also arisen from simplifications in the shear modelling as well. This would be consistent with our findings in section 3.4, in which large systematic errors could be seen in $\gamma_{\rm LOS}$ measurements for not-unrealistic values of the foreground shear. It seems that the degeneracy between the main lens ellipticity and foreground shear is not strong enough to prevent a loss of accuracy when modeling strong lens systems with elliptical power law lenses and external shear, even when all additional complexity is removed from the main lens in the mock data.

Despite these findings, the work of [59] (discussed in section 1.3.4) convincingly demonstrates that any non-degeneracy between e and γ_{od} is simply model-dependent, and a sufficiently rich main lens model could fully absorb the effects of the foreground shear. When using more complex, two-component lens models to model similarly complex mock lenses with foreground and line-of-sight shear present, [77] found no disadvantage when omitting the foreground shear from the model, but a significant loss of accuracy when oversimplifying the main lens model. A sensible interpretation of our results therefore seems to be that the foreground shear can offer advantages in eliminating biased results when included as a free parameter in a simple lens model. However, its role should be understood to be as a nuisance parameter to add a degree of freedom to the lens, rather than a measured quantity, as it is fundamentally linked to the intrinsic ellipticity of the lens.

Chapter 4

Line-of-Sight Effects on Time Delay Measurements and the Hubble Constant

Any discussion of line-of-sight effects on strong lensing time delays would be incomplete without considering how these effects influence measurements of H_0. We have seen in chapter 3 that even a simple power law lens model in the presence of foreground shears of realistic strengths can lead to strong biases in the measurement of $\gamma_{\rm LOS}$ when fitting with a power lens and external shear only. The power law assumption is now understood to introduce similar biases into H_0 measurements via its inability to capture the full complexity of the main lens model. In some cases, the power law plus external shear model is supplemented with an external convergence, constrained with velocity dispersion data, itself a potentially biased approach, as discussed in section 1.3.4. More recent studies have used more complex, two-component lens models in the presence of external tidal corrections constrained via imaging and velocity dispersion measurements, to better account for deviations from a power law in the main lens model.

Giving a constructive answer to the question of the relative importance of uncertainties introduced by the line-of-sight shear versus those attributed to the mass-sheet degeneracy when measuring H_0 is not straightforward. There is little question that the MSD represents the biggest challenge to constraining H_0 with gravitational lensing time delays ([51, 66, 133, 178, 138]), and, as such, has been the focus of many recent studies ([62, 19, 181, 151]). However, it is conceivable that, as a better and better theoretical understanding of the problem is achieved, and as the observation and modelling of velocity dispersion data improves, a naive handling of the external shear may either dominate the uncertainty budget, or lead to systematic biases or underestimations of error variances in reported H_0 measurements. In this chapter, we will attempt to contribute towards answering this question, by asking whether the power law plus external shear assumption can account for the effects of foreground shear via the degeneracy discussed in section 1.3.4, or whether new biases are introduced by this shear, even when no additional complexity is present in the main lens and the line-of-sight convergence is perfectly constrained. In section 4.1, we investigate the effect of this shear on modelled time delays, and in section 4.2, we model mock lenses and their time delays with and without the post-Born correction included in the lens model, and compare results.

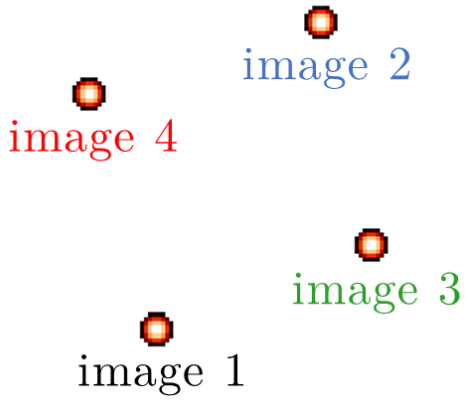

Figure 4.1: *Image labelling convention with a quadruply-imaged point source. Time delays are relative to image 1.*

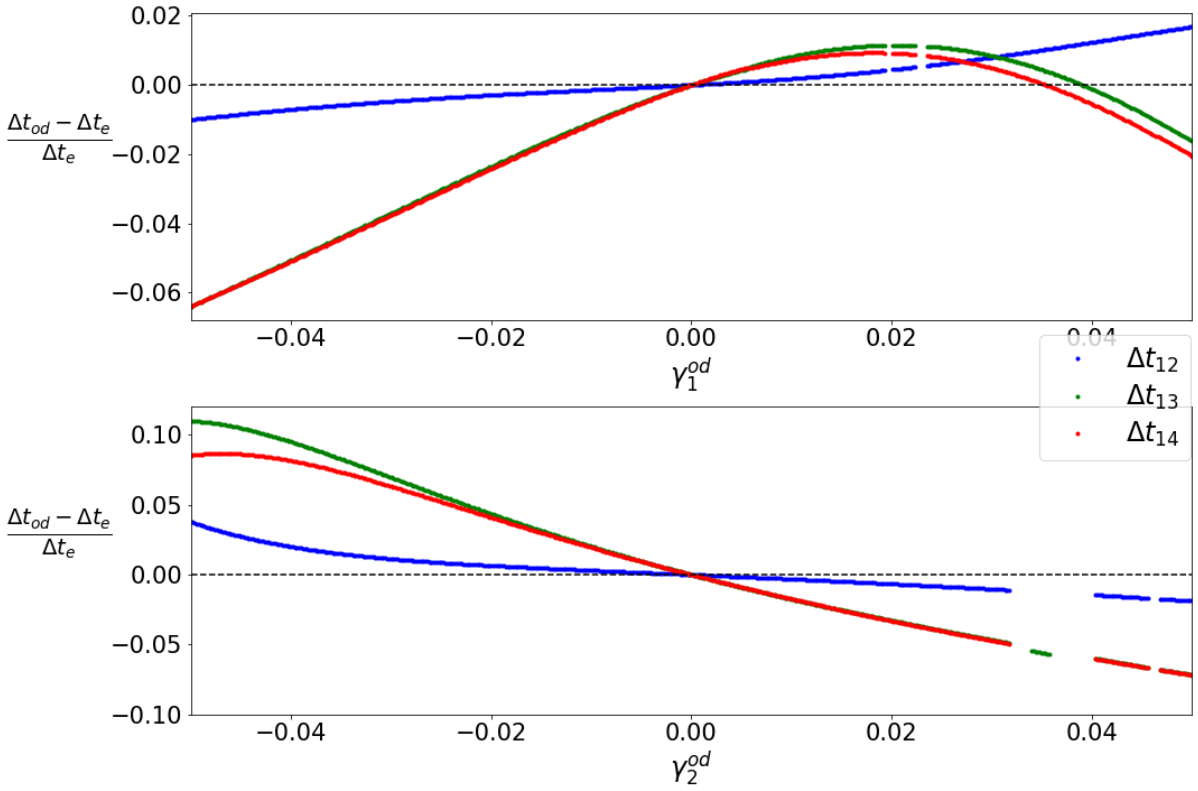

Figure 4.2: *The fractional difference in relative time delays between an elliptical power law lens with a foreground shear term included (Δt_{od}), and one with an ellipticity modified according to eqs. (3.20) and (3.21) (Δt_e), which is therefore degenerate with the first lens to linear order. These differences are plotted for the relative time delays between the first image and each of the three others, labelled according to the conventions illustrated in fig. 4.1. In the top plot, the γ_1^{od} component is assigned various values, while γ_2^{od} is fixed to 0. In the bottom plot, γ_1^{od} is fixed to zero, while γ_2^{od} is varied. For certain specific combinations of parameters, the number of solutions to the lens equation drops from 4 to 3 or 2. Here, the Δt_{13} and Δt_{14} terms are no longer applicable, and Δt_{12} can change dramatically, leading to the gaps visible in the plotted relationships.*

4.1 The importance of shear terms within the Hubble constant uncertainty budget

While the effects of the line-of-sight shear described by the $\gamma_{\rm LOS}$ term is itself a source of uncertainty within measurements of H_0, this is accounted for via the external shear correction which is typically included in studies. What is often excluded, however, is the foreground shear term, whose influence cannot be absorbed into a $\gamma_{\rm ext}$ correction. As discussed in prior sections, this choice is justified by the degeneracy between γ_{od} and the main lens profile. However, the extent to which this degeneracy is complete depends on the choice of profile, and simple models may fail to account for the effects of γ_{od} on imaging observables, relative time delay measurements, and hence the measured value of H_0. As an illustration of the effects of omitting the foreground shear from a model of strong lensing time delays, we define an elliptical power law lens with $\gamma_{\rm PL} = 1.6$, $\theta_{\rm E} = 2$, $e_1 = -0.06$ and $e_2 = 0.07$. To this, the `LOS_MINIMAL` profile is added, with $\gamma_1^{\rm LOS} = 0.02$ and $\gamma_2^{\rm LOS} = -0.08$. In a first model, a foreground shear γ_{od} is added. In a second, the foreground shear is set to zero, but the ellipticity of the ellipticial power law profile is modified according to eqs. (3.20) and (3.21).

Figure 4.2 shows the fractional change in relative time delays between these two models as a function of the inputted value of γ_{od} parameters. While the specific relationship between these variables will differ between different lenses, it is clear from these plots that foreground shears less than 5% can result in measured relative time delays which differ by more than 5% from those which would be predicted by a model of the main lens as an elliptical power law which fails to include γ_{od}, but rather absorbs its effects into the ellipticity of the lens profile. This difference is larger than the uncertainties reported in several

recent strong lens time delay measurements (e.g. [27, 111, 102, 154, 37, 148]), and thus could be an important contributor to the uncertainty budget in H_0 measurements when the main lens model is not sufficiently rich to render the degeneracy between its ellipticity and the foreground shear exact.

4.2 Is the post-Born correction necessary when measuring the Hubble constant?

ignoring the post-Born correction in τ_{ds} measurements					
		EPS		**PS**	
		PB included	PB ignored	PB included	PB ignored
constraints	$p_{\mathcal{C}}$	99 %	95 %	85 %	82 %
precision	$p_{\sigma<5\%}$	16 %	13 %	0 %	2 %
	$p_{\sigma<20\%}$	81 %	74 %	26 %	14 %
accuracy	$p_{<\sigma}$	71 %	57 %	80 %	66 %
	$p_{>2\sigma}$	3 %	16 %	0 %	5 %
	χ^2	0.85	662	0.25	2414

Table 4.1: *The results of measuring τ_{ds} when extended sources are present (EPS) or absent (PS), comparing the case where the post-Born correction is included in the model, versus that in which it is omitted.*

From the results of chapter 3 and the discussion above, it is of interest to consider the effects of excluding the foreground shear when attempting to measure the Hubble constant with an elliptical power law main-lens profile. To this end, we follow a similar approach to that adopted in section 3.4, making use of the dataset of lenses created in section 3.2 to obtain measurements of the time delay distance τ_{ds}, and hence H_0. As this procedure requires relative time delay measurements (and hence point sources), and requires τ_{ds} to be sampled, only the EPS and PS fit types are considered (as defined in table 3.1). As in section 3.4, each fit is run twice: once with γ_{od} sampled (including the post-Born correction), and once where it is fixed to zero (ignoring the post-Born correction). We once again choose to fix the line-of-sight convergence terms in both the mock data and the fit to zero. While this assumption cannot be made when analysing real data, it does allow the specific effects of the foreground shear to be more clearly ascertained. We do not expect the constraints on H_0 achieved with this mock data to be comparable with those measured from real data, but there are nonetheless insights to be gained in a comparative sense between the different approaches to the fit.

Table 4.1 shows the results of these fits. In both the EPS and PS cases, including γ_{od} in the fit improves the probability of τ_{ds} being constrained, though the improvement is small. Improvements to the precision are also seen in both cases. However, it is the accuracy of these constraints which is most affected. In the case where an extended source is present and γ_{od} is included in the model, $p_{<\sigma}$, $p_{>2\sigma}$ and χ^2 all indicate that these constraints are well-defined, and the set of measurements is close to the expected distribution around the true values. However, when γ_{od} is omitted from the model, the number of outliers is triple the expected value, $p_{<\sigma}$ is lower than expected, and the massive χ^2 value indicates very poor fits. This effect can be very clearly seen in fig. 4.3, comparing the two plots on the right. When the post-Born correction is omitted, many outliers are clearly visible, an effect which does not seem to be mitigated by having 4 point source images present. In some cases, very tight constraints are τ_{ds} are reported, despite the measured value lying far from the true value. Outliers are more common when the foreground shear is stronger, but are seen even at small values of γ_{od}.

When measuring τ_{ds} from images featuring point sources only, constraints are, on average, substantially tighter when γ_{od} is included in the model, as can be seen from both the $p_{\sigma<20\%}$ parameter in table 4.1 and the error bars plotted in the left hand plots in fig. 4.3. This effect is not limited to lenses in which the foreground shear is large, and is seen regardless of the number of point source images present. Furthermore, including the post-Born correction seems to pull τ_{ds} measurements tightly towards the inputted value, removing almost all outliers. This effect is seen in every measure of accuracy shown in table 4.1, which indicate that errors are being overestimated when the post-Born correction is included, and in fig. 4.3, in which the reported values and their errors lie tightly along the line of the true value.

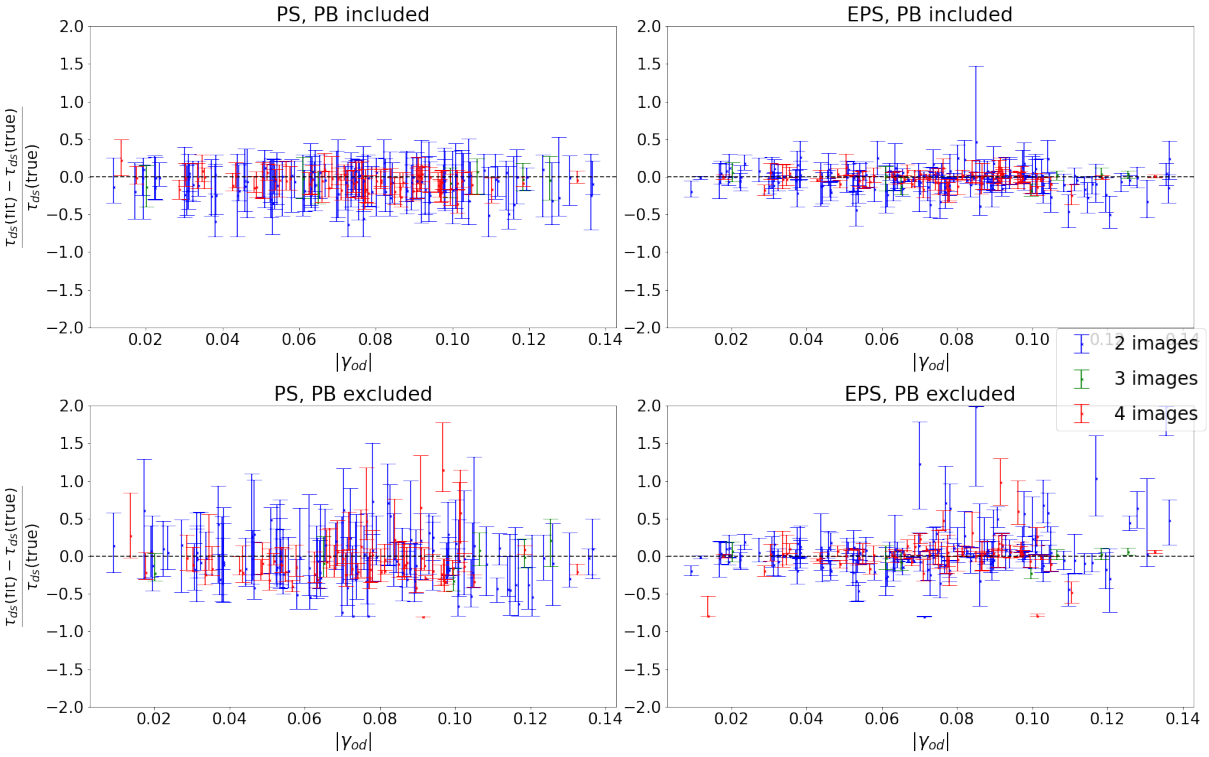

Figure 4.3: *Measurements of the time delay distance τ_{ds} from images with (EPS, right) and without (PS, left) an extended source present, comparing cases where the post-Born correction is included (top) or excluded (bottom). The number of point source images present is indicated by the colour of the error bars. On the x-axis, the quadratic sum of the components of γ_{od} used when generating the image are shown.*

When the post-Born correction is excluded, there is a greater scatter in the measurements, and while $p_{<\sigma}$ and $p_{>2\sigma}$ indicate very well defined error bars for the dataset as a whole, the very large χ^2 value suggests that, at least in a few cases, these errors may be underestimated.

The corresponding effects on measurements of H_0, obtained from τ_{ds} measurements and the source and lens redshifts using `lenstronomy`'s `ddt2h0` function, are plotted in fig. 4.4. Figure 4.5 shows only the measurements obtained when 4 point source images are present, plotted as a function of the uncertainty in the measured time delays. It is clear from these figures and the results discussed above that including γ_{od} when modelling lensed point sources only can somewhat reduce the scatter in measured H_0 values, but also that this effect is small. However, when an extended source is also present, failing to account for γ_{od} can lead to strongly biased measurements of H_0, and reported errors are frequently greatly underestimated.

What is also clear from these results is that the presence of an extended source and source-lens alignments leading to quadruple-imaging are those which provide the best chance of measuring H_0 accurately and precisely. While the links between these results and measurements of H_0 from real data are rudimentary at best, this does at least serve as a confirmation that, despite the apparent degeneracy between e and γ_{od} for elliptical power law lens models, failing to include the post-Born correction when measuring H_0, in particular when the main lens model is relatively simple, can lead to substantial biases in the reported results, even when γ_{od} is on the order of a percent or two.

