

# Source-position transformation – an approximate invariance in strong gravitational lensing

Peter Schneider<sup>1</sup> & Dominique Sluse<sup>1</sup>

Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany  
 peter, dsluse@astro.uni-bonn.de

November 17, 2021

## ABSTRACT

The main obstacle for gravitational lensing to determine accurate masses of deflectors, or to determine precise estimates for the Hubble constant, is the degeneracy of lensing observables with respect to the mass-sheet transformation (MST). The MST is a global modification of the mass distribution which leaves all image positions, shapes and flux ratios invariant, but which changes the time delay. Here we show that another global transformation of lensing mass distributions exists which almost leaves image positions and flux ratios invariant, and of which the MST is a special case. Whereas for axi-symmetric lenses this source position transformation exactly reproduces all strong lensing observables, it does so only approximately for more general lens situations. We provide crude estimates for the accuracy with which the transformed mass distribution can reproduce the same image positions as the original lens model, and present an illustrative example of its performance. This new invariance transformation most likely is the reason why the same strong lensing information can be accounted for with rather different mass models.

**Key words.** cosmological parameters – gravitational lensing: strong

## 1. Introduction

Multiple image systems in strong gravitational lensing systems provide an invaluable tool for the determination of mass properties of cosmic objects, specifically of galaxies and galaxy clusters (see, e.g., Kochanek 2006; Bartelmann 2010 and references therein). The determination of the mass inside the Einstein radius of a multiple image system is the most accurate mass measurement available for galaxies and cluster cores.

Mass estimates at larger and smaller radii are, however, less accurate. The reason for this is that a given system of multiple images can be fitted by more than one mass model – i.e., the mass model is not unique. For example, a four-image system provide a total of six positional constraints on the lensing mass distribution, and many different density profiles can satisfy these constraints. The variety of mass distribution which can reproduce a set of lensed images can be seen by adding angular structures to the lens potential (Trotter et al. 2000; Evans & Witt 2003), or by modelling the mass distribution with a grid of variable pixels or a sum of basis functions (e.g. Saha & Williams 1997; Diego et al. 2005; Coe et al. 2008; Liesenborgs & De Rijcke 2012). Therefore, a finite set of individual lensed compact images clearly cannot uniquely determine the lensing mass distribution.

When extended source components are lensed, e.g., into a partial or full Einstein ring, the constraints on the lens model become considerably stronger. Here, a point-by-point modification of a mass model (like in the LensPerfect code of Coe et al. 2008) no longer can be used to fit the observed brightness profile with a lens model. However, as was pointed out first by Falco et al. (1985), even in this case the mass model is not unique; there exists a transformation of the mass distribution, called mass-sheet transformation (MST), which leaves all image positions and image flux ratios invariant. If  $\kappa(\theta)$  denotes the dimensionless sur-

face mass density of the lens at angular position  $\theta$ , then the whole family of mass models

$$\kappa_\lambda(\theta) = \lambda\kappa(\theta) + (1 - \lambda) \quad (1)$$

predicts the same imaging properties as the original mass profile  $\kappa(\theta)$ . The MST keeps the mass inside the Einstein radius invariant, but changes the enclosed mass at all other radii. Furthermore, the MST changes the predicted product of time delay between images and the Hubble constant,  $\tau = H_0 \Delta t$ , to  $\tau_\lambda = \lambda\tau$ , for all pairs of images. As we pointed out in Schneider & Sluse (2013; hereafter SS13), this MST may strongly impact on the ability to use time-delay lens systems for accurate determinations of the Hubble constant.

In SS13, we furthermore considered an illustrative case where a composite lens, consisting of a Hernquist profile to resemble the distribution of stellar mass in a lens galaxy, plus a modified Navarro, Frank and White profile for the description of the dark matter in the inner part of the galaxy, yields almost identical imaging properties as a power-law mass profile. The relation between these two mass models is not described by a MST; in particular, we found that the time delay ratios of image pairs between these two models are not constant, as would be predicted from a MST. It thus appeared as if there exists a more general transformation between lensing mass models which leaves observed image positions almost invariant. Hints of the existence of such a transformation were pointed out earlier by several authors (Saha and Williams 2006; Read et al. 2007, their Appendix A3; Coe et al. 2008, their Sect. 3.4), but to our knowledge it has never been identified as a transformation of the source plane or derived explicitly.

In this paper, we will show the existence of such a transformation, i.e., a transformation of the deflection law which leaves the strong lensing properties invariant not only for a finite set of

source positions, but for all source positions at the same time; for reasons that will become obvious in the following, we call it the source-position transformation (SPT). The general concept of the SPT is outlined in Sect. 2. We will then show in Sect. 3 that for axi-symmetric lenses, the SPT is indeed an exact invariance transformation which leaves all relative image positions and flux ratios invariant. Thus, there exists a much larger set of mass models than described by the MST which lead to the same strong lensing predictions as the original mass distribution. We then turn to the more general case in Sect. 4 and show that the lens models obtained through an SPT in general lead to different imaging properties, but that these differences can be quite small in realistic cases. We consider the same example as that in SS13 to show how the SPT works in practice; a more detailed investigation of the SPT will be deferred to a later publication. We briefly discuss our findings and conclude in Sect. 5; in particular we will discuss the issue that the SPT only yields an approximate invariance transformation, and its relevance to application in strong lensing systems.

## 2. The principle of the source position transformation

A given mass distribution  $\kappa(\theta)$  defines a mapping from the lens plane  $\theta$  to the source plane,  $\beta = \theta - \alpha(\theta)$ ; throughout this paper, we use standard gravitational lensing notation (see, e.g., Schneider 2006). Provided the mass is sufficiently concentrated, there will be regions in the source plane such that if a source is located there, it has multiple images, that is, several points  $\theta_i$  correspond to the same source position. The source position corresponding to these images is not observable; hence, the constraint imposed on the lens from observing  $n$  such multiple images is

$$\theta_i - \alpha(\theta_i) = \theta_j - \alpha(\theta_j), \quad (2)$$

for all  $1 \leq i < j \leq n$ . Hence, the constraints we obtain from observing a strong lensing system is at best a relation between points corresponding to the same source position – i.e., a mapping  $\theta_i(\theta_1)$ ,  $i \geq 2$ , for all images  $i$  corresponding to the same source position as  $\theta_1$ . Images corresponding to singly-imaged source location carry no strong lensing information about the lens mapping.

We now ask whether there exists another deflection law  $\hat{\alpha}(\theta)$  which yields exactly the same mapping  $\theta_i(\theta_1)$  as the original one. If such a mass distribution exists, then the condition

$$\theta_i - \hat{\alpha}(\theta_i) = \theta_1 - \hat{\alpha}(\theta_1) \quad (3)$$

must be satisfied, for all images  $i$  ( $\geq 2$ ) corresponding to the same source position  $\beta$  as  $\theta_1$ .

The above consideration shows that the new deflection law  $\hat{\alpha}(\theta)$  provides the same mapping  $\theta_i(\theta_1)$  as the original one if (1) all image pairs  $\theta_1, \theta_2$  that belong to the same source position in the original mapping are also multiple images under the new deflection law, and (2) any two points  $\theta_1, \theta_2$  which do not correspond to the same source position in the original mapping are also not matched by the new one. Two deflection laws which satisfy this condition are called ‘equivalent’.

If such an equivalent deflection law to  $\alpha$  indeed exists, then the new deflection  $\hat{\alpha}$  defines a new lens mapping

$$\hat{\beta} = \theta - \hat{\alpha}(\theta) \quad (4)$$

from the lens plane to the source plane. The different images  $\theta_i$  corresponding to the same source position  $\beta$  must also have

the same source position  $\hat{\beta}$  in the new mapping, according to (3). Therefore, the new deflection law  $\hat{\alpha}$  defines a mapping  $\hat{\beta}(\beta)$ , implicitly given by

$$\theta = \hat{\alpha}(\theta) + \hat{\beta} = \alpha(\theta) + \beta, \quad (5)$$

where  $\theta$  is any of the possible multiple images corresponding to the source position  $\beta$ .<sup>1</sup>

We can reverse the argument and consider a mapping  $\hat{\beta}(\beta)$  from the original source coordinates to the new ones; such a mapping in the source plane gives rise to a modified deflection law, as seen by (5),

$$\hat{\alpha}(\theta) = \alpha(\theta) + \beta - \hat{\beta} = \theta - \hat{\beta}(\theta - \alpha(\theta)), \quad (6)$$

where in the last step we inserted the original lens equation  $\beta = \theta - \alpha(\theta)$ . Therefore, any source-position transformation (SPT)  $\hat{\beta}(\beta)$  defines a deflection law  $\hat{\alpha}(\theta)$  such that all images of the same source under the original lens mapping are also multiple images with the new lens equation (4). Thus, the two lens mappings caused by  $\alpha$  and  $\hat{\alpha}$  predict the same multiple images, for all source positions  $\beta$  (or  $\hat{\beta}$ ).