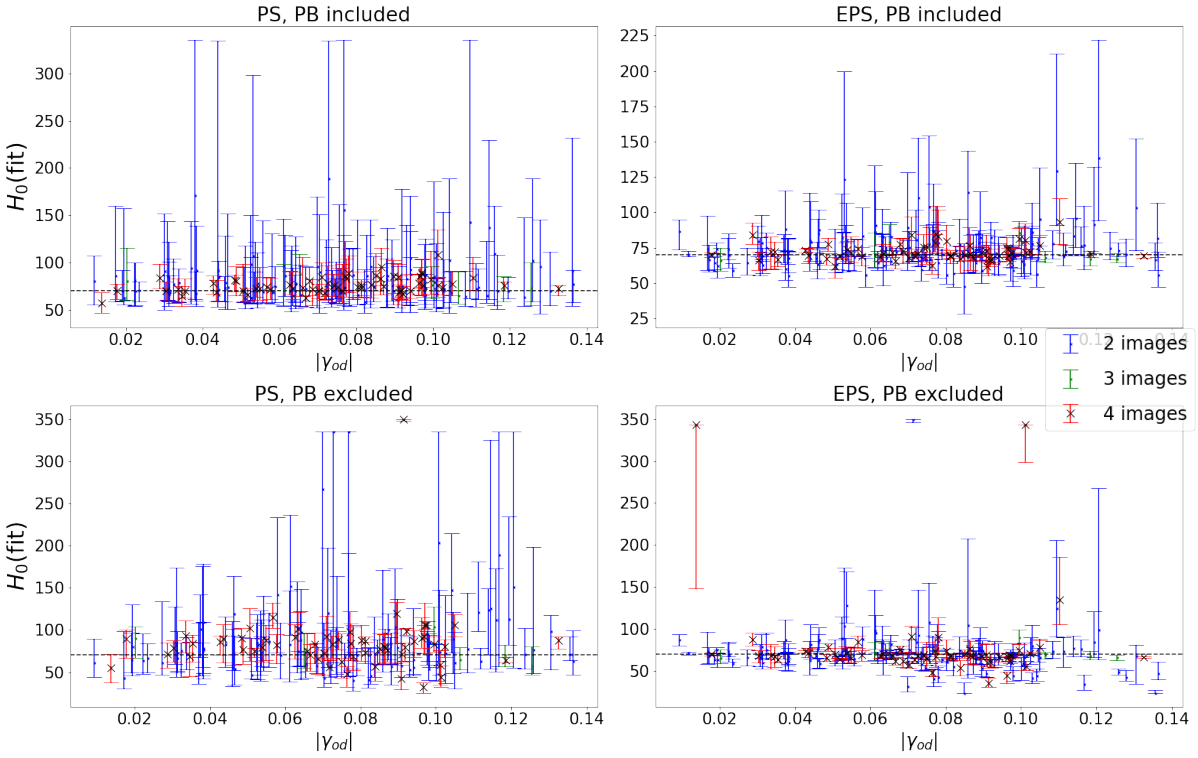

Figure 4.4: *Measurements of the Hubble constant H_0 from images with (EPS, right) and without (PS, left) an extended source present, comparing cases where the post-Born correction is included (top) or excluded (bottom). The number of point source images present is indicated by the colour of the error bars. On the x-axis, the quadratic sum of the components of γ_{od} used when generating the image are shown.*

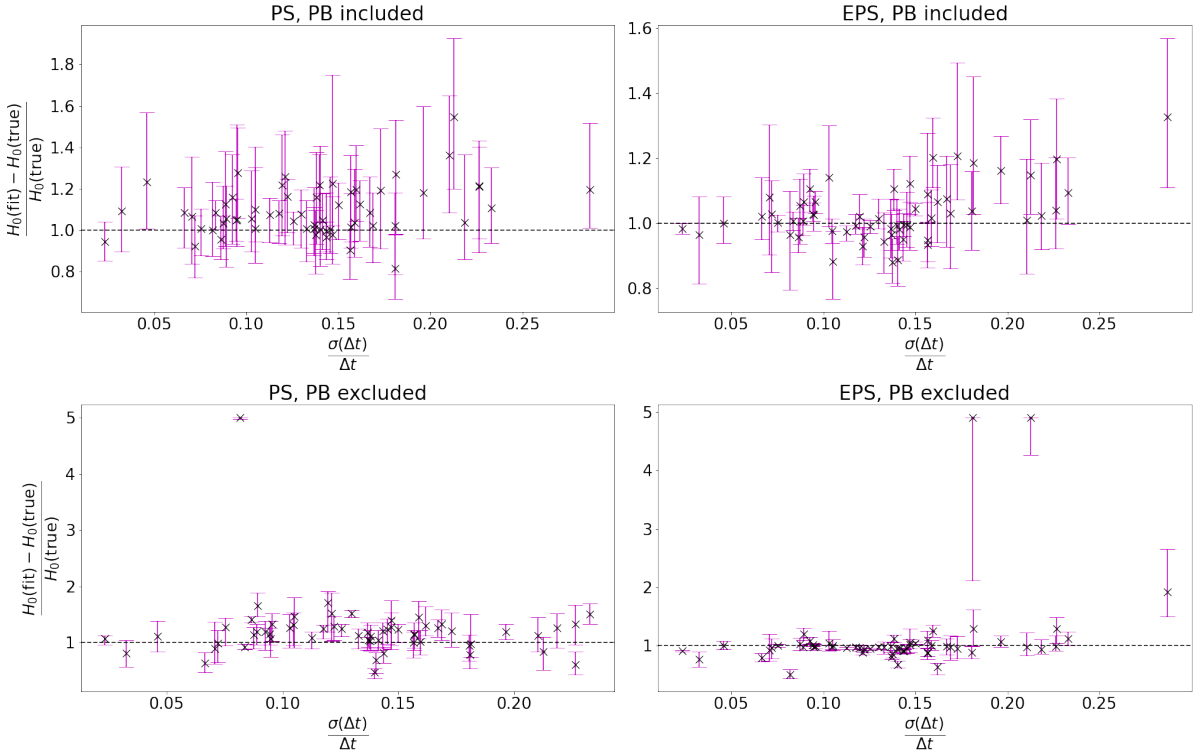

Figure 4.5: *Measurements of the Hubble constant H_0 from images with (EPS, right) and without (PS, left) an extended source present, comparing cases where the post-Born correction is included (top) or excluded (bottom), for images featuring quadruply-lensed point sources. On the horizontal axis, the relative uncertainty in the time delay measurements used in the fit is shown. Note the differences in vertical axis limits between the different graphs.*

Conclusion

The last century has seen our understanding of the Universe change on a fundamental level. Astrophysics and cosmology have answered questions about the past, present and future of the cosmos once thought to be far beyond our reach, but with each advancement has come new questions posing an almost existential threat to current scientific paradigms, and around 95% of the content of the Universe remains hidden from view. Answering these questions will require new methods of scientific inquiry, and advancements in theory to match those in its implementation. Gravitational lensing has emerged as one of the most compelling probes of the content and history of the Universe at some of its largest scales, and could be the key to resolving some of the most fascinating challenges faced by science today. Before this is possible, however, new formalisms to model weak and strong lensing effects simultaneously must be integrated into the field, and systematic sources of bias and error must be carefully identified and constrained.

With these challenges as its motivation, this project aimed to present and expand on the current understanding of the importance of line-of-sight effects in strong gravitational lensing, with a particular interest in how these effects influence strong lensing time delays, and how they could best be measured. In chapter 1, we presented and interpreted the theoretical basis of these effects. This included a new derivation of the time delay between strongly lensed images of the same source for a single lens plane, the first explicit demonstration of the degeneracy between the ellipticity and foreground shear for the widely-used elliptical power law lens, an illustration of the danger of under-parameterised mass constraints from velocity dispersion data, and a brief discussion of the possibility of using this data to measure the foreground shear. These derivations lead us to conclude that the modelling of line-of-sight effects, even when time delay data are included, is limited by several key degeneracies. Line-of-sight convergence terms are fundamentally unmeasurable without independent constraints, and we argued that these constraints must be carefully parameterised to avoid the introduction of systematic biases. The shear terms γ_{os} and γ_{ds} are degenerate with each other and with γ_{od}, and can only be measured in combination as $\gamma_{\rm LOS}$. The weighting function for $\gamma_{\rm LOS}$ is different to that of the weak lensing parameter γ_{os}, which is often treated as being equivalent to the $\gamma_{\rm ext}$ parameter included in many lens models, and thus care is needed when comparing measurements of these parameters. The foreground shear γ_{od} is degenerate with the ellipticity of an elliptical power law lens to first order, but could be fully degenerate with a more complicated lens model. We therefore argued that the two terms of the complex shear $\gamma_{\rm LOS}$ are the only truly independent parameters introduced into the lens equation and time delay function by the presence of tidal line-of-sight effects, and thus the only parameters which could be measurable from lensing observables. Velocity dispersion data, however, hold the potential to lift the degeneracy on both the convergence and foreground shear terms.

In chapter 2, we presented an implementation of tidal line-of-sight effects on strong lensing time delays into `lenstronomy`, an open-source software package for strong lens modelling. This implementation is available in the most recent version of `lenstronomy`, and forms a part of the subpackage used in [77]. In chapter 3, this theory and code was applied to the problem of constraining the line-of-sight shear from mock datasets, paying particular attention to the changes to the precision and accuracy of results when the sources and fitting approaches were changed. Our results indicate that high SNR Einstein ring images present the best opportunity for precise constraints on $\gamma_{\rm LOS}$, but that multiply-imaged point sources and time delay measurements can improve both the precision and accuracy of noisier images. In the case where only a point source is present, time delays can also offer improvements on constraints. The value of including time delays increased as their measurements were more precise and as the extended source became harder to detect above the background noise, but in all cases, these improvements were subtle. It therefore seems that the cost and effort of time delay measurements cannot be justified for line-of-sight shear measurements, but that they have value as additional constraints when such measurements have already been made. Their value may be most appreciated in large statistical studies of the line-of-sight shear, where strongly lensed quasar or supernova images may be an important source of productive data.

We also investigated the effects on $\gamma_{\rm LOS}$ recovery when omitting the foreground shear from an EPL main lens model, and found that large systematic biases were possible and errorbars were significantly underestimated. While this should not be taken to mean that γ_{od} can escape degeneracy with the main lens, it does highlight its potential value as an additional source of flexibility in simpler lens models. Lastly, we illustrated the degeneracy between e and γ_{od} by attempting to measure e from models were γ_{od} was fixed to zero.

In chapter 4, the final component of this project focused on the effects of ignoring the foreground shear when trying to measure the Hubble constant. We showed that, despite their first-order degeneracy, the time delays between images from a lens in the presence of foreground shear could differ measurably from those from a lens with the corresponding value of e, and went on to show that lenses featuring both an extended and a point source could give significantly biased H_0 values when modelled without the foreground shear. Interestingly, measurements from point sources did not seem to suffer from this bias, although there was a small loss in precision.

It is our hope that these results could contribute to a more careful approach to modelling line-of-sight effects in strong lensing, and inform both the target choices and interpretations of results in future attempts to measure the line-of-sight shear. Nonetheless, many questions remain to be answered. While an interesting prospect, the value of velocity dispersion measurements in constraining the foreground shear has yet to be tested in the context of strong lensing, and the extent to which these constraints are valuable for other related measurements is not yet clear. The simulated data used to measure the line-of-sight shear were somewhat simplified and idealised, and the extent to which our conclusions generalise to more complex lens models or low SNR point source images remains to be seen. This is equally true for the results relating to the omission of the foreground shear, results which we would expect to change if a two-component lens model were used. Finally, it has been suggested that higher-order line-of-sight effects may be ripe with information [9, 108, 57, 77]. However, their consequences for measurements of $\gamma_{\rm LOS}$ in the presence of time delays has yet to be explored. Our hope is that this project will serve as another layer in the foundations on which these future studies could be built.

Acknowledgements

I would like to thank my supervisors Lucia Marchetti, whose door was always open whenever I needed help or advice, and whose generosity and enthusiasm for the prospects of the project made this thesis possible, Julien Larena, who stayed centrally involved in the project despite having left UCT for the University of Montpellier, who made visiting Montpellier possible and who supported me every step of the way, and to Pierre Fleury, who went from joining a virtual meeting to becoming inextricably involved in each of its stages, and who welcomed me to the University of Paris-Saclay for my visit. I am immensely grateful for the patience and generosity of each of their efforts in making this project possible, and I have gained so much from their insights over these last years. I would also like to thank Natalie Hogg for her invaluable support during the work on the lenstronomy subpackage. I am very grateful to Julien, Pierre, Natalie, Simon Birrer, James Nightingale and the other participants in the 2022 Line-of-Sight workshop in Montpellier, whose own research and feedback has been very influential on shaping the latter stages of this project. I would like to express my thanks to Simon Birrer and Adam Amara for the use of `lenstronomy`, and for enabling the line-of-sight effects subpackage to be implemented, to Daniel Gilman for reviewing the pull request for this subpackage, and to Jeremy Smith and the Ilifu support team for their assistance during the use of the ilifu cloud computing facility. I would also like to thank Julien Lavalle for making it possible for me to attend the Dark Matter Subhaloes conference in Montpellier at very short notice, and express my gratitude to the members of the labs in Montpellier and Paris who made me feel so welcome during my visit. Finally, I would like to express my deepest gratitude to Saalih Allie and Rushana Fagodien for their tireless support and assistance throughout my Master's degree, whose work to secure funding for me over the last two years has made this project possible.

This work was supported by the David & Elaine Potter Foundation and the National Astrophysics and Space Science Program (NASSP). We acknowledge the use of the ilifu cloud computing facility – www.ilifu.ac.za, a partnership between the University of Cape Town, the University of the Western Cape, the University of Stellenbosch, Sol Plaatje University, the Cape Peninsula University of Technology and the South African Radio Astronomy Observatory. The ilifu facility is supported by contributions from the Inter-University Institute for Data Intensive Astronomy (IDIA – a partnership between the University of Cape Town, the University of Pretoria, the University of the Western Cape and the South African Radio astronomy Observatory), the Computational Biology division at UCT and the Data Intensive Research Initiative of South Africa (DIRISA)."

Bibliography

[1] T M C Abbott, F B Abdalla, J Annis, K Bechtol, J Blazek, B A Benson, R A Bernstein, G M Bernstein, E Bertin, D Brooks, D L Burke, A Carnero Rosell, M Carrasco Kind, J Carretero, F J Castander, C L Chang, T M Crawford, C E Cunha, C B D'Andrea, L N da Costa, C Davis, J DeRose, S Desai, H T Diehl, J P Dietrich, P Doel, A Drlica-Wagner, A E Evrard, E Fernandez, B Flaugher, P Fosalba, J Frieman, J García-Bellido, E Gaztanaga, D W Gerdes, T Giannantonio, D Gruen, R A Gruendl, J Gschwend, G Gutierrez, W G Hartley, J W Henning, K Honscheid, B Hoyle, D Huterer, B Jain, D J James, M Jarvis, T Jeltema, M D Johnson, M W G Johnson, E Krause, K Kuehn, S Kuhlmann, N Kuropatkin, O Lahav, A R Liddle, M Lima, H Lin, N MacCrann, M A G Maia, A Manzotti, M March, J L Marshall, R Miquel, J J Mohr, T Natoli, P Nugent, R L C Ogando, Y Park, A A Plazas, C L Reichardt, K Reil, A Roodman, A J Ross, E Rozo, E S Rykoff, E Sanchez, V Scarpine, M Schubnell, D Scolnic, I Sevilla-Noarbe, E Sheldon, M Smith, R C Smith, M Soares-Santos, F Sobreira, E Suchyta, G Tarle, D Thomas, M A Troxel, A R Walker, R H Wechsler, J Weller, W Wester, W L K Wu, J Zuntz, and The Dark Energy Survey and the South Pole Telescope Collaborations). Dark Energy Survey Year 1 Results: A Precise H0 Estimate from DES Y1, BAO, and D/H Data. *Monthly Notices of the Royal Astronomical Society*, 480(3):3879–3888, 07 2018.

[2] A. Agnello, M. W. Auger, and N. W. Evans. Lensing and dynamics in two simple steps. *Monthly Notices of the Royal Astronomical Society*, 429:L35–L39, 02 2013.

[3] Adam Amara, R. Benton Metcalf, Thomas J. Cox, and Jeremiah P. Ostriker. Simulations of strong gravitational lensing with substructure. *Monthly Notices of the Royal Astronomical Society*, 367(4):1367–1378, 03 2006.

[4] Luca Amendola, Stephen Appleby, Anastasios Avgoustidis, David Bacon, Tessa Baker, Marco Baldi, Nicola Bartolo, Alain Blanchard, Camille Bonvin, Stefano Borgani, Enzo Branchini, Clare Burrage, Stefano Camera, Carmelita Carbone, Luciano Casarini, Mark Cropper, Claudia de Rham, Jörg P. Dietrich, Cinzia Di Porto, Ruth Durrer, Anne Ealet, Pedro G. Ferreira, Fabio Finelli, Juan García-Bellido, Tommaso Giannantonio, Luigi Guzzo, Alan Heavens, Lavinia Heisenberg, Catherine Heymans, Henk Hoekstra, Lukas Hollenstein, Rory Holmes, Zhiqi Hwang, Knud Jahnke, Thomas D. Kitching, Tomi Koivisto, Martin Kunz, Giuseppe La Vacca, Eric Linder, Marisa March, Valerio Marra, Carlos Martins, Elisabetta Majerotto, Dida Markovic, David Marsh, Federico Marulli, Richard Massey, Yannick Mellier, Francesco Montanari, David F. Mota, Nelson J. Nunes, Will Percival, Valeria Pettorino, Cristiano Porciani, Claudia Quercellini, Justin Read, Massimiliano Rinaldi, Domenico Sapone, Ignacy Sawicki, Roberto Scaramella, Constantinos Skordis, Fergus Simpson, Andy Taylor, Shaun Thomas, Roberto Trotta, Licia Verde, Filippo Vernizzi, Adrian Vollmer, Yun Wang, Jochen Weller, and Tom Zlosnik. Cosmology and fundamental physics with the euclid satellite. *Living Reviews in Relativity*, 21(1), 04 2018.

[5] A Amvrosiadis, S A Eales, M Negrello, L Marchetti, M W L Smith, N Bourne, D L Clements, G De Zotti, L Dunne, S Dye, C Furlanetto, R J Ivison, S J Maddox, E Valiante, M Baes, A J Baker, A Cooray, S M Crawford, D Frayer, A Harris, M J Michałowski, H Nayyeri, S Oliver, D A Riechers, S Serjeant, and M Vaccari. ALMA observations of lensed herschel sources: testing the dark matter halo paradigm. *Monthly Notices of the Royal Astronomical Society*, 475(4):4939–4952, 01 2018.

[6] and N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J.-F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba,

L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J.-M. Delouis, E. Di Valentino, J. M. Diego, O. Doré , M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J.-M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y.-Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M.-A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J.-L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A.-S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, sep 2020.

[7] Marika Asgari, Chieh-An Lin, Benjamin Joachimi, Benjamin Giblin, Catherine Heymans, Hendrik Hildebrandt, Arun Kannawadi, Benjamin Stölzner, Tilman Tröster, Jan Luca van den Busch, Angus H. Wright, Maciej Bilicki, Chris Blake, Jelte de Jong, Andrej Dvornik, Thomas Erben, Fedor Getman, Henk Hoekstra, Fabian Köhlinger, Konrad Kuijken, Lance Miller, Mario Radovich, Peter Schneider, HuanYuan Shan, and Edwin Valentijn. KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics. *Astronomy & Astrophysics*, 645:A104, 01 2021.

[8] M. W. Auger, C. D. Fassnacht, A. L. Abrahamse, L. M. Lubin, and G. K. Squires. The gravitational lens-galaxy group connection. II. groups associated with b2319+051 and b1600+434. *The Astronomical Journal*, 134(2):668, 06 2007.

[9] D. J. Bacon, D. M. Goldberg, B. T. P. Rowe, and A. N. Taylor. Weak gravitational flexion. *Monthly Notices of the Royal Astronomical Society*, 365(2):414–428, 01 2006.

[10] David J. Bacon, Richard A. Battye, Philip Bull, Stefano Camera, Pedro G. Ferreira, Ian Harrison, David Parkinson, Alkistis Pourtsidou, Má rio G. Santos, Laura Wolz, Filipe Abdalla, Yashar Akrami, David Alonso, Sambatra Andrianomena, Mario Ballardini, José Luis Bernal, Daniele Bertacca, Carlos A. P. Bengaly, Anna Bonaldi, Camille Bonvin, Michael L. Brown, Emma Chapman, Song Chen, Xuelei Chen, Steven Cunnington, Tamara M. Davis, Clive Dickinson, José Fonseca, Keith Grainge, Stuart Harper, Matt J. Jarvis, Roy Maartens, Natasha Maddox, Hamsa Padmanabhan, Jonathan R. Pritchard, Alvise Raccanelli, Marzia Rivi, Sambit Roychowdhury, Martin Sahlén, Dominik J. Schwarz, Thilo M. Siewert, Matteo Viel, Francisco Villaescusa-Navarro, Yidong Xu, Daisuke Yamauchi, and Joe Zuntz. Cosmology with phase 1 of the square kilometre array red book 2018: Technical specifications and performance forecasts. *Publications of the Astronomical Society of Australia*, 37, 2020.

[11] David J. Bacon, Alexandre Refregier, Douglas Clowe, and Richard S. Ellis. Numerical simulations of weak lensing measurements. *Monthly Notices of the Royal Astronomical Society*, 325(3):1065–1074, 2001.

[12] Rennan Bar-Kana. Effect of Large-Scale Structure on Multiply Imaged Sources. *Astrophysical Journal*, 468:17, 09 1996.

[13] Matthias Bartelmann. Gravitational lensing. *Classical and Quantum Gravity*, 27(23):233001, 11 2010.

[14] Matthias Bartelmann, Andreas Huss, Joerg M. Colberg, Adrian Jenkins, and Frazer R. Pearce. Arc statistics with realistic cluster potentials. iv. clusters in different cosmologies. *Astronomy & Astrophysics*, 1997.

[15] Matthias Bartelmann and Peter Schneider. Weak gravitational lensing. *Physics Reports*, 340(4-5):291–472, 01 2001.

[16] Anne H. Bauer, Charles Baltay, Nancy Ellman, Jonathan Jerke, David Rabinowitz, and Richard Scalzo. The Mass-Richness Relation of MaxBCG Clusters from Quasar Lensing Magnification Using Variability. *Astrophysical Journal*, 749(1):56, 04 2012.

[17] Matthew B. Bayliss, Traci Johnson, Michael D. Gladders, Keren Sharon, and Masamune Oguri. Line-of-sight structure toward strong lensing galaxy clusters. *The Astrophysical Journal*, 783(1):41, 02 2014.

[18] S. Birrer, M. Millon, D. Sluse, A. J. Shajib, F. Courbin, L. V. E. Koopmans, S. H. Suyu, and T. Treu. Time-delay cosmography: Measuring the hubble constant and other cosmological parameters with strong gravitational lensing, 2022.

[19] S. Birrer, A. J. Shajib, A. Galan, M. Millon, T. Treu, A. Agnello, M. Auger, G. C.-F. Chen, L. Christensen, T. Collett, F. Courbin, C. D. Fassnacht, L. V. E. Koopmans, P. J. Marshall, J.-W. Park, C. E. Rusu, D. Sluse, C. Spiniello, S. H. Suyu, S. Wagner-Carena, K. C. Wong, M. Barnabè , A. S. Bolton, O. Czoske, X. Ding, J. A. Frieman, and L. Van de Vyvere. TDCOSMO. *Astronomy & Astrophysics*, 643:A165, 11 2020.