The mass-sheet transformation (MST) is a special case of this more general SPT, obtained by setting  $\hat{\beta} = \lambda\beta$ , which gives rise to the transformed deflection law

$$\hat{\alpha}(\theta) = \theta - \lambda[\theta - \alpha(\theta)] = \lambda\alpha(\theta) + (1 - \lambda)\theta, \quad (7)$$

which we recognize being the deflection of the mass-sheet transformed deflection  $\alpha$ , corresponding to the transformed convergence  $\kappa_\lambda$  in (1).

The Jacobi matrix of the new lens equation (4) reads

$$\hat{\mathcal{A}}(\theta) = \frac{\partial \hat{\beta}}{\partial \theta} = \frac{\partial \hat{\beta}}{\partial \beta} \frac{\partial \beta}{\partial \theta} \equiv \mathcal{B}(\beta(\theta)) \mathcal{A}(\theta), \quad (8)$$

where  $\mathcal{B}$  is the Jacobi matrix of the SPT and  $\mathcal{A}$  the Jacobi matrix of the original lens equation. This implies that

$$\det \hat{\mathcal{A}} = \det \mathcal{B} \det \mathcal{A}. \quad (9)$$

Hence, if the SPT  $\hat{\beta}(\beta)$  is a one-to-one mapping (with no loss of generality, we will require  $\det \mathcal{B} > 0$  for all  $\beta$ ), the critical curves of the modified lens mapping are exactly the same as those of the original lens mapping.<sup>2</sup>

From (8), we infer that the relative magnification matrices between image pairs from the same source  $\hat{\beta}$  remain unchanged,

$$\hat{\mathcal{A}}(\theta_1) \hat{\mathcal{A}}^{-1}(\theta_2) = \mathcal{A}(\theta_1) \mathcal{A}^{-1}(\theta_2), \quad (10)$$

which implies that magnification ratios of image pairs are preserved, as well as their relative image shapes.

To summarize this section: Any bijective SPT  $\hat{\beta}(\beta)$  (with  $\det \mathcal{B} > 0$ ) defines an equivalent deflection law  $\hat{\alpha}(\theta)$  given by (6), i.e., which yields the same strong lensing properties as the original lens mapping. However, this does in general not imply that there is a corresponding mass distribution  $\hat{\kappa}(\theta)$  which yields

<sup>1</sup> Note that for any source position  $\beta$ , at least one image  $\theta$  exists, according to the odd-number theorem (Burke 1981).

<sup>2</sup> If the mapping  $\hat{\beta}(\beta)$  is not one-to-one, then there exist pairs of positions  $\beta^{(1)}$  and  $\beta^{(2)}$  which are mapped onto the same  $\hat{\beta}$ . This implies that all images  $\theta_i^{(1)}$  and  $\theta_i^{(2)}$  which correspond to these two different source positions are images of the same source  $\hat{\beta}$  in the new mapping, modifying the pairing of images. Hence, we will assume  $\det \mathcal{B} > 0$  in the following.

the deflection law  $\hat{\alpha}$ , owing to the fact that in general, the Jacobian matrix  $\hat{\mathcal{A}}$  will be non-symmetric (and thus the deflection  $\hat{\alpha}$  cannot be derived as a gradient of a deflection potential); we will discuss this issue in more detail in Sect. 4. However, for the special case of axi-symmetric lenses, such modified mass distributions do exist, as discussed next.

### 3. The axi-symmetric case

We first consider the case of an axi-symmetric lens, for which the SPT yields an exact invariance transformation between different mass profiles  $\kappa(\theta)$ , which will be explicitly derived in Sect. 3.1. In Sect. 3.2, we provide a few examples for such modified density profiles.

#### 3.1. The SPT-transformed mass profile

Let  $\kappa(\theta)$  be the radial mass profile of a lens, corresponding to the lens mapping  $\beta = \theta - \alpha(\theta)$ , and let

$$\hat{\beta} = [1 + f(\beta)]\beta \quad (11)$$

describe the SPT. To preserve axi-symmetry, the deformation function  $f(\beta)$  must be even,  $f(-\beta) = f(\beta)$ . Furthermore, we require the SPT (11) to be one-to-one, i.e.,  $\hat{\beta}' = 1 + f + \beta f' > 0$ . The modified deflection law is, according to (6),

$$\begin{aligned} \hat{\alpha}(\theta) &= \theta - [1 + f(\theta - \alpha(\theta))] [\theta - \alpha(\theta)] \\ &= \alpha(\theta) - f(\theta - \alpha(\theta))[\theta - \alpha(\theta)]. \end{aligned} \quad (12)$$

We will show next that this deflection can be derived from a mass profile  $\hat{\kappa}(\theta)$ . For this, we first write the deflection as  $\hat{\alpha}(\theta) = \hat{m}(\theta)/\theta$ , where

$$\hat{m}(\theta) = 2 \int_0^\theta d\theta' \theta' \hat{\kappa}(\theta') \quad (13)$$

is the enclosed dimensionless mass within  $\theta$ . This yields

$$\hat{m}(\theta) = \theta \hat{\alpha}(\theta) = \theta \alpha(\theta) - \theta [\theta - \alpha(\theta)] f(\theta - \alpha(\theta)). \quad (14)$$

The mass profile  $\hat{\kappa}$  is obtained from  $\hat{m}$  through  $\hat{\kappa}(\theta) = \hat{m}'(\theta)/(2\theta)$ ; calculating the derivative, we find

$$\begin{aligned} \hat{m}' &= \theta \alpha' + \alpha - (2\theta - \alpha - \theta \alpha') f - \theta(\theta - \alpha)(1 - \alpha') f' \\ &= 2\theta \kappa - 2\theta(1 - \kappa)f - \theta^2 \det \mathcal{A} f', \end{aligned} \quad (15)$$

where we now dropped the arguments of the functions, keeping in mind that  $\alpha$  depends on  $\theta$  and  $f$  on  $\beta = \theta - \alpha$ ; furthermore, we used that  $d\beta/d\theta = 1 - \alpha'$ , and in the last step, we employed the relations  $\alpha' + \alpha/\theta = 2\kappa$  and  $\det \mathcal{A} = (1 - \alpha/\theta)(1 - \alpha')$  which apply for the axi-symmetric case.<sup>3</sup> Hence,

$$\begin{aligned} \hat{\kappa}(\theta) &= \frac{\hat{m}'(\theta)}{2\theta} = \kappa(\theta) - [1 - \kappa(\theta)] f(\theta - \alpha(\theta)) \\ &\quad - \frac{\theta}{2} \det \mathcal{A}(\theta) f'(\theta - \alpha(\theta)). \end{aligned} \quad (16)$$

This equation now yields an explicit expression for the mass profile  $\hat{\kappa}(\theta)$  of the transformed lens mapping, in terms of the original mass distribution  $\kappa(\theta)$  and the source-plane deformation  $f(\theta)$ . The mass profiles  $\kappa(\theta)$  and  $\hat{\kappa}(\theta)$  thus predict exactly the same lensing properties concerning multiple images and flux ratios of compact images, as well as the multiple images of extended source components. Whereas the magnifications are affected, as well as the corresponding shapes of the sources, these properties are unobservable in general in strong lensing systems.

<sup>3</sup> Since  $\alpha = m/\theta$ ,  $\alpha' = m'/\theta - m/\theta^2 = 2\kappa - \alpha/\theta$  and  $\det \mathcal{A} = (\beta/\theta)(d\beta/d\theta)$ .

#### 3.2. Some examples of transformed density profiles

In order to get more insight into this transformation, we first consider some special locations in the lens plane, starting with the tangential critical curve at  $\theta = \theta_E$ , where  $\alpha(\theta_E) = \theta_E$  and thus  $\beta = 0$ . Writing  $\theta \det \mathcal{A} = (\theta - \alpha)(1 - \alpha')$  and using  $\alpha' = 2\kappa - \alpha/\theta$ , we find from differentiating (16) that at the Einstein radius

$$\begin{aligned} \hat{\kappa}(\theta_E) &= \kappa(\theta_E) [1 + f(0)] - f(0), \\ \hat{\kappa}'(\theta_E) &= \kappa'(\theta_E) [1 + f(0)], \\ \hat{\kappa}''(\theta_E) &= \kappa''(\theta_E) [1 + f(0)] - 12 [1 - \kappa(\theta_E)]^3 f''(0), \\ \hat{\kappa}'''(\theta_E) &= \kappa'''(\theta_E) [1 + f(0)] + 12 [1 - \kappa(\theta_E)]^2 f''(0) \\ &\quad \times \left\{ 6\kappa'(\theta_E) + \frac{5[1 - \kappa(\theta_E)]}{\theta_E} \right\}, \end{aligned} \quad (17)$$