[20] S Birrer, T Treu, C E Rusu, V Bonvin, C D Fassnacht, J H H Chan, A Agnello, A J Shajib, G C-F Chen, M Auger, F Courbin, S Hilbert, D Sluse, S H Suyu, K C Wong, P Marshall, B C Lemaux, and G Meylan. H0LiCOW – IX. cosmographic analysis of the doubly imaged quasar SDSS 1206+4332 and a new measurement of the hubble constant. *Monthly Notices of the Royal Astronomical Society*, 484(4):4726–4753, 01 2019.

[21] Simon Birrer and Adam Amara. lenstronomy: Multi-purpose gravitational lens modelling software package. *Physics of the Dark Universe*, 22:189–201, 12 2018.

[22] Simon Birrer, Adam Amara, and Alexandre Refregier. The mass-sheet degeneracy and time-delay cosmography: analysis of the strong lens RXJ1131-1231. *Journal of Cosmology and Astroparticle Physics*, 2016(08):020–020, 08 2016.

[23] Simon Birrer, Cyril Welschen, Adam Amara, and Alexandre Refregier. Line-of-sight effects in strong lensing: putting theory into practice. *Journal of Cosmology and Astroparticle Physics*, 2017(04):049–049, 04 2017.

[24] A. W. Blain. Detecting gravitational lensing cosmic shear from samples of several galaxies using two-dimensional spectral imaging. *The Astrophysical Journal*, 570(2):L51, 04 2002.

[25] Roger Blandford and Ramesh Narayan. Fermat's Principle, Caustics, and the Classification of Gravitational Lens Images. *Astrophysical Journal*, 310:568, 11 1986.

[26] Adam S. Bolton, Scott Burles, Leon V. E. Koopmans, Tommaso Treu, and Leonidas A. Moustakas. The sloan lens ACS survey. i. a large spectroscopically selected sample of massive early-type lens galaxies. *The Astrophysical Journal*, 638(2):703–724, 02 2006.

[27] V. Bonvin, F. Courbin, S. H. Suyu, P. J. Marshall, C. E. Rusu, D. Sluse, M. Tewes, K. C. Wong, T. Collett, C. D. Fassnacht, T. Treu, M. W. Auger, S. Hilbert, L. V. E. Koopmans, G. Meylan, N. Rumbaugh, A. Sonnenfeld, and C. Spiniello. H0LiCOW V. new COSMOGRAIL time delays of HE 0435-1223: H_0 to 3.8 percent precision from strong lensing in a flat λCDM model. *Monthly Notices of the Royal Astronomical Society*, 465(4):4914–4930, 11 2016.

[28] Victor Boscá, Pierre Fleury, and Juan García-Bellido. Accurate modelling of extragalactic microlensing by compact objects. *Journal of Cosmology and Astroparticle Physics*, 2022(10):098, oct 2022.

[29] Claudio Bruderer, Justin I. Read, Jonathan P. Coles, Dominik Leier, Emilio E. Falco, Ignacio Ferreras, and Prasenjit Saha. Light versus dark in strong-lens galaxies: dark matter haloes that are rounder than their stars. *Monthly Notices of the Royal Astronomical Society*, 456:870–884, 2015.

[30] G. B. Caminha, S. H. Suyu, A. Mercurio, G. Brammer, P. Bergamini, A. Acebron, and E. Vanzella. First JWST observations of a gravitational lens. mass model from new multiple images with near-infrared observations of SMACS˜j0723.3-7327. *Astronomy & Astrophysics*, 09 2022.

[31] Xiaoyue Cao, Ran Li, J. W. Nightingale, Richard Massey, Andrew Robertson, Carlos S. Frenk, Aristeidis Amvrosiadis, Nicola C. Amorisco, Qiuhan He, Amy Etherington, Shaun Cole, and Kai Zhu. Systematic errors induced by the elliptical power-law model in galaxy–galaxy strong lens modeling. *Research in Astronomy and Astrophysics*, 22(2):025014, 02 2022.

[32] N. Caon, M. Capaccioli, and M. D'Onofrio. On the shape of the light profiles of early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 265:1013–1021, 12 1993.

[33] Michele Cappellari, Nicholas Scott, Katherine Alatalo, Leo Blitz, Maxime Bois, Frédéric Bournaud, M. Bureau, Alison F. Crocker, Roger L. Davies, Timothy A. Davis, P. T. de Zeeuw, Pierre-Alain Duc, Eric Emsellem, Sadegh Khochfar, Davor Krajnović, Harald Kuntschner, Richard M. McDermid, Raffaella Morganti, Thorsten Naab, Tom Oosterloo, Marc Sarzi, Paolo Serra, Anne-Marie Weijmans, and Lisa M. Young. The ATLAS^{3D} project - XV. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Mass Plane. *Monthly Notices of the Royal Astronomical Society*, 432(3):1709–1741, 07 2013.

[34] Sean M Carroll. *Spacetime and geometry*. Cambridge University Press, 2004.

[35] Kyu-Hyun Chae, Shude Mao, and Pedro Augusto. Modelling the first probable two-plane lens system B2114+022: reproducing two compact radio cores A and D. *Monthly Notices of the Royal Astronomical Society*, 326(3):1015–1026, 09 2001.

[36] Geoff C-F Chen, James H H Chan, Vivien Bonvin, Christopher D Fassnacht, Karina Rojas, Martin Millon, Fred Courbin, Sherry H Suyu, Kenneth C Wong, Dominique Sluse, Tommaso Treu, Anowar J Shajib, Jen-Wei Hsueh, David J Lagattuta, Léon V E Koopmans, Simona Vegetti, and John P McKean. Constraining the microlensing effect on time delays with a new time-delay prediction model in H0 measurements. *Monthly Notices of the Royal Astronomical Society*, 481(1):1115–1125, 09 2018.

[37] Geoff C-F Chen, Christopher D Fassnacht, Sherry H Suyu, Cristian E Rusu, James H H Chan, Kenneth C Wong, Matthew W Auger, Stefan Hilbert, Vivien Bonvin, Simon Birrer, Martin Millon, Lé on V E Koopmans, David J Lagattuta, John P McKean, Simona Vegetti, Frederic Courbin, Xuheng Ding, Aleksi Halkola, Inh Jee, Anowar J Shajib, Dominique Sluse, Alessandro Sonnenfeld, and Tommaso Treu. A SHARP view of H0LiCOW: H0 from three time-delay gravitational lens systems with adaptive optics imaging. *Monthly Notices of the Royal Astronomical Society*, 490(2):1743–1773, 09 2019.

[38] Jacqueline Chen, Savvas M. Koushiappas, and Andrew R. Zentner. The effects of halo-to-halo variation on substructure lensing. *The Astrophysical Journal*, 741(2):117, 10 2011.

[39] Thomas E. Collett. The population of galaxy–galaxy strong lenses in forthcoming optical imaging surveys. *The Astrophysical Journal*, 811(1):20, 09 2015.

[40] Thomas E. Collett, Philip J. Marshall, Matthew W. Auger, Stefan Hilbert, Sherry H. Suyu, Zachary Greene, Tommaso Treu, Christopher D. Fassnacht, Léon V. E. Koopmans, Maruša Bradač, and Roger D. Blandford. Reconstructing the lensing mass in the Universe from photometric catalogue data. *Monthly Notices of the Royal Astronomical Society*, 432(1):679–692, 04 2013.

[41] Neal Dalal and Casey R. Watson. What are the environments of lens galaxies?, 2004.

[42] C. O. de Burgh-Day, E. N. Taylor, R. L. Webster, and A. M. Hopkins. Direct shear mapping - a new weak lensing tool. *Monthly Notices of the Royal Astronomical Society*, 451(2):2161–2173, 08 2015.

[43] C. O. de Burgh-Day, E. N. Taylor, R. L. Webster, and A. M. Hopkins. Direct shear mapping: Prospects for weak lensing studies of individual galaxy–galaxy lensing systems. *Publications of the Astronomical Society of Australia*, 32:e040, 2015.

[44] Giulia Despali, Simona Vegetti, Simon D. M. White, Carlo Giocoli, and Frank C. van den Bosch. Modelling the line-of-sight contribution in substructure lensing. *Monthly Notices of the Royal Astronomical Society*, 475(4):5424–5442, 04 2018.

[45] Birendra Dhanasingham, Francis-Yan Cyr-Racine, Annika H G Peter, Andrew Benson, and Daniel Gilman. Interlopers speak out: studying the dark universe using small-scale lensing anisotropies. *Monthly Notices of the Royal Astronomical Society*, 518(4):5843–5861, 10 2022.

[46] Veronica Dike, Daniel Gilman, and Tommaso Treu. Strong lensing constraints on primordial black holes as a dark matter candidate, 2022.

[47] Gregory Dobler, Christopher D. Fassnacht, Tommaso Treu, Phil Marshall, Kai Liao, Alireza Hojjati, Eric Linder, and Nicholas Rumbaugh. Strong lens time delay challenge I. experimental design. *The Astrophysical Journal*, 799(2):168, 01 2015.

[48] S Dye, C Furlanetto, L Dunne, S A Eales, M Negrello, H Nayyeri, P P van der Werf, S Serjeant, D Farrah, M J Michałowski, M Baes, L Marchetti, A Cooray, D A Riechers, and A Amvrosiadis. Modelling high-resolution ALMA observations of strongly lensed highly star-forming galaxies detected by *Herschel*. *Monthly Notices of the Royal Astronomical Society*, 476(4):4383–4394, 02 2018.

[49] F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, 220:291–333, 01 1920.

[50] R. Fadely, C. R. Keeton, R. Nakajima, and G. M. Bernstein. Improved constraints on the gravitational lens q0957+561. II. strong lensing. *The Astrophysical Journal*, 711(1):246, 02 2010.

[51] E. E. Falco, M. V. Gorenstein, and I. I. Shapiro. On model-dependent bounds on H0 from gravitational images : application to Q 0957+561 A, B. *Astrophysical Journal, Letters*, 289:L1–L4, 02 1985.

[52] C. D. Fassnacht, R. R. Gal, L. M. Lubin, J. P. McKean, Gordon K. Squires, and A. C. S. Readhead. Mass along the line of sight to the gravitational lens B1608+656: Galaxy groups and implications for H0*. *The Astrophysical Journal*, 642(1):30, 05 2006.

[53] C. D. Fassnacht, L. V. E. Koopmans, and K. C. Wong. Galaxy number counts and implications for strong lensing. *Monthly Notices of the Royal Astronomical Society*, 410(4):2167–2179, 01 2011.

[54] Tehani K. Finch, Lisa P. Carlivati, Joshua N. Winn, and Paul L. Schechter. Analytic expressions for mean magnification by a quadrupole gravitational lens. *The Astrophysical Journal*, 577(1):51–56, 09 2002.

[55] Philippe Fischer, Gary Bernstein, George Rhee, and J. Anthony Tyson. The Mass Distribution of the Cluster 0957+561 From Gravitational Lensing. *The Astrophysical Journal*, 113:521, 02 1997.

[56] Pierre Fleury, Julien Larena, and Jean-Philippe Uzan. Cosmic convergence and shear with extended sources. *Physical Review D*, 99(2), 01 2019.

[57] Pierre Fleury, Julien Larena, and Jean-Philippe Uzan. Weak lensing distortions beyond shear. *Physical Review D*, 99(2), 01 2019.

[58] Pierre Fleury, Julien Larena, and Jean-Philippe Uzan. Gravitational lenses in arbitrary space-times. *Classical and Quantum Gravity*, 38(8):085002, 03 2021.

[59] Pierre Fleury, Julien Larena, and Jean-Philippe Uzan. Line-of-sight effects in strong gravitational lensing. *JCAP*, 08:024, 2021.

[60] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306, 03 2013.

[61] Raphaël Gavazzi, Tommaso Treu, Léon V. E. Koopmans, Adam S. Bolton, Leonidas A. Moustakas, Scott Burles, and Philip J. Marshall. The Sloan Lens ACS Survey. VI. Discovery and Analysis of a Double Einstein Ring. *Astrophysical Journal*, 677(2):1046–1059, 04 2008.

[62] M. Gomer and L.L.R. Williams. Galaxy-lens determination of H0: constraining density slope in the context of the mass sheet degeneracy. *Journal of Cosmology and Astroparticle Physics*, 2020(11):045–045, 11 2020.

[63] Matthew R. Gomer, Dominique Sluse, Lyne Van de Vyvere, Simon Birrer, and Frederic Courbin. Tdcosmo VIII: A key test of systematics in the hierarchical method of time-delay cosmography, 2022.

[64] Matthew R Gomer and Liliya L R Williams. Galaxy-lens determination of H0: the effect of the ellipse + shear modelling assumption. *Monthly Notices of the Royal Astronomical Society*, 504(1):1340–1354, 04 2021.

[65] A. Goobar, R. Amanullah, S. R. Kulkarni, P. E. Nugent, J. Johansson, C. Steidel, D. Law, E. Mörtsell, R. Quimby, N. Blagorodnova, A. Brandeker, Y. Cao, A. Cooray, R. Ferretti, C. Fremling, L. Hangard, M. Kasliwal, T. Kupfer, R. Lunnan, F. Masci, A. A. Miller, H. Nayyeri, J. D. Neill, E. O. Ofek, S. Papadogiannakis, T. Petrushevska, V. Ravi, J. Sollerman, M. Sullivan, F. Taddia, R. Walters, D. Wilson, L. Yan, and O. Yaron. iPTF16geu: A multiply imaged, gravitationally lensed type ia supernova. *Science*, 356(6335):291–295, 2017.

[66] M. V. Gorenstein, E. E. Falco, and I. I. Shapiro. Degeneracies in Parameter Estimates for Models of Gravitational Lens Systems. *Astrophysical Journal*, 327:693, 04 1988.

[67] Zach S. Greene, Sherry H. Suyu, Tommaso Treu, Stefan Hilbert, Matthew W. Auger, Thomas E. Collett, Philip J. Marshall, Christopher D. Fassnacht, Roger D. Blandford, Maruša Bradač, and Léon V. E. Koopmans. Improving the precision of time-delay cosmography with observations of galaxies along the line of sight. *The Astrophysical Journal*, 768(1):39, 04 2013.

[68] James E. Gunn. On the Propagation of Light in Inhomogeneous Cosmologies. I. Mean Effects. *Astrophysical Journal*, 150:737, 12 1967.

[69] Pol Gurri, Edward N Taylor, and Christopher J Fluke. The first shear measurements from precision weak lensing. *Monthly Notices of the Royal Astronomical Society*, 499(4):4591–4604, 09 2020.

[70] N. P. Hathi, S. Malhotra, and J. E. Rhoads. Starburst intensity limit of galaxies at z $\simeq$ 5-6. *The Astrophysical Journal*, 673(2):686–693, 02 2008.

[71] Qiuhan He, James Nightingale, Andrew Robertson, Aristeidis Amvrosiadis, Shaun Cole, Carlos S Frenk, Richard Massey, Ran Li, Nicola C Amorisco, R Benton Metcalf, Xiaoyue Cao, and Amy Etherington. Testing strong lensing subhalo detection with a cosmological simulation. *Monthly Notices of the Royal Astronomical Society*, 518(1):220–239, 09 2022.

[72] Yashar D. Hezaveh, Laurence Perreault Levasseur, and Philip J. Marshall. Fast automated analysis of strong gravitational lenses with convolutional neural networks. *Nature*, 548(7669):555–557, 08 2017.

[73] Chiaki Hikage, Masamune Oguri, Takashi Hamana, Surhud More, Rachel Mandelbaum, Masahiro Takada, Fabian Köhlinger, Hironao Miyatake, Atsushi J Nishizawa, Hiroaki Aihara, Robert Armstrong, James Bosch, Jean Coupon, Anne Ducout, Paul Ho, Bau-Ching Hsieh, Yutaka Komiyama, François Lanusse, Alexie Leauthaud, Robert H Lupton, Elinor Medezinski, Sogo Mineo, Shoken Miyama, Satoshi Miyazaki, Ryoma Murata, Hitoshi Murayama, Masato Shirasaki, Cristóbal Sifón, Melanie Simet, Joshua Speagle, David N Spergel, Michael A Strauss, Naoshi Sugiyama, Masayuki Tanaka, Yousuke Utsumi, Shiang-Yu Wang, and Yoshihiko Yamada. Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data. *Publications of the Astronomical Society of Japan*, 71(2), 03 2019. 43.

[74] H. Hildebrandt, F. Köhlinger, J. L. van den Busch, B. Joachimi, C. Heymans, A. Kannawadi, A. H. Wright, M. Asgari, C. Blake, H. Hoekstra, S. Joudaki, K. Kuijken, L. Miller, C. B. Morrison, T. Tröster, A. Amon, M. Archidiacono, S. Brieden, A. Choi, J. T. A. de Jong, T. Erben, B. Giblin, A. Mead, J. A. Peacock, M. Radovich, P. Schneider, C. Sifó n, and M. Tewes. KiDS+VIKING-450: Cosmic shear tomography with optical and infrared data. *Astronomy & Astrophysics*, 633:A69, 01 2020.

[75] Samuel R. Hinton. ChainConsumer. *The Journal of Open Source Software*, 1(4):00045, 08 2016.

[76] David W. Hogg and Daniel Foreman-Mackey. Data analysis recipes: Using markov chain monte carlo. *The Astrophysical Journal Supplement Series*, 236(1):11, 05 2018.

[77] Natalie B. Hogg, Pierre Fleury, Julien Larena, and Matteo Martinelli. Measuring line-of-sight shear with einstein rings: a proof of concept, 2022.

[78] Gilbert P. Holder and Paul L. Schechter. External shear in quadruply imaged lens systems. *The Astrophysical Journal*, 589(2):688–692, 06 2003.

[79] X. Huang, C. Storfer, A. Gu, V. Ravi, A. Pilon, W. Sheu, R. Venguswamy, S. Banka, A. Dey, M. Landriau, D. Lang, A. Meisner, J. Moustakas, A. D. Myers, R. Sajith, E. F. Schlafly, and D. J. Schlegel. Discovering new strong gravitational lenses in the desi legacy imaging surveys. *The Astrophysical Journal*, 909(1):27, 03 2021.

[80] X. Huang, C. Storfer, V. Ravi, A. Pilon, M. Domingo, D. J. Schlegel, S. Bailey, A. Dey, R. R. Gupta, D. Herrera, S. Juneau, M. Landriau, D. Lang, A. Meisner, J. Moustakas, A. D. Myers, E. F. Schlafly, F. Valdes, B. A. Weaver, J. Yang, and C. Yèche. Finding strong gravitational lenses in the desi decam legacy survey. *The Astrophysical Journal*, 894(1):78, 05 2020.

[81] Naohisa Inada, Masamune Oguri, Min-Su Shin, Issha Kayo, Michael A. Strauss, Tomoki Morokuma, Cristian E. Rusu, Masataka Fukugita, Christopher S. Kochanek, Gordon T. Richards, Donald P. Schneider, Donald G. York, Neta A. Bahcall, Joshua A. Frieman, Patrick B. Hall, and Richard L. White. The sloan digital sky survey quasar lens search. v. final catalog from the seventh data release. *The Astronomical Journal*, 143(5):119, 04 2012.

[82] C. Jacobs, T. Collett, K. Glazebrook, E. Buckley-Geer, H. T. Diehl, H. Lin, C. McCarthy, A. K. Qin, C. Odden, M. Caso Escudero, P. Dial, V. J. Yung, S. Gaitsch, A. Pellico, K. A. Lindgren, T. M. C. Abbott, J. Annis, S. Avila, D. Brooks, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, L. N. da Costa, J. De Vicente, P. Fosalba, J. Frieman, J. García-Bellido, E. Gaztanaga, D. A. Goldstein, D. Gruen, R. A. Gruendl, J. Gschwend, D. L. Hollowood, K. Honscheid, B. Hoyle, D. J. James, E. Krause, N. Kuropatkin, O. Lahav, M. Lima, M. A. G. Maia, J. L. Marshall, R. Miquel, A. A. Plazas, A. Roodman, E. Sanchez, V. Scarpine, S. Serrano, I. Sevilla-Noarbe, M. Smith, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, V. Vikram, A. R. Walker, Y. Zhang, and (DES Collaboration). An extended catalog of galaxy–galaxy strong gravitational lenses discovered in des using convolutional neural networks. *The Astrophysical Journal Supplement Series*, 243(1):17, 07 2019.

[83] M. Jaroszynski and Z. Kostrzewa-Rutkowska. The influence of the matter along the line of sight and in the lens environment on the strong gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 439(3):2432–2441, 04 2014.

[84] Michal Jaroszynski, Changbom Park, Bohdan Paczynski, and III Gott, J. Richard. Weak Gravitational Lensing Due to Large-Scale Structure of the Universe. *Astrophysical Journal*, 365:22, 12 1990.

[85] Michał Jaroszyński. Gravitational lensing in a universe model with realistic mass distribution. *Monthly Notices of the Royal Astronomical Society*, 249(3):430–438, 04 1991.

[86] Nick Kaiser and Gordon Squires. Mapping the Dark Matter with Weak Gravitational Lensing. *The Astrophysical Journal*, 404:441, February 1993.

[87] William C. Keel, Rogier A. Windhorst, Rolf A. Jansen, Seth H. Cohen, Benne Holwerda, Sarah T. Bradford, Clayton D. Robertson, Giovanni Ferrami, Stuart Wyithe, Haojing Yan, Christopher J. Conselice, Simon P. Driver, Norman A. Grogin, Christopher N. A. Willmer, Anton M. Koekemoer, Brenda L. Frye, Nimish P. Hathi, Russell E. Ryan, Nor Pirzkal, Madeline A. Marshall, Dan Coe, Jose M. Diego, Thomas J. Broadhurst, Michael J. Rutkowski, Lifan Wang, S. P. Willner, Andreea Petric, Cheng Cheng, and Adi Zitrin. Webb's pearls: dust attenuation and gravitational lensing in the backlit-galaxy system vv 191, 2022.

[88] C. R. Keeton, C. S. Kochanek, and U. Seljak. Shear and ellipticity in gravitational lenses. *The Astrophysical Journal*, 482(2):604–620, 06 1997.

[89] Charles R. Keeton. GRAVLENS: Computational Methods for Gravitational Lensing. Astrophysics Source Code Library, record ascl:1102.003, 02 2011.

[90] Patrick L. Kelly, Jose M. Diego, Steven Rodney, Nick Kaiser, Tom Broadhurst, Adi Zitrin, Tommaso Treu, Pablo G. Perez-Gonzalez, Takahiro Morishita, Mathilde Jauzac, Jonatan Selsing, Masamune Oguri, Laurent Pueyo, Timothy W. Ross, Alexei V. Filippenko, Nathan Smith, Jens Hjorth, S. Bradley Cenko, Xin Wang, D. Andrew Howell, Johan Richard, Brenda L. Frye, Saurabh W. Jha, Ryan J. Foley, Colin Norman, Marusa Bradac, Weikang Zheng, Gabriel Brammer, Alberto Molino Benito, Antonio Cava, Lise Christensen, Selma E. de Mink, Or Graur, Claudio Grillo, Ryota Kawamata, Jean-Paul Kneib, Thomas Matheson, Curtis McCully, Mario Nonino, Ismael Perez-Fournon, Adam G. Riess, Piero Rosati, Kasper Borello Schmidt, Keren Sharon, and Benjamin J. Weiner. Extreme magnification of a star at redshift 1.5 by a galaxy-cluster lens, 2017.

[91] Patrick L. Kelly, Steven A. Rodney, Tommaso Treu, Ryan J. Foley, Gabriel Brammer, Kasper B. Schmidt, Adi Zitrin, Alessandro Sonnenfeld, Louis-Gregory Strolger, Or Graur, Alexei V. Filippenko, Saurabh W. Jha, Adam G. Riess, Marusa Bradac, Benjamin J. Weiner, Daniel Scolnic, Matthew A. Malkan, Anja von der Linden, Michele Trenti, Jens Hjorth, Raphael Gavazzi, Adriano Fontana, Julian C. Merten, Curtis McCully, Tucker Jones, Marc Postman, Alan Dressler, Brandon Patel, S. Bradley Cenko, Melissa L. Graham, and Bradley E. Tucker. Multiple images of a highly magnified supernova formed by an early-type cluster galaxy lens. *Science*, 347(6226):1123–1126, 2015.

[92] Jean-Paul Kneib, Henri Bonnet, Ghyslain Golse, David Sand, Eric Jullo, and Phil Marshall. LENSTOOL: A Gravitational Lensing Software for Modeling Mass Distribution of Galaxies and Clusters (strong and weak regime). Astrophysics Source Code Library, record ascl:1102.004, 02 2011.

[93] C. S. Kochanek. What do gravitational lens time delays measure? *The Astrophysical Journal*, 578(1):25–32, 10 2002.

[94] C S Kochanek. Overconstrained gravitational lens models and the Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 493(2):1725–1735, 02 2020.

[95] Christopher S. Kochanek and John Apostolakis. The two-screen gravitational lens. *Monthly Notices of the Royal Astronomical Society*, 235(4):1073–1109, 12 1988.

[96] L. V. E. Koopmans. Gravitational lensing & stellar dynamics: Dark-matter and baryons in early-type galaxies to z=1. *European Astronomical Society Publications Series*, 2004.

[97] L. V. E. Koopmans, A. Bolton, T. Treu, O. Czoske, M. W. Auger, M. Barnabè, S. Vegetti, R. Gavazzi, L. A. Moustakas, and S. Burles. The Structure and Dynamics of Massive Early-Type Galaxies: On Homology, Isothermality, and Isotropy Inside One Effective Radius. *Astrophysical Journal, Letters*, 703(1):L51–L54, 09 2009.

[98] L.V.E. Koopmans. Gravitational lensing & stellar dynamics. *EAS Publications Series*, 20:161–166, 2006.

[99] R. Kormann, P. Schneider, and M. Bartelmann. Isothermal elliptical gravitational lens models. *Astronomy & Astrophysics*, 284:285–299, 04 1994.

[100] Israel Kovner. The Thick Gravitational Lens: A Lens Composed of Many Elements at Different Distances. *Astrophysical Journal*, 316:52, 05 1987.

[101] Felix Arjun Kuhn, Simon Birrer, Claudio Bruderer, Adam Amara, and Alexandre Refregier. Combining strong and weak lensing estimates in the cosmos field. *Journal of Cosmology and Astroparticle Physics*, 2021(04):010, 04 2021.

[102] S. Rathna Kumar, C. S. Stalin, and T. P. Prabhu. H0 from ten well-measured time delay lenses. *Astronomy & Astrophysics*, 580:A38, 07 2015.

[103] Cameron A. Lemon, Matthew W. Auger, Richard G. McMahon, and Fernanda Ostrovski. Gravitationally lensed quasars in Gaia - II. Discovery of 24 lensed quasars. *Monthly Notices of the Royal Astronomical Society*, 479(4):5060–5074, 10 2018.

[104] Kai Liao, Tommaso Treu, Phil Marshall, Christopher D. Fassnacht, Nick Rumbaugh, Gregory Dobler, Amir Aghamousa, Vivien Bonvin, Frederic Courbin, Alireza Hojjati, Neal Jackson, Vinay Kashyap, S. Rathna Kumar, Eric Linder, Kaisey Mandel, Xiao-Li Meng, Georges Meylan, Leonidas A. Moustakas, Tushar P. Prabhu, Andrew Romero-Wolf, Arman Shafieloo, Aneta Siemiginowska, Chelliah S. Stalin, Hyungsuk Tak, Malte Tewes, and David van Dyk. Strong lens time delay challenge II. results of TDC1. *The Astrophysical Journal*, 800(1):11, 02 2015.

[105] J. Liesenborgs and S. De Rijcke. Lensing degeneracies and mass substructure. *Monthly Notices of the Royal Astronomical Society*, 425(3):1772–1780, 2012.

[106] Richard Luhtaru, Paul L. Schechter, and Kaylee M. de Soto. What makes quadruply lensed quasars quadruple? *The Astrophysical Journal*, 915(1):4, 06 2021.

[107] Shude Mao, Yipeng Jing, Jeremiah P. Ostriker, and Jochen Weller. Anomalous Flux Ratios in Gravitational Lenses: For or against Cold Dark Matter? *Astrophysical Journall*, 604(1):L5–L8, 03 2004.

[108] Curtis McCully, Charles R. Keeton, Kenneth C. Wong, and Ann I. Zabludoff. A new hybrid framework to efficiently model lines of sight to gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 443(4):3631–3642, 08 2014.

[109] Curtis McCully, Charles R. Keeton, Kenneth C. Wong, and Ann I. Zabludoff. Quantifying environmental and line-of-sight effects in models of strong gravitational lens systems. *The Astrophysical Journal*, 836(1):141, 02 2017.

[110] M. Meneghetti, M. Bartelmann, H. Dahle, and M. Limousin. Arc statistics. *Space Science Reviews*, 177(1-4):31–74, 05 2013.

[111] M. Millon, F. Courbin, V. Bonvin, E. Buckley-Geer, C. D. Fassnacht, J. Frieman, P. J. Marshall, S. H. Suyu, T. Treu, T. Anguita, V. Motta, A. Agnello, J. H. H. Chan, D. C.-Y. Chao, M. Chijani, D. Gilman, K. Gilmore, C. Lemon, J. R. Lucey, A. Melo, E. Paic, K. Rojas, D. Sluse, P. R. Williams, A. Hempel, S. Kim, R. Lachaume, and M. Rabus. TDCOSMO. *Astronomy & Astrophysics*, 642:A193, 10 2020.

[112] Jordi Miralda-Escude and Arif Babul. Gravitational lensing in clusters of galaxies: New clues regarding the dynamics of intracluster gas. *The Astrophysical Journal*, 449:18, aug 1995.

[113] Houjun Mo, Frank C. van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010.

[114] Ivelina Momcheva, Kurtis Williams, Charles Keeton, and Ann Zabludoff. A spectroscopic study of the environments of gravitational lens galaxies. *The Astrophysical Journal*, 641(1):169, 04 2006.

[115] Miguel F. Morales. A technique for weak lensing with velocity maps: Eliminating ellipticity noise in h i radio observations. *The Astrophysical Journal*, 650(1):L21, 09 2006.

[116] S. T. Myers, N. J. Jackson, I. W. A. Browne, A. G. de Bruyn, T. J. Pearson, A. C. S. Readhead, P. N. Wilkinson, A. D. Biggs, R. D. Blandford, C. D. Fassnacht, L. V. E. Koopmans, D. R. Marlow, J. P. McKean, M. A. Norbury, P. M. Phillips, D. Rusin, M. C. Shepherd, and C. M. Sykes. The cosmic lens all-sky survey - i. source selection and observations. *Monthly Notices of the Royal Astronomical Society*, 341(1):1–12, 05 2003.

[117] Chris Nagele, John D. Silverman, Tilman Hartwig, Junyao Li, Connor Bottrell, Xuheng Ding, and Yoshiki Toba. A machine learning approach to assessing the presence of substructure in quasar host galaxies using the hyper suprime-cam subaru strategic program, 2022.

[118] R. Nakajima, G. M. Bernstein, R. Fadely, C. R. Keeton, and T. Schrabback. Improved constraints on the gravitational lens q0957+561. i. weak lensing. *The Astrophysical Journal*, 697(2):1793, 05 2009.

[119] James. Nightingale, Richard Hayes, Ashley Kelly, Aristeidis Amvrosiadis, Amy Etherington, Qiuhan He, Nan Li, XiaoYue Cao, Jonathan Frawley, Shaun Cole, Andrea Enia, Carlos Frenk, David Harvey, Ran Li, Richard Massey, Mattia Negrello, and Andrew Robertson. PyAutoLens: Open-source strong gravitational lensing. *Journal of Open Source Software*, 6(58):2825, 02 2021.

[120] James W. Nightingale, Qiuhan He, Xiaoyue Cao, Aristeidis Amvrosiadis, Amy Etherington, Carlos S. Frenk, Richard G. Hayes, Andrew Robertson, Shaun Cole, Samuel Lange, Ran Li, and Richard Massey. Scanning for dark matter subhalos in hubble space telescope imaging of 54 strong lenses, 2022.

[121] Masamune Oguri. glafic: Software Package for Analyzing Gravitational Lensing. Astrophysics Source Code Library, record ascl:1010.012, 10 2010.

[122] Fernanda Ostrovski, Richard G. McMahon, Andrew J. Connolly, Cameron A. Lemon, Matthew W. Auger, Manda Banerji, Johnathan M. Hung, Sergey E. Koposov, Christopher E. Lidman, Sophie L. Reed, Sahar Allam, Aurélien Benoit-Lévy, Emmanuel Bertin, David Brooks, Elizabeth Buckley-Geer, Aurelio Carnero Rosell, Matias Carrasco Kind, Jorge Carretero, Carlos E. Cunha, Luiz N. da Costa, Shantanu Desai, H. Thomas Diehl, Jörg P. Dietrich, August E. Evrard, David A. Finley, Brenna Flaugher, Pablo Fosalba, Josh Frieman, David W. Gerdes, Daniel A. Goldstein, Daniel Gruen, Robert A. Gruendl, Gaston Gutierrez, Klaus Honscheid, David J. James, Kyler Kuehn, Nikolay Kuropatkin, Marcos Lima, Huan Lin, Marcio A. G. Maia, Jennifer L. Marshall, Paul Martini, Peter Melchior, Ramon Miquel, Ricardo Ogando, Andrés Plazas Malagón, Kevin Reil, Kathy Romer, Eusebio Sanchez, Basilio Santiago, Vic Scarpine, Ignacio Sevilla-Noarbe, Marcelle Soares-Santos, Flavia Sobreira, Eric Suchyta, Gregory Tarle, Daniel Thomas, Douglas L. Tucker, and Alistair R. Walker. VDES J2325-5229 a z = 2.7 gravitationally lensed quasar discovered using morphology-independent supervised machine learning. *Monthly Notices of the Royal Astronomical Society*, 465(4):4325–4334, 11 2016.

[123] Ewald Puchwein and Stefan Hilbert. Cluster strong lensing in the millennium simulation: the effect of galaxies and structures along the line-of-sight. *Monthly Notices of the Royal Astronomical Society*, 398(3):1298–1308, 2009.

[124] J. I. Read, P. Saha, and A. V. Macciò. Radial density profiles of time-delay lensing galaxies. *The Astrophysical Journal*, 667(2):645, 10 2007.

[125] S. Refsdal. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect. *Monthly Notices of the Royal Astronomical Society*, 128:307, 01 1964.

[126] S. Refsdal and S. Rosseland. On the Possibility of Determining the Distances and Masses of Stars from the Gravitational Lens Effect. *Monthly Notices of the Royal Astronomical Society*, 134(3):315–319, 12 1966.

[127] Rhea-Silvia Remus, Andreas Burkert, Klaus Dolag, Peter H. Johansson, Thorsten Naab, Ludwig Oser, and Jens Thomas. The dark halo—spheroid conspiracy and the origin of elliptical galaxies. *The Astrophysical Journal*, 766(2):71, 03 2013.

[128] Adam G. Riess, Stefano Casertano, Wenlong Yuan, Lucas M. Macri, and Dan Scolnic. Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the hubble constant and stronger evidence for physics beyond λcdm. *The Astrophysical Journal*, 876(1):85, 05 2019.

[129] Steven A. Rodney, Brandon Patel, Daniel Scolnic, Ryan J. Foley, Alberto Molino, Gabriel Brammer, Mathilde Jauzac, Maruša Bradač, Tom Broadhurst, Dan Coe, Jose M. Diego, Or Graur, Jens Hjorth, Austin Hoag, Saurabh W. Jha, Traci L. Johnson, Patrick Kelly, Daniel Lam, Curtis McCully, Elinor Medezinski, Massimo Meneghetti, Julian Merten, Johan Richard, Adam Riess, Keren Sharon, Louis-Gregory Strolger, Tommaso Treu, Xin Wang, Liliya L. R. Williams, and Adi Zitrin. Illuminating a dark lens: A type ia supernova magnified by the frontier fields galaxy cluster abell 2744. *The Astrophysical Journal*, 811(1):70, 09 2015.

[130] Aaron J. Romanowsky and Christopher S. Kochanek. Constraints on H0 from the central velocity dispersions of lens galaxies. *The Astrophysical Journal*, 516(1):18, 05 1999.

[131] D. Rusin and C. S. Kochanek. The evolution and structure of early-type field galaxies: A combined statistical analysis of gravitational lenses. *The Astrophysical Journal*, 623(2):666, 04 2005.

[132] Cristian E Rusu, Kenneth C Wong, Vivien Bonvin, Dominique Sluse, Sherry H Suyu, Christopher D Fassnacht, James H H Chan, Stefan Hilbert, Matthew W Auger, Alessandro Sonnenfeld, Simon Birrer, Frederic Courbin, Tommaso Treu, Geoff C-F Chen, Aleksi Halkola, Lé on V E Koopmans,

Philip J Marshall, and Anowar J Shajib. H0LiCOW XII. lens mass model of WFI2033 - 4723 and blind measurement of its time-delay distance and H0. *Monthly Notices of the Royal Astronomical Society*, 498(1):1440–1468, 09 2019.

[133] Prasenjit Saha. Lensing degeneracies revisited. *The Astronomical Journal*, 120(4):1654–1659, 10 2000.

[134] Prasenjit Saha, Jonathan Coles, Andrea V. Macciò, and Liliya L. R. Williams. The Hubble Time Inferred from 10 Time Delay Lenses. *Astrophysical Journal, Letters*, 650(1):L17–L20, 10 2006.

[135] Prasenjit Saha and Liliya L. R. Williams. PixeLens: A Portable Modeler of Lensed Quasars. Astrophysics Source Code Library, record ascl:1102.007, 02 2011.

[136] Peter Schneider. Generalized multi-plane gravitational lensing: time delays, recursive lens equation, and the mass-sheet transformation. *Astronomy & Astrophysics*, 624:A54, 04 2019.

[137] Peter Schneider, Jürgen Ehlers, and Emilio E. Falco. *Gravitational Lenses*. Springer Berlin, Heidelberg, 1992.

[138] Peter Schneider and Dominique Sluse. Mass-sheet degeneracy, power-law models and external convergence: Impact on the determination of the hubble constant from gravitational lensing. *Astronomy & Astrophysics*, 559:A37, 11 2013.

[139] Tim Schrabback, Stefan Hilbert, Henk Hoekstra, Patrick Simon, Edo van Uitert, Thomas Erben, Catherine Heymans, Hendrik Hildebrandt, Thomas D. Kitching, Yannick Mellier, Lance Miller, Ludovic Van Waerbeke, Philip Bett, Jean Coupon, Liping Fu, Michael J. Hudson, Benjamin Joachimi, Martin Kilbinger, and Konrad Kuijken. CFHTLenS: weak lensing constraints on the ellipticity of galaxy-scale matter haloes and the galaxy-halo misalignment. *Monthly Notices of the Royal Astronomical Society*, 454(2):1432–1452, 10 2015.

[140] N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J.-P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson. The cosmic evolution survey (COSMOS): Overview. *The Astrophysical Journal Supplement Series*, 172(1):1–8, 09 2007.

[141] Stella Seitz and Peter Schneider. Some remarks on multiple deflection gravitational lensing. *Astronomy & Astrophysics*, 287:349–360, 07 1994.

[142] Uros Seljak. Large-scale structure effects on the gravitational lens image positions and time delay. *The Astrophysical Journal*, 436:509, 12 1994.

[143] Atinç Çagan Sengül, Cora Dvorkin, Bryan Ostdiek, and Arthur Tsang. Substructure detection reanalysed: dark perturber shown to be a line-of-sight halo. *Monthly Notices of the Royal Astronomical Society*, 515(3):4391–4401, 07 2022.

[144] Anowar J. Shajib, Pritom Mozumdar, Geoff C. F. Chen, Tommaso Treu, Michele Cappellari, Shawn Knabel, Sherry H. Suyu, Vardha N. Bennert, Joshua A. Frieman, Dominique Sluse, Simon Birrer, Frederic Courbin, Christopher D. Fassnacht, Lizvette Villafaña, and Peter R. Williams. Tdcosmo. xiii. improved hubble constant measurement from lensing time delays using spatially resolved stellar kinematics of the lens galaxy, 2023.

[145] Alessandro Sonnenfeld. On the choice of lens density profile in time delay cosmography. *Monthly Notices of the Royal Astronomical Society*, 474(4):4648–4659, 12 2017.

[146] Volker Springel, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, John A. Peacock, Shaun Cole, Peter Thomas, Hugh Couchman, August Evrard, Jörg Colberg, and Frazer Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435(7042):629–636, 06 2005.

[147] Katherine A. Suess, Mariska Kriek, Sedona H. Price, and Guillermo Barro. Half-mass Radii of Quiescent and Star-forming Galaxies Evolve Slowly from $0 \lesssim z \leq 2.5$: Implications for Galaxy Assembly Histories. *Astrophysical Journal, Letters*, 885(1):L22, 11 2019.

[148] S. H. Suyu, M. W. Auger, S. Hilbert, P. J. Marshall, M. Tewes, T. Treu, C. D. Fassnacht, L. V. E. Koopmans, D. Sluse, R. D. Blandford, F. Courbin, and G. Meylan. Two accurate time-delay distances from strong lensing: implications for cosmology. *The Astrophysical Journal*, 766(2):70, 03 2013.

[149] S. H. Suyu, P. J. Marshall, M. W. Auger, S. Hilbert, R. D. Blandford, L. V. E. Koopmans, C. D. Fassnacht, and T. Treu. Dissecting the gravitational lens b1608+656. II. precision measurements of the hubble constant, spatial curvature, and the dark energy equation of state*. *The Astrophysical Journal*, 711(1):201, 02 2010.

[150] Ryuichi Takahashi and Kaiki Taro Inoue. Weak lensing by intergalactic ministructures in quadruple lens systems: simulation and detection. *Monthly Notices of the Royal Astronomical Society*, 440(1):870–888, 03 2014.

[151] Luca Teodori, Kfir Blum, Emanuele Castorina, Marko Simonović , and Yotam Soreq. Comments on the mass sheet degeneracy in cosmography analyses. *Journal of Cosmology and Astroparticle Physics*, 2022(07):027, 07 2022.

[152] Nicolas Tessore, Fabio Bellagamba, and R. Benton Metcalf. lensed: a code for the forward reconstruction of lenses and sources from strong lensing observations. *Monthly Notices of the Royal Astronomical Society*, 463(3):3115–3128, 09 2016.

[153] Nicolas Tessore and R. Benton Metcalf. The elliptical power law profile lens. *Astronomy & Astrophysics*, 580:A79, 08 2015.

[154] M. Tewes, F. Courbin, and G. Meylan. COSMOGRAIL: the COSmological MOnitoring of GRAv-Itational lenses. *Astronomy & Astrophysics*, 553:A120, 05 2013.

[155] The Dark Energy Survey Collaboration. The dark energy survey, 2005.

[156] S. S. Tie and C. S. Kochanek. Microlensing makes lensed quasar time delays significantly time variable. *Monthly Notices of the Royal Astronomical Society*, 473(1):80–90, 09 2017.

[157] O Tihhonova, F Courbin, D Harvey, S Hilbert, C E Rusu, C D Fassnacht, V Bonvin, P J Marshall, G Meylan, D Sluse, S H Suyu, T Treu, and K C Wong. H0LiCOW VIII. A weak-lensing measurement of the external convergence in the field of the lensed quasar HE 0435-1223. *Monthly Notices of the Royal Astronomical Society*, 477(4):5657–5669, 04 2018.

[158] T Treu, A Agnello, M A Baumer, S Birrer, E J Buckley-Geer, F Courbin, Y J Kim, H Lin, P J Marshall, B Nord, P L Schechter, P R Sivakumar, L E Abramson, T Anguita, Y Apostolovski, M W Auger, J H H Chan, G C F Chen, T E Collett, C D Fassnacht, J-W Hsueh, C Lemon, R G McMahon, V Motta, F Ostrovski, K Rojas, C E Rusu, P Williams, J Frieman, G Meylan, S H Suyu, T M C Abbott, F B Abdalla, S Allam, J Annis, S Avila, M Banerji, D Brooks, A Carnero Rosell, M Carrasco Kind, J Carretero, F J Castander, C B D'Andrea, L N da Costa, J De Vicente, P Doel, T F Eifler, B Flaugher, P Fosalba, J García-Bellido, D A Goldstein, D Gruen, R A Gruendl, G Gutierrez, W G Hartley, D Hollowood, K Honscheid, D J James, K Kuehn, N Kuropatkin, M Lima, M A G Maia, P Martini, F Menanteau, R Miquel, A A Plazas, A K Romer, E Sanchez, V Scarpine, R Schindler, M Schubnell, I Sevilla-Noarbe, M Smith, R C Smith, M Soares-Santos, F Sobreira, E Suchyta, M E C Swanson, G Tarle, D Thomas, D L Tucker, and A R Walker. The STRong lensing Insights into the Dark Energy Survey (STRIDES) 2016 follow-up campaign – I. Overview and classification of candidates selected by two techniques. *Monthly Notices of the Royal Astronomical Society*, 481(1):1041–1054, 08 2018.

[159] T. Treu and L. V. E. Koopmans. The internal structure of the lens PG1115+080: breaking degeneracies in the value of the Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 337(2):L6–L10, 12 2002.

[160] Tommaso Treu. Strong lensing by galaxies. *Annual Review of Astronomy and Astrophysics*, 48(1):87–125, 08 2010.

[161] M. A. Troxel, N. MacCrann, J. Zuntz, T. F. Eifler, E. Krause, S. Dodelson, D. Gruen, J. Blazek, O. Friedrich, S. Samuroff, J. Prat, L. F. Secco, C. Davis, A. Ferté, J. DeRose, A. Alarcon, A. Amara, E. Baxter, M. R. Becker, G. M. Bernstein, S. L. Bridle, R. Cawthon, C. Chang, A. Choi, J. De

Vicente, A. Drlica-Wagner, J. Elvin-Poole, J. Frieman, M. Gatti, W. G. Hartley, K. Honscheid, B. Hoyle, E. M. Huff, D. Huterer, B. Jain, M. Jarvis, T. Kacprzak, D. Kirk, N. Kokron, C. Krawiec, O. Lahav, A. R. Liddle, J. Peacock, M. M. Rau, A. Refregier, R. P. Rollins, E. Rozo, E. S. Rykoff, C. Sánchez, I. Sevilla-Noarbe, E. Sheldon, A. Stebbins, T. N. Varga, P. Vielzeuf, M. Wang, R. H. Wechsler, B. Yanny, T. M. C. Abbott, F. B. Abdalla, S. Allam, J. Annis, K. Bechtol, A. Benoit-Lévy, E. Bertin, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, D. L. DePoy, S. Desai, H. T. Diehl, J. P. Dietrich, P. Doel, E. Fernandez, B. Flaugher, P. Fosalba, J. García-Bellido, E. Gaztanaga, D. W. Gerdes, T. Giannantonio, D. A. Goldstein, R. A. Gruendl, J. Gschwend, G. Gutierrez, D. J. James, T. Jeltema, M. W. G. Johnson, M. D. Johnson, S. Kent, K. Kuehn, S. Kuhlmann, N. Kuropatkin, T. S. Li, M. Lima, H. Lin, M. A. G. Maia, M. March, J. L. Marshall, P. Martini, P. Melchior, F. Menanteau, R. Miquel, J. J. Mohr, E. Neilsen, R. C. Nichol, B. Nord, D. Petravick, A. A. Plazas, A. K. Romer, A. Roodman, M. Sako, E. Sanchez, V. Scarpine, R. Schindler, M. Schubnell, M. Smith, R. C. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, D. L. Tucker, V. Vikram, A. R. Walker, J. Weller, and Y. Zhang. Dark energy survey year 1 results: Cosmological constraints from cosmic shear. *Physical Review D*, 98(4), 08 2018.

[162] Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F Mota, Adam G Riess, and Joseph Silk. In the realm of the hubble tension—a review of solutions*. *Classical and Quantum Gravity*, 38(15):153001, 07 2021.

[163] Glenn Van De Ven, Rachel Mandelbaum, and Charles R. Keeton. Galaxy density profiles and shapes – I. Simulation pipeline for lensing by realistic galaxy models. *Monthly Notices of the Royal Astronomical Society*, 398(2):607–634, 09 2009.

[164] A. van der Wel, G. van de Ven, M. Maseda, H. W. Rix, G. H. Rudnick, A. Grazian, S. L. Finkelstein, D. C. Koo, S. M. Faber, H. C. Ferguson, A. M. Koekemoer, N. A. Grogin, and D. D. Kocevski. Discovery of a quadruple lens in CANDELS with a record lens redshift $z = 1.53$. *The Astrophysical Journal*, 777(1):L17, 10 2013.

[165] Ludovic Van Waerbeke, Takashi Hamana, Román Scoccimarro, Stephane Colombi, and Francis Bernardeau. Weak lensing predictions at intermediate scales. *Monthly Notices of the Royal Astronomical Society*, 322(4):918–926, 04 2001.

[166] S. Vegetti, L. V. E. Koopmans, A. Bolton, T. Treu, and R. Gavazzi. Detection of a dark substructure through gravitational imaging. *Monthly Notices of the Royal Astronomical Society*, 408(4):1969–1981, 10 2010.

[167] S. Vegetti, D. J. Lagattuta, J. P. McKean, M. W. Auger, C. D. Fassnacht, and L. V. E. Koopmans. Gravitational detection of a low-mass dark satellite galaxy at cosmological distance. *Nature*, 481(7381):341–343, 01 2012.

[168] D. Walsh, R. F. Carswell, and R. J. Weymann. 0957+561 A, B: twin quasistellar objects or gravitational lens? *Nature*, 279:381–384, 05 1979.

[169] Brian Welch, Dan Coe, Erik Zackrisson, S. E. de Mink, Swara Ravindranath, Jay Anderson, Gabriel Brammer, Larry Bradley, Jinmi Yoon, Patrick Kelly, Jose M. Diego, Rogier Windhorst, Adi Zitrin, Paola Dimauro, Yolanda Jiménez-Teja, Abdurro'uf, Mario Nonino, Ana Acebron, Felipe Andrade-Santos, Roberto J. Avila, Matthew B. Bayliss, Alex Benítez, Tom Broadhurst, Rachana Bhatawdekar, Maruša Bradač, Gabriel B. Caminha, Wenlei Chen, Jan Eldridge, Ebraheem Farag, Michael Florian, Brenda Frye, Seiji Fujimoto, Sebastian Gomez, Alaina Henry, Tiger Y.-Y Hsiao, Taylor A. Hutchison, Bethan L. James, Meridith Joyce, Intae Jung, Gourav Khullar, Rebecca L. Larson, Guillaume Mahler, Nir Mandelker, Stephan McCandliss, Takahiro Morishita, Rosa Newshore, Colin Norman, Kyle O'Connor, Pascal A. Oesch, Masamune Oguri, Masami Ouchi, Marc Postman, Jane R. Rigby, Russell E. Ryan Jr, Soniya Sharma, Keren Sharon, Victoria Strait, Louis-Gregory Strolger, F. X. Timmes, Sune Toft, Michele Trenti, Eros Vanzella, and Anton Vikaeus. Jwst imaging of earendel, the extremely magnified star at redshift z = 6.2. *The Astrophysical Journal Letters*, 940(1):L1, 11 2022.

[170] Olivier Wertz and Bastian Orthen. A dedicated source-position transformation package: pySPT. *Astronomy & Astrophysics*, 619:A117, 11 2018.

[171] P. R. Williams, A. Agnello, T. Treu, L. E. Abramson, T. Anguita, Y. Apostolovski, G. C. F. Chen, C. D. Fassnacht, J. W. Hsueh, B. C. Lemaux, V. Motta, L. Oldham, K. Rojas, C. E. Rusu, A. J. Shajib, and X. Wang. Discovery of three strongly lensed quasars in the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 477(1):L70–L74, 06 2018.

[172] Michelle L. Wilson, Ann I. Zabludoff, Charles R. Keeton, Kenneth C. Wong, Kurtis A. Williams, K. Decker French, and Ivelina G. Momcheva. A spectroscopic survey of the fields of 28 strong gravitational lenses: Implications for H0. *The Astrophysical Journal*, 850(1):94, 11 2017.

[173] David M. Wittman, J. Anthony Tyson, David Kirkman, Ian Dell'Antonio, and Gary Bernstein. Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405(6783):143–148, 05 2000.

[174] Kenneth C. Wong, C. R. Keeton, K. A. Williams, I. G. Momcheva, and A. I. Zabludoff. The Effect of Environment on Shear in Strong Gravitational Lenses. In *American Astronomical Society Meeting Abstracts #216*, volume 216 of *American Astronomical Society Meeting Abstracts*, page 318.05, 05 2010.

[175] Kenneth C. Wong, Charles R. Keeton, Kurtis A. Williams, Ivelina G. Momcheva, and Ann I. Zabludoff. The effect of environment on shear in strong gravitational lenses*. *The Astrophysical Journal*, 726(2):84, 12 2010.

[176] Kenneth C. Wong, Sherry H. Suyu, Matthew W. Auger, Vivien Bonvin, Frederic Courbin, Christopher D. Fassnacht, Aleksi Halkola, Cristian E. Rusu, Dominique Sluse, Alessandro Sonnenfeld, Tommaso Treu, Thomas E. Collett, Stefan Hilbert, Leon V. E. Koopmans, Philip J. Marshall, and Nicholas Rumbaugh. H0LiCOW – IV. lens mass model of HE 0435-1223 and blind measurement of its time-delay distance for cosmology. *Monthly Notices of the Royal Astronomical Society*, 465(4):4895–4913, 11 2016.

[177] Kenneth C Wong, Sherry H Suyu, Geoff C-F Chen, Cristian E Rusu, Martin Millon, Dominique Sluse, Vivien Bonvin, Christopher D Fassnacht, Stefan Taubenberger, Matthew W Auger, Simon Birrer, James H H Chan, Frederic Courbin, Stefan Hilbert, Olga Tihhonova, Tommaso Treu, Adriano Agnello, Xuheng Ding, Inh Jee, Eiichiro Komatsu, Anowar J Shajib, Alessandro Sonnenfeld, Roger D Blandford, Léon V E Koopmans, Philip J Marshall, and Georges Meylan. H0LiCOW – XIII. A 2.4 per cent measurement of H0 from lensed quasars: 5.3σ tension between early- and late-Universe probes. *Monthly Notices of the Royal Astronomical Society*, 498(1):1420–1439, 09 2019.

[178] O. Wucknitz. Degeneracies and scaling relations in general power-law models for gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 332(4):951–961, 06 2002.

[179] D. D. Xu, Shude Mao, Andrew P. Cooper, Liang Gao, Carlos S. Frenk, Raul E. Angulo, and John Helly. On the effects of line-of-sight structures on lensing flux-ratio anomalies in a λCDM universe. *Monthly Notices of the Royal Astronomical Society*, 421(3):2553–2567, 02 2012.

[180] C. K. Young and M. J. Currie. A new extragalactic distance indicator based on the surface brightness profiles of dwarf elliptical galaxies. *Monthly Notices of the Royal Astronomical Society*, 268:L11–L15, 05 1994.

[181] A. Yıldırım, S. H. Suyu, G. C. F. Chen, and E. Komatsu. TDCOSMO VIII: Cosmological distance measurements in light of the mass-sheet degeneracy – forecasts from strong lensing and ifu stellar kinematics, 2021.

[182] Hu Zhan and J Anthony Tyson. Cosmology with the large synoptic survey telescope: an overview. *Reports on Progress in Physics*, 81(6):066901, 04 2018.

[183] F. Zwicky. Nebulae as gravitational lenses. *Phys. Rev.*, 51:290–290, 02 1937.

Appendix A

Supplementary Derivations

A.1 Derivations for simple lens models

A.1.1 The point mass lens deflection angle

Here, we show that the deflection angle of a point mass lens eq. (1.133) follows from its gravitational potential eq. (1.132) and the expression for the deflection angle eq. (1.64).

Working in cylindrical coordinates, letting b be the distance of nearest approach from the mass to the unperturbed light ray (the "impact parameter"), and letting z represent the distance along the line of sight (i.e. $z = 0$ defines the lens plane, and $r^2 = z^2 + b^2$), we can rewrite eq. (1.132) as

$$\Phi(r) = -\frac{GM}{(b^2 + z^2)^{\frac{1}{2}}}. \tag{A.1}$$

Now, recall that $\nabla_\perp$ is the 2D gradient in the lens plane, or, in the case of a point mass, the gradient in the plane perpendicular to the line joining the source and the observer. Therefore

$$\nabla_\perp \Phi(r) = \frac{\partial \Phi(r)}{b}\hat{\boldsymbol{b}} = -\frac{GMb}{(b^2 + z^2)^{\frac{3}{2}}}\hat{\boldsymbol{b}}. \tag{A.2}$$

We can then plug this into equation 1.64 to obtain

$$\hat{\boldsymbol{\alpha}} = 2\int_{-\infty}^{\infty} dz\, \frac{GMb}{(b^2 + z^2)^{\frac{3}{2}}}\hat{\boldsymbol{b}}, \tag{A.3}$$

$$= \frac{2GM}{b}\int_{-\infty}^{\infty} \frac{dx}{(1 + x^2)^{\frac{3}{2}}}\hat{\boldsymbol{b}}, \tag{A.4}$$

$$= \frac{2GM}{b}\left[\frac{x}{(1 + x^2)^{\frac{1}{2}}}\right]_{-\infty}^{\infty}\hat{\boldsymbol{b}}, \tag{A.5}$$

$$= \frac{4GM}{b}\hat{\boldsymbol{b}}, \tag{A.6}$$

thus confirming eq. (1.133). Note that the contribution to the integral drops off rapidly as z grows, confirming the assumption that the majority of the deflection takes place close to the lens plane (and thus supporting the validity of the thin-screen approximation).

A.1.2 Singular isothermal sphere

Here, we show that the deflection angle for a SIS lens, eq. (1.141), can be derived from eq. (1.64) when the galaxy is modelled as a spherically symmetric ideal gas in thermal and hydrostatic equilibrium.

For an ideal gas, we have

$$p = \frac{\rho kT}{m}, \tag{A.7}$$

$$m\sigma_v^2 = kT, \tag{A.8}$$

where σ_v is the velocity dispersion of particle (stars etc) velocities around the mean velocity.

If we assume a state of hydrostatic equilibrium, we are assuming that the centrally-directed force of gravity is balanced by the outwardly-directed force of pressure. The equation of hydrostatic equilibrium is

$$\frac{\mathrm{d}p(r)}{\mathrm{d}r} = -\frac{GM(r)\rho(r)}{r^2}. \tag{A.9}$$

A simple mass distribution which fulfills these conditions is the single isothermal sphere,

$$\rho(r) = \frac{\sigma_v^2}{2\pi G}\frac{1}{r^2}. \tag{A.10}$$

To obtain the surface mass density, we must project this along the line of sight, to obtain

$$\Sigma(\boldsymbol{\xi}) = \int \mathrm{d}z\, \rho(\boldsymbol{\xi}, z), \tag{A.11}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}z\, \frac{\sigma_v^2}{2\pi G}\frac{1}{\xi^2 + z^2}, \tag{A.12}$$

$$= \frac{\sigma_v^2}{2\pi G\xi^2}\int_{-\infty}^{\infty} \frac{\mathrm{d}u}{1+u^2}. \tag{A.13}$$

Letting $u = \tan x$, $\mathrm{d}u = \sec^2 x \mathrm{d}x$, and using the fact that $\sec^2 x = 1 + \tan^2 x$, the integral simply becomes

$$\int \frac{du}{1+u^2} = \int \frac{\sec^2 x \mathrm{d}x}{1+\tan^2 x}, \tag{A.14}$$

$$= \int \mathrm{d}x, \tag{A.15}$$

$$= \arctan(u) + C. \tag{A.16}$$

Now, since $\tan x \to \infty$ at $x = +\pi/2$ and $\tan x \to -\infty$ at $x = -\pi/2$, we see that $\int_{-\infty}^{\infty} \mathrm{d}u/(1+u^2) = \pi$, and so (since $u = z/\xi$),

$$\Sigma(\boldsymbol{\xi}) = \frac{\sigma_v^2}{2G\xi}. \tag{A.17}$$

Since the gravitational field due to a circularly-symmetric mass distribution at some radius R is identical to that of a point mass at the centre of the distribution with a mass equal to that contained within the R, it follows that the deflection angle for such a distribution, which depends only on this gravitational field, will be the same as that for a point mass (eq. (1.133)), and so

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM(\xi)}{\xi}\hat{\boldsymbol{\xi}}, \tag{A.18}$$

where the mass enclosed within radius ξ is

$$M(\xi) = 2\pi\int_0^{\xi} \mathrm{d}\xi'\, \Sigma(\xi')\xi'. \tag{A.19}$$

Plugging the expression for $\Sigma(\xi)$, eq. (1.140), into this,

$$M(\xi) = 2\pi\int_0^{\xi} \mathrm{d}\xi'\, \frac{\sigma_v^2}{2G}, \tag{A.20}$$

$$= \frac{\pi\sigma_v^2\xi}{G}, \tag{A.21}$$

from which eq. (1.141) follows.

A.1.3 The singular isothermal ellipse

In the following, we derive the various quantities associated with the singular isothermal ellipse, as presented in section 1.1.5.

The normalisation of the surface-mass density

To understand the normalisation of eq. (1.149), consider the fact that an iso-density contour must have ζ constant. To find the total mass contained within this ellipse, we must integrate the surface mass density over an elliptical surface. To do this, we rearrange

$$\xi_2 = \pm\sqrt{\zeta^2 - f^2\xi_1^2}, \tag{A.22}$$

which gives us our limits of integration. The integral then takes the form

$$M(\zeta) = \frac{\sqrt{f}\sigma_v^2}{2G}\int_{-\zeta/f}^{\zeta/f}\int_{-\sqrt{\zeta^2-f^2\xi_1^2}}^{\sqrt{\zeta^2-f^2\xi_1^2}} \frac{\mathrm{d}\xi_1\mathrm{d}\xi_2}{\sqrt{f^2\xi_1^2+\xi_2^2}}. \tag{A.23}$$

Making the change of variables

$$\xi_1 = \frac{\zeta}{f}x, \tag{A.24}$$

$$\xi_2 = \zeta y, \tag{A.25}$$

such that

$$\mathrm{d}\xi_1 = \frac{\zeta}{f}\mathrm{d}x, \tag{A.26}$$

$$\mathrm{d}\xi_2 = \zeta\mathrm{d}y, \tag{A.27}$$

and, for the limits of integration,

$$\zeta/f \to 1, \tag{A.28}$$

$$\sqrt{\zeta^2 - f^2\xi_1^2} \to \sqrt{1-x^2}, \tag{A.29}$$

the integral simplifies to

$$M(\zeta) = \frac{\sqrt{f}\sigma_v^2}{2G}\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{\zeta}{f}\frac{\mathrm{d}x\mathrm{d}y}{\sqrt{x^2+y^2}}, \tag{A.30}$$

$$= C\frac{\zeta}{\sqrt{f}}, \tag{A.31}$$

$$\propto \frac{1}{\Sigma}. \tag{A.32}$$

Each value of the surface mass density Σ will correspond to an iso-density contour, whose distance from the centre of the lens, due to the ellipticity of the model, will vary with different angular positions. While the dimensions of this contour will depend on f, the results of the above integral show that the total mass contained within this contour is inversely proportional to Σ, and thus will be fixed for a given value of the surface density, irrespective of f.

Lensing potential, deflection angle and lens equation

Here, we show that the lensing potential eq. (1.153) and displacement angle eq. (1.154) of an SIE lens follow from the expression for the surface mass density eq. (1.149) and the definitions in eq. (1.152).

With the definition

$$\phi = \arctan\left(\frac{\xi_2}{\xi_1}\right), \tag{A.33}$$

we can write

$$\sqrt{f^2\xi_1^2+\xi_2^2} = \sqrt{f^2\xi^2\cos^2\phi + \xi^2\sin^2\phi}, \tag{A.34}$$

$$= \xi\sqrt{f^2\cos^2\phi + \sin^2\phi}, \tag{A.35}$$

$$= \xi\Delta(\phi), \tag{A.36}$$

where

$$\Delta(\phi) = \sqrt{f^2\cos^2\phi + \sin^2\phi}, \tag{A.37}$$

and so

$$\Sigma(\boldsymbol{\xi}) = \frac{\sqrt{f}\sigma_v^2}{2G}\frac{1}{\xi\Delta(\phi)}. \tag{A.38}$$

Now, the convergence of the SIE is given by

$$\kappa(x) = \frac{\Sigma(\xi_o\boldsymbol{\theta})}{\Sigma_{cr}}, \tag{A.39}$$

where

$$\boldsymbol{x} = \frac{\boldsymbol{\xi}}{\xi_0}, \tag{A.40}$$

$$\Sigma_{cr} = \frac{D_s}{4\pi G D_d D_{ds}}. \tag{A.41}$$

It then follows that

$$\kappa(x,\phi) = \frac{\sqrt{f}\sigma_v^2}{2G}\frac{1}{\xi_0\boldsymbol{x}\Delta(\phi)}\frac{4\pi G D_d D_{ds}}{D_s}, \tag{A.42}$$

$$= \frac{\sqrt{f}}{2b}\frac{1}{\xi_0}\frac{4\pi\sigma_v^2 D_d D_{ds}}{D_s}, \tag{A.43}$$

$$= \frac{\sqrt{f}}{2b}, \tag{A.44}$$

where we have introduced

$$b = \frac{\zeta}{\xi_0} = \boldsymbol{\theta}\Delta(\phi) = \sqrt{f^2x_1^2 + x_2^2}, \tag{A.45}$$

and have chosen the length scale

$$\xi_0 = \frac{D_s}{4\pi\sigma_v^2 D_d D_{ds}}. \tag{A.46}$$

The convergence is related to the lensing potential via Poisson's equation in eq. (1.85),

$$\nabla_\theta^2\psi = 2\kappa. \tag{A.47}$$

Since x and ϕ are the coordinates in the lens plane (i.e. $\theta = \theta(x,\phi)$), we can write this in polar coordinates as

$$\frac{1}{x}\frac{\partial}{\partial x}\left(x\frac{\partial\psi}{\partial x}\right) + \frac{1}{x^2}\frac{\partial^2\psi}{\partial\phi^2} = \frac{\sqrt{f}}{2x\Delta(\phi)}. \tag{A.48}$$

Using the method of separation of variables, we make the anzatz $\psi(x,\phi) = x\tilde{\psi}(\phi)$, and solving using Green's method,

$$\psi(x,\phi) = \frac{\sqrt{f}x}{f'}\left[\sin\phi\,\operatorname{arcsinh}\left(\frac{f'}{f}\sin\phi\right) + \cos\phi\,\arcsin\left(f'\cos\phi\right)\right], \tag{A.49}$$

$$= \frac{\sqrt{f}x}{f'}\left[|\sin\phi|\operatorname{arcosh}\frac{\Delta}{f} + |\cos\phi|\arccos\Delta\right], \tag{A.50}$$

where $f' = \sqrt{1-f^2}$. Taking the gradient of this, we obtain the deflection angle

$$\boldsymbol{\alpha} = \frac{\sqrt{f}}{f'}\left[\arcsin(f'\cos\phi)\boldsymbol{e_1} + \operatorname{arsinh}\left(\frac{f'}{f}\sin\phi\right)\boldsymbol{e_2}\right], \tag{A.51}$$

and thus the lens equation is

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{\sqrt{f}}{f'}\left[\arcsin(f'\cos\phi)\boldsymbol{e_1} + \operatorname{arsinh}\left(\frac{f'}{f}\sin\phi\right)\boldsymbol{e_2}\right]. \tag{A.52}$$

$\boldsymbol{e_1}$ and $\boldsymbol{e_2}$ are unit vectors in the directions of $\boldsymbol{\xi_1}$ and $\boldsymbol{\xi_2}$ respectively. Since

$$\lim_{f\to1}\left[\frac{\sqrt{f}}{f'}\arcsin(f'\cos\phi)\right] = \cos\phi, \tag{A.53}$$

$$\lim_{f\to1}\left[\frac{\sqrt{f}}{f'}\operatorname{arsinh}\left(\frac{f'}{f}\sin\phi\right)\right] = \sin\phi, \tag{A.54}$$

the lens equation for the SIE reduces to that of the SIS in the limit $f \to 1$.

A.2 Derivations in the DL approximation

A.2.1 The lens equation

The lens equation given in 1.218 can be split into the contributions for the foreground, dominant and background lenses

$$\boldsymbol{\theta} - \boldsymbol{\beta} = \boldsymbol{\alpha} = \sum_{l<d} \boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) + \boldsymbol{\alpha}_{ods}(\boldsymbol{\beta}_{od}) + \sum_{l>d} \boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}). \tag{A.55}$$

The angles $\boldsymbol{\beta}_{ol} = \boldsymbol{x}_l/D_{ol}$, in terms of which the deflection angles have been expressed, are related to each other via the recursion relation which follows from 1.219,

$$\beta_{ol} = \boldsymbol{\theta} - \sum_{m<l} \boldsymbol{\alpha}_{oml}(\boldsymbol{\beta}_{om}). \tag{A.56}$$

In the analysis which follows, the DL approximation will be used to expand equation 1.233, with the goal of arriving at an equation of $\boldsymbol{\theta}$ only.

Foreground displacement - For the lens located in front of the dominant lens, i.e. $l < d$, the difference between $\boldsymbol{\beta}_{ol}$ and $\boldsymbol{\theta}$ is independent of the main lens. This can be seen from equation 1.234. We therefore have $\boldsymbol{\beta}_{ol} = \boldsymbol{\theta} + \mathcal{O}(\epsilon^2)$ (as the sum of the $\boldsymbol{\alpha}_{oml}$ terms is $\mathcal{O}(\epsilon^2)$), and so, Taylor expanding, each partial displacement reads

$$\boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) = \boldsymbol{\alpha}_{ols}(\boldsymbol{\theta} + \mathcal{O}(\epsilon^2)) = \boldsymbol{\alpha}_{ols}(\boldsymbol{\theta}) + \frac{\mathrm{d}\boldsymbol{\alpha}_{ols}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\beta}_{ol}}\mathcal{O}(\epsilon^2) \tag{A.57}$$

Since the derivative of $\boldsymbol{\alpha}_{ols}$ is linearly related to κ_{ols}, which in turn is on the order ϵ^2 for $l \neq d$, we have

$$\boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) = \boldsymbol{\alpha}_{ols}(\boldsymbol{\theta}) + \mathcal{O}(\epsilon^4). \tag{A.58}$$

What this means is that the contribution to the deflection by foreground lenses can be evaluated looking backwards along the incoming light ray as it's observed at Earth, to linear order in ϵ^2. In other words, the Born approximation applies.

Main displacement - The main-lens displacement $\boldsymbol{\alpha}_{ods}$ is unique in that it is not $\mathcal{O}(\epsilon^2)$, and thus we must go beyond the born approximation for this term - in other words, we cannot evaluate the main lens at position $\boldsymbol{\theta}$. Setting $l = d$ in equation 1.234 and Taylor-expanding, and noting that, for $l < d$, we can apply equation A.58, we get

$$\boldsymbol{\alpha}_{ods}(\boldsymbol{\beta}_{od}) = \boldsymbol{\alpha}_{ods}\left[\boldsymbol{\theta} - \sum_{m<d} \boldsymbol{\alpha}_{omd}(\boldsymbol{\beta}_{om})\right], \tag{A.59}$$

$$= \boldsymbol{\alpha}_{ods}(\boldsymbol{\theta}) - \frac{\mathrm{d}\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\beta}_{od}} \sum_{m<d} \boldsymbol{\alpha}_{omd}(\boldsymbol{\theta}) + \mathcal{O}(\epsilon^4) + \frac{1}{2}\frac{\mathrm{d}^2\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\beta}_{od}^2}\left[\sum_{m<d} \boldsymbol{\alpha}_{omd}(\boldsymbol{\theta}) + \mathcal{O}(\epsilon^4)\right]^2 + ..., \tag{A.60}$$

$$= \boldsymbol{\alpha}_{ods}(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_{ods}(\boldsymbol{\theta}) \sum_{m<d} \boldsymbol{\alpha}_{omd}(\boldsymbol{\theta}) + \mathcal{O}(\epsilon^4). \tag{A.61}$$

We ignore higher order terms because the angles for $m < d$ must is order ϵ^2. This expansion is permissible, provided that $\boldsymbol{\theta} \mapsto \boldsymbol{\alpha}_{ods}(\boldsymbol{\theta})$ is differentiable, or in other words, provided that the partial derivatives encoded in $\boldsymbol{\Gamma}_{ods}(\boldsymbol{\theta})$ remain finite. This is just saying that $\mathrm{d}\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta})/\mathrm{d}\boldsymbol{\beta}_{od}$ must exist. This assumption breaks down near the centre of pathological lens models (e.g. point or SIS are non-differentiable at $\boldsymbol{\theta} = 0$). Note, however, they are not differentiable when the *image* is located at the exact centre. These locations in the image plane are practically irrelevant, as sources must be very far from the line of sight, and will have vanishing surface brightnesses.

Equation A.61 can be abbreviated by writing the perturbation within the argument, and introducing the notation $\boldsymbol{\alpha}_{od}$ as

$$\boldsymbol{\alpha}_{ods}(\boldsymbol{\beta}_{od}) = \boldsymbol{\alpha}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})] + \mathcal{O}(\epsilon^4), \qquad \boldsymbol{\alpha}_{od}(\boldsymbol{\theta}) \equiv \sum_{m<d} \boldsymbol{\alpha}_{omd}(\boldsymbol{\theta}). \tag{A.62}$$

Background displacements - here, the light ray is travelling from the source before crossing the dominant lens plane, and so the correction to $\boldsymbol{\theta}$ as the point of evaluation must include the contribution from this dominant lens,

$$\boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) = \boldsymbol{\alpha}_{ols}\left[\boldsymbol{\theta} - \sum_{m<l}\boldsymbol{\alpha}_{oml}(\boldsymbol{\beta}_{om})\right], \tag{A.63}$$

$$= \boldsymbol{\alpha}_{ols}\left[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\beta}_{om}) - \sum_{m<d}\boldsymbol{\alpha}_{oml}(\boldsymbol{\beta}_{om}) - \sum_{d<m<l}\boldsymbol{\alpha}_{oml}(\boldsymbol{\beta}_{om})\right]. \tag{A.64}$$

We then Taylor expand $\boldsymbol{\alpha}_{ols}$, with first two terms in the argument being the reference position. Using the same arguments as above, with the first derivative of α_{ols} as well as the $\boldsymbol{\alpha}_{oml}$ terms within the two sums being $\mathcal{O}(\epsilon^2)$, we obtain

$$\boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) = \boldsymbol{\alpha}_{ols}\left[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\beta}_{om})\right] + \mathcal{O}(\epsilon^4). \tag{A.65}$$

Because $\boldsymbol{\alpha}_{ols}$ is $\mathcal{O}(\epsilon^2)$, the $\boldsymbol{\alpha}_{odl}(\boldsymbol{\beta}_{om})$ term in its argument can be replaced with $\boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})$, as the difference between these is a sum of $\mathcal{O}(\epsilon^2)$ terms (equation A.61).

$$\boldsymbol{\alpha}_{ols}(\boldsymbol{\beta}_{ol}) = \boldsymbol{\alpha}_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})] + \mathcal{O}(\epsilon^4). \tag{A.66}$$

Combining these expressions for the background, foreground and dominant contributions to $\boldsymbol{\alpha}$, we obtain a lens equation which is a function of $\boldsymbol{\theta}$ only (eq. (1.236)).

A.2.2 The shear matrix

Here, we present the derivation of the shear matrix given in eq. (1.238), which is defined in eq. (1.237). This derivation is taken from [59]. Taking the derivative of the $\boldsymbol{\alpha}(\boldsymbol{\theta})$ in the DL approximation (equation 1.236),

$$\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_{ods} - \left(\sum_{l<d}\boldsymbol{\alpha}_{old}\cdot\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\right)\boldsymbol{\Gamma}_{ods} - \boldsymbol{\Gamma}_{ods}\sum_{l<d}\boldsymbol{\Gamma}_{old} + \sum_{l<d}\boldsymbol{\Gamma}_{ols} + \sum_{l>d}(\boldsymbol{\Gamma}_{ols} - \boldsymbol{\Gamma}_{ols}\boldsymbol{\Gamma}_{odl}) + \mathcal{O}(\epsilon^4), \tag{A.67}$$

where all the quantities are evaluated at $\boldsymbol{\theta}$, except the $\boldsymbol{\Gamma}_{ols}$ for $l > d$, which must be evaluated at $\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})$ to allow for the non-perturbative main displacement. The last term (before $\mathcal{O}(\epsilon^2)$) comes from the chain rule on the last term of equation 1.236

$$\sum_{l>d}\left[\frac{\mathrm{d}}{\mathrm{d}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta}))}\boldsymbol{\alpha}_{ols}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta}))\frac{d}{d\boldsymbol{\theta}}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\right]. \tag{A.68}$$

The first two terms in A.67 encode the direct contribution of the main deflector, evaluated at a position which is displaced by an amount

$$\boldsymbol{\alpha}_{od}(\boldsymbol{\theta}) \equiv \sum_{l<d}\boldsymbol{\alpha}_{old}(\boldsymbol{\theta}) \tag{A.69}$$

by the foreground lenses. The derivative of $\boldsymbol{\Gamma}_{ods}$ wrt $\boldsymbol{\theta}$ tells us how a small change in $\boldsymbol{\theta}$ affects $\boldsymbol{\Gamma}_{ods}$, and $\boldsymbol{\alpha}_{od}$ is the change in the light path relative to $\boldsymbol{\theta}$ caused by all the foreground lenses.. Because of this interpretation, we can combine these terms into $\boldsymbol{\Gamma}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})]$.

The third term encodes the non-linear coupling between the foreground (old) and main (ods) lenses, which, like the second term, relates physically to the departure from the Born approximation. The sum over l can be rewritten as

$$\boldsymbol{\Gamma}_{od}(\boldsymbol{\theta}) \equiv \sum_{l<d}\boldsymbol{\Gamma}_{old}(\boldsymbol{\theta}). \tag{A.70}$$

This shear matrix describes the distortions of an infinitesimal source located in the main deflector's plane due to all the foreground lenses.

The fourth and fifth terms are the direct contributions of the non-dominant (foreground and background) lenses, which may be gathered into

$$\boldsymbol{\Gamma}_{os}(\boldsymbol{\theta}) \equiv \sum_{l<d}\boldsymbol{\Gamma}_{ols}(\boldsymbol{\theta}) + \sum_{l>d}\boldsymbol{\Gamma}_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]. \tag{A.71}$$

Notice that, as stated previously, $\mathbf{\Gamma}_{ols}$ for the background lenses is evaluated at $\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})$, because the effect of the main lens is non-perturbative. This matrix would describe the distortions of an infinitesimal source in the source plane, due to all the lenses but the dominant one. However, because the background terms are evaluated along the main lens deflected path, $\mathbf{\Gamma}_{os}(\boldsymbol{\theta})$ is not truly independent of the main lens.

The last term is analogous to the third term, in that it encodes the non-linear couplings between the dominant and background lenses. $\mathbf{\Gamma}_{ols}$ is the distortion to a source at s as seen from o due to lens l, and $\mathbf{\Gamma}_{odl}$ is the distortion to a source in lens l due to the main lens. Using eqs. (1.223) and (1.224),

$$\mathbf{\Gamma}_{ols}\mathbf{\Gamma}_{odl} = \frac{\mathbf{\Sigma}_l}{\Sigma^{\rm crit}_{ols}}\frac{\mathbf{\Sigma}_d}{\Sigma^{\rm crit}_{odl}} = \frac{\mathbf{\Sigma}_l}{\Sigma^{\rm crit}_{dls}}\frac{\mathbf{\Sigma}_d}{\Sigma^{\rm crit}_{ods}} = \mathbf{\Gamma}_{dls}\mathbf{\Gamma}_{ods}. \tag{A.72}$$

The equivalence of the critical densities follows from their definition in equation 1.224,

$$\Sigma^{\rm crit}_{ols}\Sigma^{\rm crit}_{odl} = \left(\frac{4\pi G D_{ol} D_{ls}}{D_{os}}\right)^{-1}\left(\frac{4\pi G D_{od} D_{dl}}{D_{ol}}\right)^{-1}, \tag{A.73}$$

$$= \left(\frac{4\pi G D_{dl} D_{ls}}{D_{ds}}\right)^{-1}\left(\frac{4\pi G D_{od} D_{ds}}{D_{os}}\right)^{-1}, \tag{A.74}$$

$$= \Sigma^{\rm crit}_{dls}\Sigma^{\rm crit}_{ods}. \tag{A.75}$$

Equation (A.72) makes the interpretation of the final term as the non-linear couplings between the dominant and background lenses even more clear. $\mathbf{\Gamma}_{dls}$ characterises how an infinitesimal source at s would be seen at d, after being distorted by lens l, and $\mathbf{\Gamma}_{ods}$ characterises how an image which reaches the dominant lens d from the source plane s (which will have been distorted by the $\mathbf{\Gamma}_{dls}$ term) will be distorted, as seen from the observer o. Since the Born approximation applies between the main deflector and source plane, the $\mathbf{\Gamma}_{dls}$ terms can simply be added.

Another effect of this representation is that one of the terms ($\mathbf{\Gamma}_{ods}$) has no dependence on the background lenses, and may thus be factored out of the sum. The remainder may then be expressed as

$$\mathbf{\Gamma}_{ds}(\boldsymbol{\theta}) \equiv \sum_{l>d} \mathbf{\Gamma}_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})], \tag{A.76}$$

which describes the distortions of an infinitesimal source in the source plane, but as observed from the main-lens plane.

Combining these new expressions and definitions, we can write the shear matrix as

$$\mathbf{\Gamma}(\boldsymbol{\theta}) = \mathbf{\Gamma}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})] - \mathbf{\Gamma}_{ods}(\boldsymbol{\theta})\mathbf{\Gamma}_{od}(\boldsymbol{\theta}) - \mathbf{\Gamma}_{ds}(\boldsymbol{\theta})\mathbf{\Gamma}_{ods}(\boldsymbol{\theta}) + \mathbf{\Gamma}_{os}(\boldsymbol{\theta}). \tag{A.77}$$

To $\mathcal{O}(\epsilon^2)$, the arguments of the $\mathbf{\Gamma}_{ods}(\boldsymbol{\theta})$ terms can all be replaced with $\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})$, as each of these terms is multiplied by a shear matrix of non-dominant lens terms. We therefore arrive at eq. (1.238).

A.2.3 Convergence and shear

Here, we derive the results presented in [59] for the convergence and shear in the DL approximation, given by eqs. (1.248) and (1.250). In what follows, we will use the complex formalism introduced in subsection section 1.1.5. We start by rewriting $\mathbf{\Gamma} = \mathrm{d}\boldsymbol{\alpha}/\mathrm{d}\boldsymbol{\theta}$ in terms of the partial derivatives with respect to $\underline{\theta}$ and $\underline{\theta}^*$, using the expression from eq. (1.246),

$$\frac{\partial \underline{\alpha}}{\partial \underline{\theta}} = \frac{1}{2}\left(\frac{\partial}{\partial\theta_1} - \mathrm{i}\frac{\partial}{\partial\theta_2}\right)(\alpha_1 + \mathrm{i}\alpha_2), \tag{A.78}$$

$$= \frac{1}{2}\left(\frac{\partial\alpha_1}{\partial\theta_1} + \mathrm{i}\frac{\partial\alpha_2}{\partial\theta_1} - \mathrm{i}\frac{\partial\alpha_1}{\partial\theta_2} + \frac{\partial\alpha_2}{\partial\theta_2}\right), \tag{A.79}$$

$$= \frac{1}{2}\left(\kappa + \mathrm{Re}(\gamma) + \mathrm{i}(\mathrm{Im}(\gamma) + \omega) - \mathrm{i}(\mathrm{Im}(\gamma) - \omega) + \kappa - \mathrm{Re}(\gamma)\right), \tag{A.80}$$

$$= \kappa + \mathrm{i}\omega. \tag{A.81}$$

$$\frac{\partial \underline{\alpha}}{\partial \underline{\theta}^*} = \frac{1}{2}\left(\frac{\partial}{\partial \theta_1} + \mathrm{i}\frac{\partial}{\partial \theta_2}\right)(\alpha_1 + \mathrm{i}\alpha_2), \tag{A.82}$$

$$= \frac{1}{2}\left(\frac{\partial \alpha_1}{\partial \theta_1} + \mathrm{i}\frac{\partial \alpha_2}{\partial \theta_1} + \mathrm{i}\frac{\partial \alpha_1}{\partial \theta_2} - \frac{\partial \alpha_2}{\partial \theta_2}\right), \tag{A.83}$$

$$= \frac{1}{2}(\kappa + \mathrm{Re}(\gamma) + \mathrm{i}(\mathrm{Im}(\gamma) + \omega) + \mathrm{i}(\mathrm{Im}(\gamma) - \omega) - \kappa + \mathrm{Re}(\gamma)), \tag{A.84}$$

$$= \gamma. \tag{A.85}$$

We have the same results for the partial convergences and shears, except that there is no rotation.

$$\kappa_{ilj} = \frac{\partial \underline{\alpha}_{ilj}}{\partial \underline{\beta}_{ij}}, \tag{A.86}$$

$$\gamma_{ilj} = \frac{\partial \underline{\alpha}_{ilj}}{\partial \underline{\beta}^*_{ij}}. \tag{A.87}$$

We also define

$$\kappa_{od} \equiv \sum_{l<d} \kappa_{old}, \qquad \gamma_{od} \equiv \sum_{l<d} \gamma_{old}, \tag{A.88}$$

$$\kappa_{ds} \equiv \sum_{l>d} \kappa_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})], \qquad \gamma_{ds} \equiv \sum_{l>d} \gamma_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})], \tag{A.89}$$

$$\kappa_{os} \equiv \sum_{l<d} \kappa_{ols} + \sum_{l>d} \kappa_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})], \qquad \gamma_{os} \equiv \sum_{l<d} \gamma_{ols} + \sum_{l>d} \gamma_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}\boldsymbol{\theta})]. \tag{A.90}$$

Now, the complex counterpart of the displacement vector 1.236 is

$$\underline{\alpha}(\underline{\theta}) = \underline{\alpha}_{ods}[\underline{\theta} - \underline{\alpha}_{od}(\underline{\theta})] + \sum_{l<d} \underline{\alpha}_{ols}(\underline{\theta}) + \sum_{l>d} \underline{\alpha}_{ols}[\underline{\theta} - \underline{\alpha}_{odl}(\underline{\theta})]. \tag{A.91}$$

By taking the derivatives of this displacement vector, using the quantities defined in A.90 and the expressions derived for the two partial derivatives of $\underline{\alpha}$ in equations A.81 and A.85, we can obtain expressions for $\kappa + \mathrm{i}\omega$ and γ as follows.

$$\frac{\partial \underline{\alpha}}{\partial \underline{\theta}} = \frac{\partial \underline{\alpha}_{ods}(u)}{\partial u}\frac{\partial(\underline{\theta} - \underline{\alpha}_{od}(\underline{\theta}))}{\partial \underline{\theta}} + \frac{\partial \underline{\alpha}_{ods}(u)}{\partial u^*}\frac{\partial(\underline{\theta} - \underline{\alpha}_{od}(\underline{\theta}))^*}{\partial \underline{\theta}} + \sum_{l<d}\frac{\partial \underline{\alpha}_{ols}(\underline{\theta})}{\partial \underline{\theta}} \tag{A.92}$$

$$+ \sum_{l>d}\left[\frac{\partial \underline{\alpha}_{ols}(u)}{\partial u}\frac{\partial(\underline{\theta} - \underline{\alpha}_{odl}(\underline{\theta}))}{\partial \underline{\theta}} + \frac{\partial \underline{\alpha}_{ols}(u)}{\partial u^*}\frac{\partial(\underline{\theta} - \underline{\alpha}_{odl}(\underline{\theta}))^*}{\partial \underline{\theta}}\right], \tag{A.93}$$

$$\tag{A.94}$$

$$= \kappa_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})](1 - \kappa_{od}) + \gamma_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})]\left[-\frac{\partial \underline{\alpha}_{od}(\underline{\theta})^*}{\partial \underline{\theta}}\right] + \sum_{l<d}\kappa_{ols} \tag{A.95}$$

$$+ \sum_{l>d}\left\{\kappa_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})](1 - \kappa_{odl}) + \gamma_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\left[-\frac{\partial \underline{\alpha}_{odl}(\underline{\theta})^*}{\partial \underline{\theta}}\right]\right\}, \tag{A.96}$$

$$\tag{A.97}$$

$$= \kappa_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})](1 - \kappa_{od}) + \sum_{l<d}\kappa_{ols} + \sum_{l>d}\kappa_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})] \tag{A.98}$$

$$- \sum_{l>d}\kappa_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\kappa_{odl} - \gamma^*_{od}\gamma_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})] - \sum_{l>d}\gamma_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\gamma^*_{odl}. \tag{A.99}$$

Using the result from eq. (A.72),

$$\kappa_{ols}\kappa_{odl} = \kappa_{dls}\kappa_{ods}, \qquad \gamma_{ols}\gamma^*_{odl} = \gamma_{dls}\gamma^*_{ods}, \tag{A.100}$$

and recognising that, to $\mathcal{O}(\epsilon^2)$, $\kappa_{ds}\kappa_{ods} = \kappa_{ds}\kappa_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}]$ and $\gamma^*_{od}\gamma_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})] = \gamma^*_{od}\gamma_{ods}$, we arrive at

$$\frac{\partial\underline{\alpha}}{\partial\underline{\theta}} = \kappa_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})](1-\kappa_{od}) + \kappa_{os} - \sum_{l>d}\kappa_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\kappa_{ods} - \gamma^*_{od}\gamma_{ods} - \sum_{l>d}\gamma_{dls}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\gamma^*_{ods}, \tag{A.101}$$

$$= \kappa_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})](1-\kappa_{od}) - \kappa_{ds}\kappa_{ods} - \gamma^*_{od}\gamma_{ods} - \gamma_{ds}\gamma^*_{ods} + \kappa_{os}, \tag{A.102}$$

$$= (1-\kappa_{od}-\kappa_{ds})\kappa_{ods}(\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}) - (\gamma^*_{od}\gamma_{ods} + \gamma_{ds}\gamma^*_{ods}) + \kappa_{os}. \tag{A.103}$$

We have also used

$$\frac{\partial z^*}{\partial z} = \frac{1}{2}\left(\frac{\partial(x - \mathrm{i}y)}{\partial x} - \mathrm{i}\frac{\partial(x-\mathrm{i}y)}{\partial y}\right) = 0, \tag{A.104}$$

$$\frac{\partial z_1^*}{\partial z_2} = \left(\frac{\partial z_1}{\partial z_2^*}\right)^*. \tag{A.105}$$

Carrying out a similar calculation for the partial derivative with respect to $\underline{\theta}^*$, and with reference to A.81 and A.85, we arrive at the results in eq. (1.250).

A.3 Derivations in the tidal regime

A.3.1 The lens equation in the tidal regime

Applying eq. (1.256) for the partial displacement angles in the tidal regime to the lens equation eq. (1.236), evaluating these angles at reference position $\mathbf{0}$, we arrive at

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\alpha}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\boldsymbol{\theta})] + \sum_{l<d}\boldsymbol{\alpha}_{ols}(\boldsymbol{\theta}) + \sum_{l>d}\boldsymbol{\alpha}_{ols}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})], \tag{A.106}$$

$$= \boldsymbol{\alpha}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0}) - \boldsymbol{\Gamma}_{od}\boldsymbol{\theta}] + \sum_{l<d}\{\boldsymbol{\alpha}_{ols}(\mathbf{0}) + \boldsymbol{\Gamma}_{ols}\boldsymbol{\theta}\} + \sum_{l>d}\{\boldsymbol{\alpha}_{ols}(\mathbf{0}) + \boldsymbol{\Gamma}_{ols}\cdot[\boldsymbol{\theta} - \boldsymbol{\alpha}_{odl}(\boldsymbol{\theta})]\}, \tag{A.107}$$

$$= \boldsymbol{\alpha}_{ods}[\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0}) - \boldsymbol{\Gamma}_{od}\boldsymbol{\theta}] - \sum_{l>d}\boldsymbol{\Gamma}_{ols}\boldsymbol{\alpha}_{odl}(\boldsymbol{\theta}) + \sum_{l\neq d}\{\boldsymbol{\alpha}_{ols}(\mathbf{0}) + \boldsymbol{\Gamma}_{ols}\boldsymbol{\theta}\}, \tag{A.108}$$

$$= \boldsymbol{\alpha}_{ods}[(1-\boldsymbol{\Gamma}_{od})\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0})] - \sum_{l>d}\boldsymbol{\Gamma}_{dls}\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta}) + \boldsymbol{\alpha}_{os}(\mathbf{0}) + \boldsymbol{\Gamma}_{os}\boldsymbol{\theta}. \tag{A.109}$$

The last line follows from the fact that $\boldsymbol{\Gamma}_{ols}\boldsymbol{\alpha}_{odl} = \boldsymbol{\Gamma}_{dls}\boldsymbol{\alpha}_{ods}$, from a rearrangement of the distance factors as in A.75. $\boldsymbol{\alpha}_{os}$ is the displacement of an image due to all the lenses except the dominant one, as defined in eq. (1.260).

Our expression can be further simplified by recognising that, to order ϵ^2, $\boldsymbol{\Gamma}_{dls}\boldsymbol{\alpha}_{ods}(\boldsymbol{\theta}) = \boldsymbol{\Gamma}_{dls}\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta} - \boldsymbol{\alpha}_{od}(\mathbf{0})]$. We also have the definition $\sum_{l>d}\boldsymbol{\Gamma}_{dls} = \boldsymbol{\Gamma}_{ds}$, and so we arrive at eq. (1.259).

A.3.2 Equivalence of the approximate and general tidal regime time delay formulas

In the small shear approximation, eq. (1.339) reduces to eq. (1.329). To see this, consider

$$\boldsymbol{\mathcal{A}}_{od}^T\boldsymbol{\mathcal{A}}_{ds}^{-1}\boldsymbol{\mathcal{A}}_{os} = (\mathbf{1} - \boldsymbol{\Gamma}_{od})^T(\mathbf{1} - \boldsymbol{\Gamma}_{ds})^{-1}(\mathbf{1} - \boldsymbol{\Gamma}_{os}), \tag{A.110}$$

$$\approx (\mathbf{1} - \boldsymbol{\Gamma}_{od})(\mathbf{1} + \boldsymbol{\Gamma}_{ds})(\mathbf{1} - \boldsymbol{\Gamma}_{os}), \tag{A.111}$$

$$\approx \mathbf{1} - \boldsymbol{\Gamma}_{os} - \boldsymbol{\Gamma}_{od} + \boldsymbol{\Gamma}_{ds}, \tag{A.112}$$

$$\approx \mathbf{1} - \boldsymbol{\Gamma}_{\mathrm{LOS}}, \tag{A.113}$$

and

$$\boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta} = (\mathbf{1} - \boldsymbol{\Gamma}_{os})^{-1}(\boldsymbol{\theta} - \boldsymbol{\alpha}), \tag{A.114}$$

$$\approx (\mathbf{1} - \boldsymbol{\Gamma}_{os})^{-1}\left[\boldsymbol{\theta} - (\mathbf{1} - \boldsymbol{\Gamma}_{ds})\cdot\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}] - \boldsymbol{\Gamma}_{os}\boldsymbol{\theta}\right], \tag{A.115}$$

$$\approx (\mathbf{1} - \boldsymbol{\Gamma}_{os})^{-1}\left[(\mathbf{1} - \boldsymbol{\Gamma}_{os})\boldsymbol{\theta} - (\mathbf{1} - \boldsymbol{\Gamma}_{ds})\cdot\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}]\right], \tag{A.116}$$

$$\approx \boldsymbol{\theta} - (\mathbf{1} - \boldsymbol{\Gamma}_{os})^{-1}(\mathbf{1} - \boldsymbol{\Gamma}_{ds})\cdot\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}], \tag{A.117}$$

$$\approx \boldsymbol{\theta} - (\mathbf{1} + \boldsymbol{\Gamma}_{os} - \boldsymbol{\Gamma}_{ds})\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}]. \tag{A.118}$$

From line 4 in the above, we see that the definition $\boldsymbol{\alpha}'(\boldsymbol{\theta}) = \boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\mathcal{A}}_{ds}\boldsymbol{\alpha}_{ods}[(\mathbf{1} - \boldsymbol{\Gamma}_{od})\boldsymbol{\theta}]$ can be used in place of $\boldsymbol{\theta} - \boldsymbol{\mathcal{A}}_{os}^{-1}\boldsymbol{\beta}$, provided the version of $\boldsymbol{\alpha}$ used in the above is an appropriate approximation.

Appendix B

Measurements of the Foreground Shear and Main Lens Ellipticity

B.1 Measuring the foreground shear

As discussed in section 1.3.4, a key degeneracy exists between elliptical power law lens ellipticities and the foreground shear. This degeneracy can be seen in plots such as fig. 3.9, as well as in the fact that both the foreground shear and lens ellipticity are a lot more poorly constrained in fits. For an elliptical power law lens, this degeneracy is only first-order. However, for profiles with elliptical iso-potential contours, e and γ_{od} are fully degenerate. Any measurement of γ_{od} is plagued by the possibility that the main lens model has insufficient degrees of freedom to capture azimuthal distortions arising from the lens itself, and thus can never be trusted without additional constraints. Nonetheless, when investigating the specific question of the information contained in point source images and time delay data, it is of some interest to consider whether the degeneracy between an EPL lens and the foreground shear can at all be alleviated.

B.1.1 Measurements from extended sources

γ_{od} ***fitting results with extended sources***						
		E	**EP**	**EPS**	**EPF**	**EPA**
constraints	$p_{\mathcal{C}}$	82 %	87 %	89 %	90 %	89 %
precision	$p_{\sigma<5\%}$	10 %	8 %	9 %	9 %	8 %
	$p_{\sigma<20\%}$	28 %	26 %	26 %	27 %	28 %
accuracy	$p_{<\sigma}$	58 %	64 %	66 %	66 %	63 %
	$p_{>2\sigma}$	6 %	5 %	4 %	5 %	5 %
	χ^2	23.7	10.1	1.3	2.5	8.3

Table B.1: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source and extended source, for the various corresponding fit types described in table 3.1, for measurements of* γ_{od}.

When attempting to measure and constrain foreground shears using a lensed extended source in the absence of a visible point source, the fit could constrain parameters only 82 % of the time, a drop from 99 % in the case of $\gamma_{\rm LOS}$. When constraints were possible, these constraints were much poorer, and the resulting fits were very poor (with $\chi^2 > 20$). This is clearly visible when comparing the top rows of fig. B.1 and fig. 3.13, where we also see that it takes much higher SNR values to achieve tightly-constrained measurements of γ_{od} than of $\gamma_{\rm LOS}$. Even when these tighly-constrained measurements are achieved, the results in table B.1 make it clear that they cannot be trusted. Interestingly, while many of the clearly identifiable outliers in the $\gamma_{\rm LOS}$ case were found at low SNR values (fig. 3.13), it is clear from fig. B.1 that some of the most significant outliers are in fact seen with the highest quality images, thus clearly indicating the strength of the degeneracy.

B.1.2 The effects of adding point sources

γ_{od} *fitting results with point sources*						
		E	**P**	**PS**	**PF**	**PA**
constraints	$p_{\mathcal{C}}$	82 %	74 %	80 %	77 %	81 %
precision	$p_{\sigma<5\%}$	10 %	0 %	0 %	0 %	0 %
	$p_{\sigma<20\%}$	28 %	1 %	1 %	2 %	2 %
accuracy	$p_{<\sigma}$	58 %	72 %	77 %	75 %	76 %
	$p_{>2\sigma}$	6 %	0 %	0 %	0 %	1 %
	χ^2	23.7	0.19	0.27	0.24	0.29

Table B.2: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source, for the various corresponding fit types described in table 3.1, for measurements of γ_{od}.*

Unsurprisingly, when the extended source is replaced with a lensed point source, the quality of the measurements drop. Here, γ_{od} could only be constrained 74 % of the time, and constraints within 5% were never possible. Interestingly though, while the constraints are very poor, this 74 % of γ_{od} parameters which could be constrained is in fact higher than the 63 % of $\gamma_{\rm LOS}$ parameters which could be constrained by an identical set of images. When only quadruply-lensed point sources were considered, foreground shears could be constrained 89 % of the time, even more often than the 83 % of times in the case of the same lensing configurations, but with an extended source taking the place of the point source. However, these constraints remained poor.

As in the case of $\gamma_{\rm LOS}$ measurements, the χ^2, $p_{<\sigma}$ and $p_{>2\sigma}$ values indicate that uncertainties are significantly overestimated. The reasons for this are likely similar to the suggestions put forward in section 3.2.4.

In the case of images featuring lensing of extended and point sources, the ability of the code to constrain parameters improves somewhat compared to the extended source only case, but the tightness of these constraints do not, as can be seen in table B.1. The χ^2, $p_{<\sigma}$ and $p_{>2\sigma}$ values all suggest a slightly better fit, though the high χ^2 confirms that the fit is still poor. Interestingly, while the role of point sources in constraining outliers was strongest for the lowest SNR valued images in measurements of $\gamma_{\rm LOS}$, as can be seen in fig. 3.13, this improvement seems present even for the highest quality images in the case of γ_{od} measurements.

As was seen for $\gamma_{\rm LOS}$ measurements, these measurements became more poorly constrained when only considering images featuring quadruply-lensed point sources and extended sources, when compared to extended sources only. Once again, this is thought to result from the increased degeneracy between the point source and host galaxy light, and the dominant effect of the PSF over the noise level for these high SNR images.

B.1.3 The effects of adding time delays

The effects of adding time delay data and sampling τ_{ds} to measurements of foreground shears from point sources only is similar to its effect when measuring $\gamma_{\rm LOS}$. For point sources, when time delay data are added, $p_{\mathcal{C}}$ becomes very similar to that in the case where extended sources only are present. Constraints, however, remained very poor, and as before, χ^2, $p_{<\sigma}$ and $p_{>2\sigma}$ each suggest that errors are overestimated.

The situation changes somewhat when only those lensing situations which would lead to quadruply-lensed point source images are considered. Here, lensed point sources with time delays allow the foreground shear to be constrained in 95 % of cases, thus offering a notable improvement on both the extended source only and quadruple point source image only cases.

The addition of time delay data and the choice to sample τ_{ds} for images which also feature an extended source seems to almost completely remove the presence of outliers from measurements of γ_{od}. While improvements to $p_{\mathcal{C}}$ and measures of the precision are slight or non-existent, the χ^2 value drops to 1.3, indicating a good fit with well-defined error bars. This is a dramatic improvement from the cases

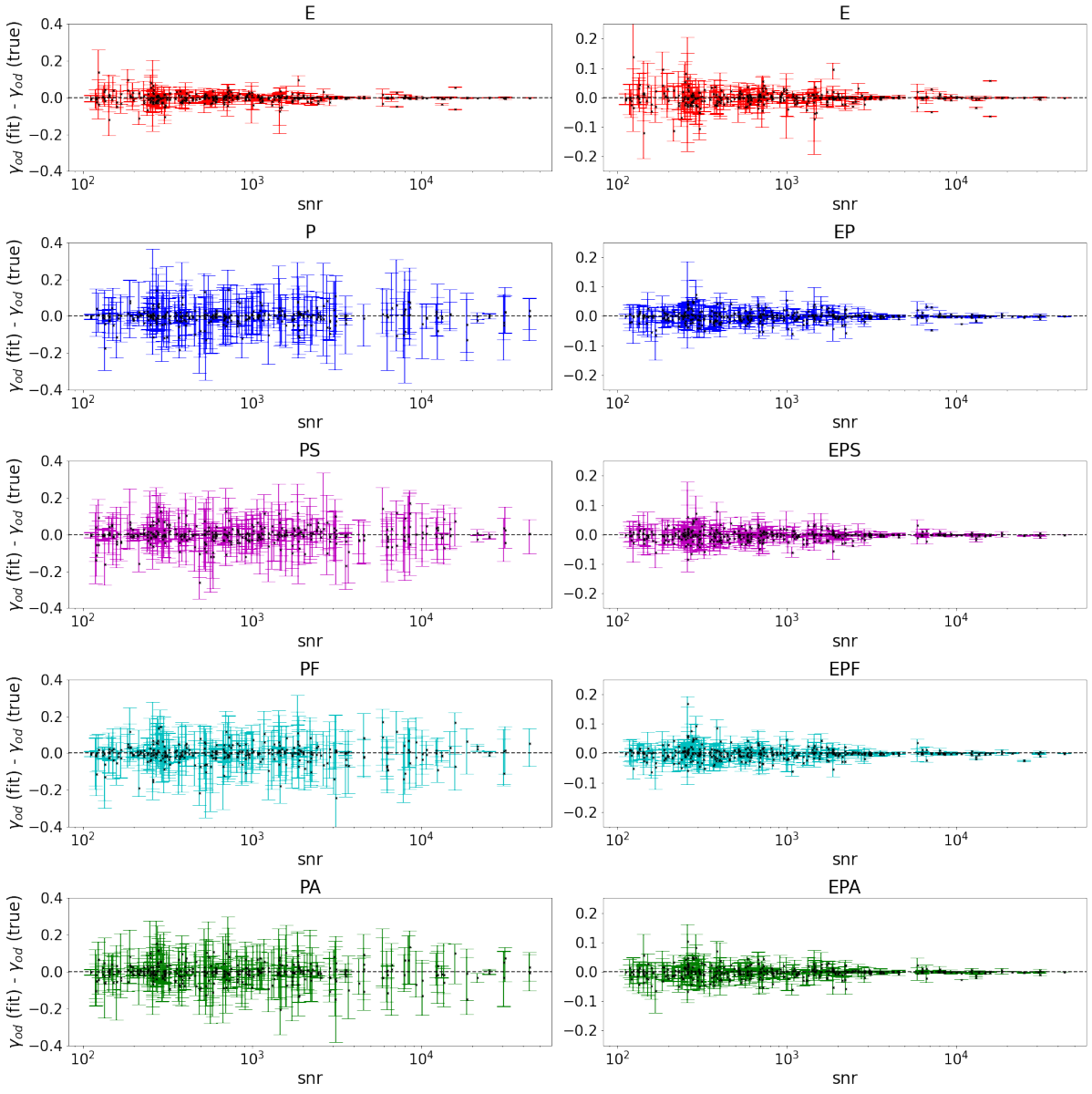

Figure B.1: *Accuracy and precision of γ_{od} measurements for the fit types in table 3.1. Measured values as plotted as black crosses, with 1σ errorbars also shown.*

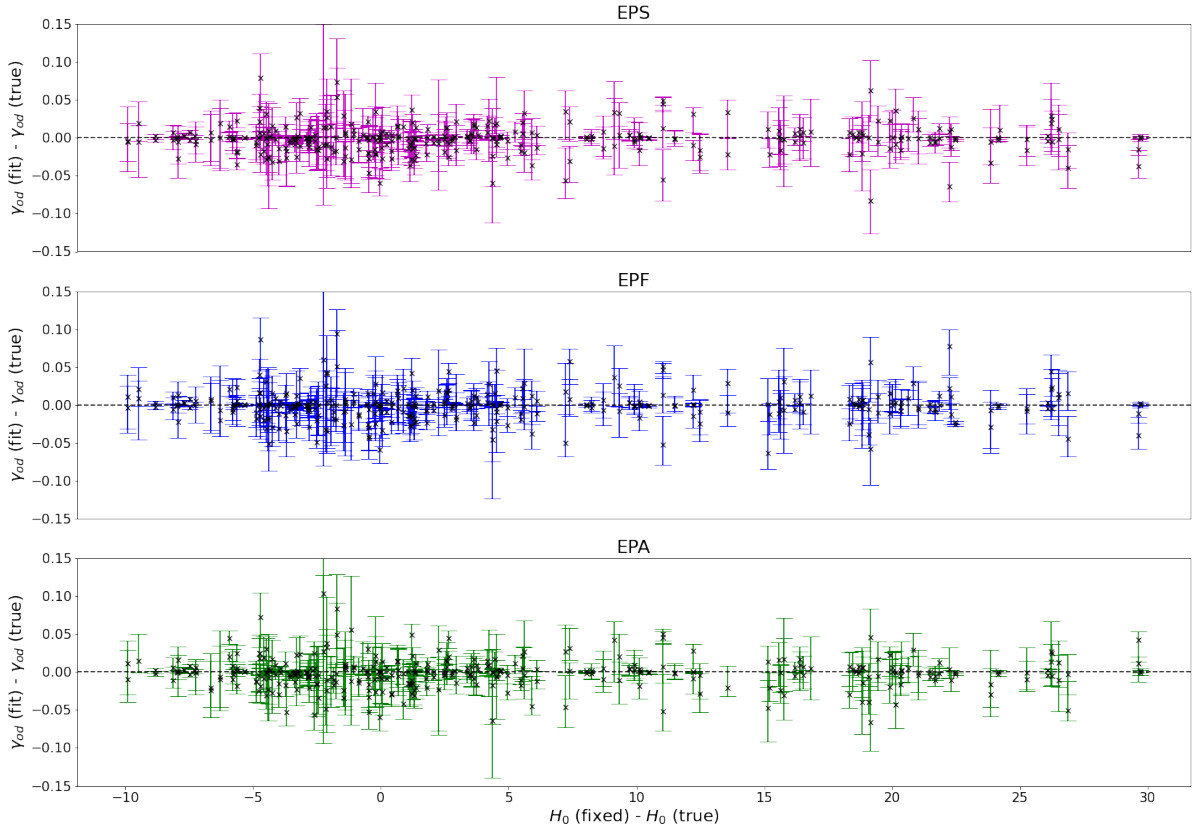

Figure B.2: *The effect of fixing H_0 on the accuracy and precision of γ_{od} measurements. In the EPS case (top), the time delay distance is sampled as a free parameter in the model. In the EPF case (centre), it is fixed to the "true" value of 70 km/s/Mpc, used when creating the mock data, and in the EPA case (bottom), it is fixed to a value selected randomly between 60 and 100 km/s/Mpc. The x-axis of the EPS and EPF plot is selected so that measurements from the same mock data align vertically with those in the other subplots. From this plot, no clear bias seems to be introduced by the choice to fix H_0, even arbitrarily.*

where time delay data are omitted, which, considering that changes to $p_{<\sigma}$ and $p_{>2\sigma}$ are slight, most likely arises from the fact that the scattering of larger outliers seen in the E and EP cases are no longer present in the EPS fits, as can be seen in fig. B.1.

In a very interesting departure from the results of the $\gamma_{\rm LOS}$ fitting, fixing τ_{ds} when measuring the foreground shear from images with both an extended and a point source present made a noticeable difference to measurements of γ_{od}. While changes to other parameters are slight, the χ^2 value more than doubles to 2.5, indicating that there is an underestimation of the uncertainties in the reported results. When τ_{ds} is fixed incorrectly, the χ^2 value increases to 8.3, seemingly suggesting a much lower quality fit. However, this effect is not seen at all in fig. B.2, where no systematic effects of introducing an incorrectly fixed value of H_0 via τ_{ds} can be clearly seen.

In conclusion, therefore, it seems that point sources and time delays can offer some slight advantages when it comes to ensuring that foreground shears to be constrained, especially when quadruply-imaged point sources are present, and, relative to the same measurements with extended sources, strongly lensed point source images are more useful when measuring γ_{od} than $\gamma_{\rm LOS}$. Point sources images, even with time delay data included, do not in general offer any improvements to the precision of foreground shear measurements over images of extended sources only, except when the signal-to-noise ratio is low, but can help to ensure that constraints placed on γ_{od} are more realistic, and improve the quality of the fit. A lensed extended source with a point source at its centre, when time delay data are included at τ_{ds} is sampled as a free parameter, seems the best class of source for properly accounting for the foreground shear - ellipticity degeneracy when attempting to constrain the foreground shear.

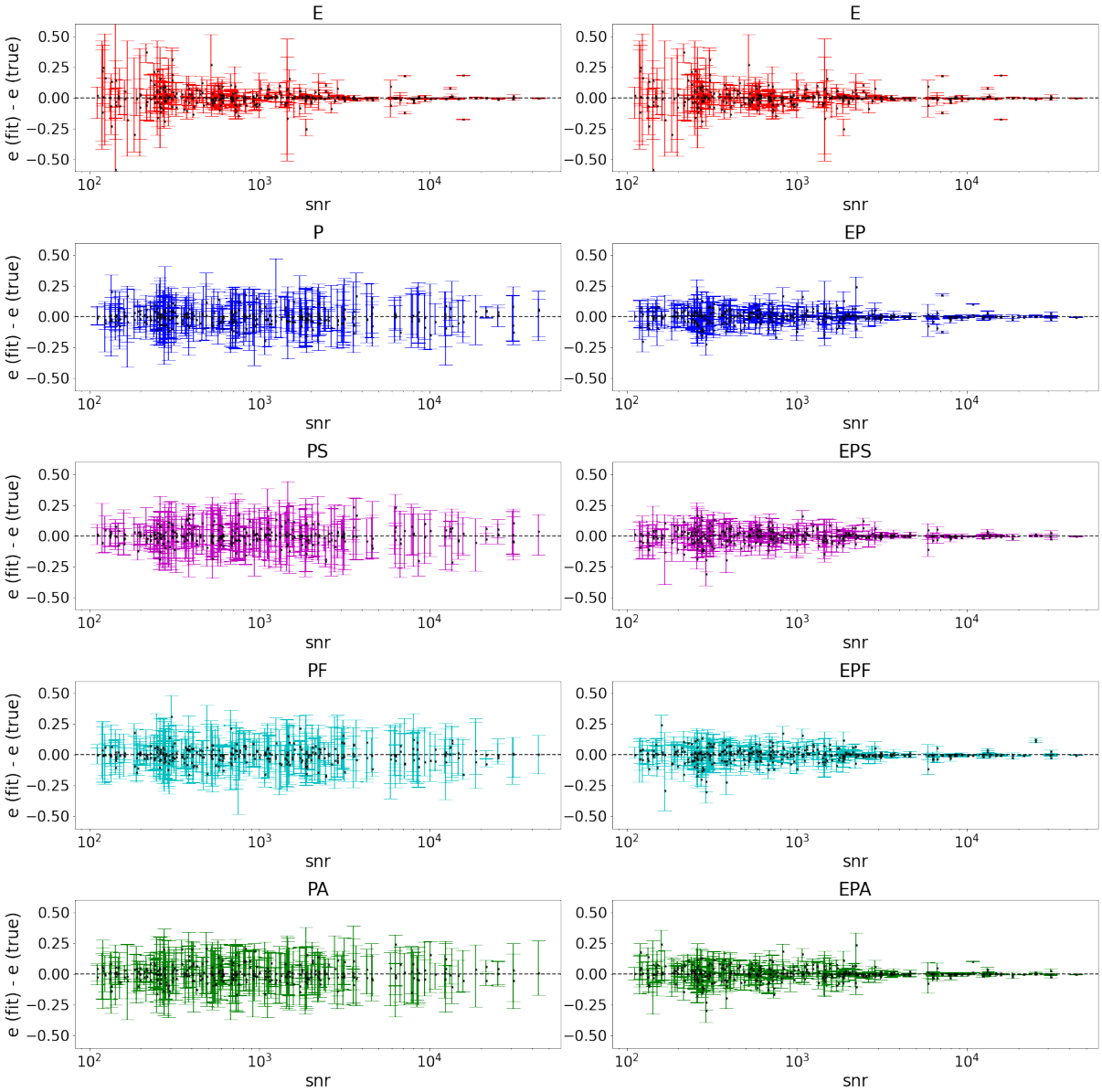

Figure B.3: *Accuracy and precision of e measurements for the fit types in table 3.1. Measured values as plotted as black crosses, with 1σ errorbars also shown.*

B.2 Measuring the lens ellipticity

The effects of different lensing datasets on measurements of the main lens ellipticity are almost identical to those described for measurements of foreground shear. As before, point sources alone were less likely than extended sources to constrain ellipticities, although the discrepancy was not as large as in the case of $\gamma_{\rm LOS}$, and quadruply-lensed point sources with time delay data in fact improved on extended sources only (constraining parameters in 93 % of cases, versus 80 % for equivalent lenses with extended sources only). Even more interestingly, quadruply-imaged point sources alone were able to constrain the lens ellipticity more often even than the equivalent extended source + point source combination. Unsurprisingly, these constraints were very poor, and extended sources offered somewhat greater precision. However, as can be seen in fig. B.3, the reported errors are large in every category of lens event, and the signal-to-noise ratio above which having an extended source present offers significant improvement to the constraints is higher than in the case of either shear measurement. When point sources or point sources and time delays are present in addition to an extended source, parameters are once again more likely to be constrained than when only an extended source is present, but the tightness of these constraints does not seem to improve versus the extended source only case. Referring to fig. B.3, we again see that it is at lower signal-to-noise

ratios where the presence of point sources is most valuable in providing tighter constraints, but at higher SNR values where they are best able to eliminate poorly-defined error bars. For point sources both with

e ***fitting results with extended sources***						
		E	**EP**	**EPS**	**EPF**	**EPA**
constraints	$p_{\mathcal{C}}$	78 %	86 %	86 %	87 %	88 %
precision	$p_{\sigma<5\%}$	5 %	3 %	3 %	3 %	3 %
	$p_{\sigma<20\%}$	17 %	13 %	13 %	13 %	13 %
accuracy	$p_{<\sigma}$	54 %	64 %	64 %	65 %	65 %
	$p_{>2\sigma}$	6 %	6 %	4 %	5 %	5 %
	χ^2	25.2	9.7	1.2	2.4	7.5

Table B.3: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed extended and point source, for the various corresponding fit types described in table 3.1, for measurements of e.*

e ***fitting results with point sources***						
		E	**P**	**PS**	**PF**	**PA**
constraints	$p_{\mathcal{C}}$	78 %	75 %	73 %	77 %	76 %
precision	$p_{\sigma<5\%}$	5 %	0 %	0 %	0 %	0 %
	$p_{\sigma<20\%}$	17 %	0 %	0 %	0 %	0 %
accuracy	$p_{<\sigma}$	54 %	72 %	69 %	74 %	72 %
	$p_{>2\sigma}$	6 %	0 %	0 %	1 %	1 %
	χ^2	25.2	0.18	0.24	0.19	0.21

Table B.4: *The results of carrying out MCMC fits on images featuring a lensed extended source only ("E") and a lensed point source, for the various corresponding fit types described in table 3.1, for measurements of e.*

and without extended sources, the addition of time delays had no real effect whatsoever to the ability of the code to constrain parameters or the tightness of these constraints. However, as in the case when measuring γ_{od}, there seems to be a suppression of outliers which greatly improves the χ^2 parameter in the EPS case. Just as for γ_{od}, the χ^2 parameter deteriorates when τ_{ds} is fixed incorrectly. Outliers are definitely visible in fig. B.4 in the EPF and EPA cases which are absent in the EPS plot, as is also apparent from fig. B.3, but, from the EPA plot, there do not appear to be any systematic biases introduced by fixing a value of H_0 far from the true value.

In conclusion, accurate and precise measurements of the ellipticity are very hard to achieve, and beset by the degeneracy discussed in section 1.3.4. Constraints from images featuring extended sources only are generally very poor, except at high signal-to-noise ratios, but for these higher-snr images, reported constraints cannot be trusted. Generally speaking, images featuring point source images only provide accurate and reliable constraints, but precise measurements are never possible. Images featuring both an extended and point source, in which time delay data are present and τ_{ds} sampled, seem to provide the best-defined constraints on e, and at high signal-to-noise ratios, good constraints are possible. The crucial caveat on this final point, however, is that this rests on the assumption that the true profile of the main lens is well approximated by an elliptical power law, an assumption which, in real observations, is almost never valid. Nonetheless, selecting images featuring point sources as well as an extended host galaxy, including any available time delay data in the model and sampling τ_{ds} seems the best method for curtailing biases introduced by degeneracies between the foreground shear and the profile of the main lens.

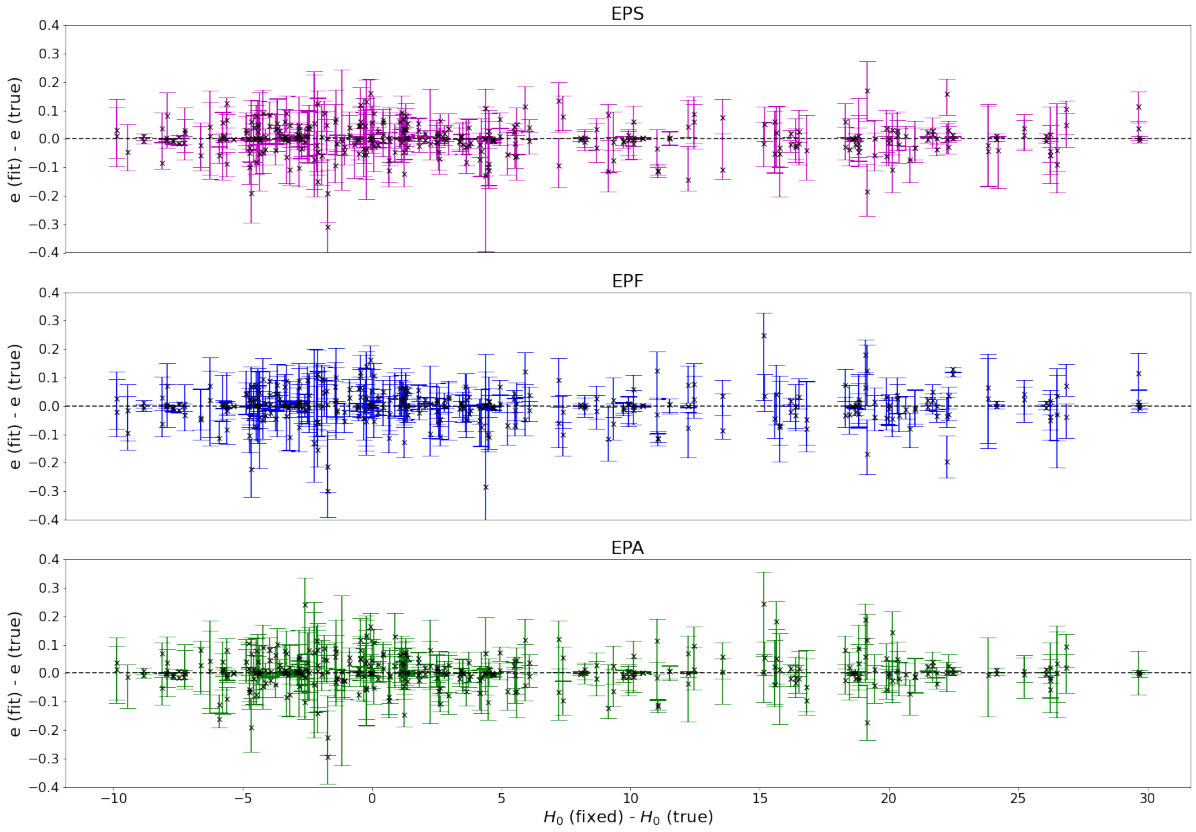

Figure B.4: *The effect of fixing H_0 on the accuracy and precision of e measurements. As in fig. B.2, the x-axis of the EPS and EPF plot is selected so that measurements from the same mock data align vertically with those in the other subplots, and once again, no clear bias seems to be introduced by the choice to fix H_0, even arbitrarily.*