where we also used the fact that  $f$  is an even function, i.e., its odd derivatives vanish at the origin. These relations show how the mass profile can be modified near the Einstein radius. We can choose  $\hat{\kappa}(\theta_E)$  freely with an appropriate choice of  $f(0)$ , but that fixes the slope of  $\hat{\kappa}$  at  $\theta_E$ , which is as well determined by  $f(0)$ . This connection between  $\hat{\kappa}(\theta_E)$  and  $\hat{\kappa}'(\theta_E)$  is the same as for the MST. The new feature of the SPT shows up for the curvature of the mass profile at  $\theta_E$  for which we again can make a choice, but then the third derivative is fixed – and so on. The fact that  $f$  is an even function implies that we can make a choice for all even derivatives, but the odd derivatives are then tied to the former.

Choosing the derivatives of the mass profile at the Einstein radius by selecting  $f(0)$  and its even derivatives then fixes the expansion of the central surface mass density through

$$\begin{aligned} \hat{\kappa}(0) &= \kappa(0) [1 + f(0)] - f(0), \\ \hat{\kappa}'(0) &= \kappa'(0) [1 + f(0)], \\ \hat{\kappa}''(0) &= \kappa''(0) [1 + f(0)] - 2 [1 - \kappa(0)]^3 f''(0), \\ \hat{\kappa}'''(0) &= \kappa'''(0) [1 + f(0)] + 15 [1 - \kappa(0)]^2 \kappa'(0) f''(0). \end{aligned} \quad (18)$$

In particular, if the mass profile is smooth at the origin, so that all odd derivatives of  $\kappa$  vanish there, the same property will be shared by the transformed mass distribution.

If there is a radial critical curve, which is the case if  $\kappa(\theta)$  is a regular function, then at  $\theta_c$ ,  $1 - \alpha'(\theta_c) = 0$ . The corresponding caustic in the source plane has radius  $\beta_c = \alpha(\theta_c) - \theta_c$ , with  $\alpha(\theta_c) = \theta_c [2\kappa(\theta_c) - 1]$ . At this location, we then obtain

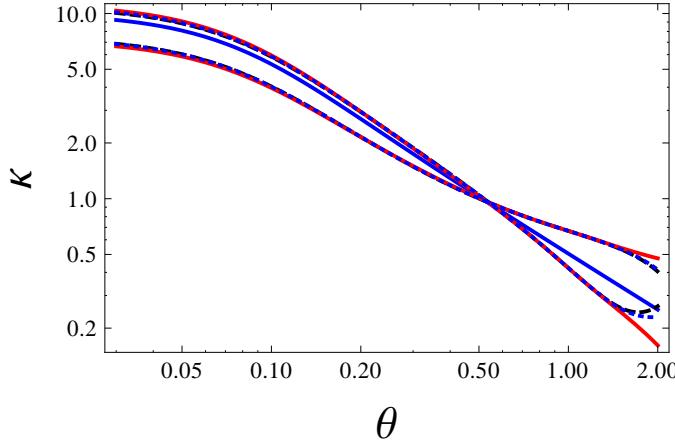
$$\begin{aligned} \hat{\kappa}(\theta_c) &= \kappa(\theta_c) [1 + f(\beta_c)] - f(\beta_c) \\ \hat{\kappa}'(\theta_c) &= \kappa'(\theta_c) [1 + f(\beta_c)] + 2 [1 - \kappa(\theta_c)] f'(\beta_c) \\ &\quad \times [\kappa(\theta_c) + \theta_c \kappa'(\theta_c) - 1] \\ \hat{\kappa}''(\theta_c) &= \kappa''(\theta_c) [1 + f(\beta_c)] + 2 [1 - \kappa(\theta_c)] f'(\beta_c) \\ &\quad \times [2 - 2\kappa(\theta_c) + \theta_c^2 \kappa''(\theta_c)] / \theta_c. \end{aligned} \quad (19)$$

#### 3.3. Examples

Our first set of examples for the action of the SPT is constructed such that near the tangential critical curve, the transformed mass distribution is approximately a power law,  $\hat{\kappa}(\theta) \approx \hat{\kappa}(\theta_E)(\theta/\theta_E)^{-\nu}$ , for  $\theta$  close to  $\theta_E$ . For such a power-law distribution, one finds that

$$\nu = -\frac{\hat{\kappa}'(\theta_E) \theta_E}{\hat{\kappa}(\theta_E)}, \quad \theta_E^2 \hat{\kappa}''(\theta_E) = \left( \frac{\hat{\kappa}'(\theta_E) \theta_E}{\hat{\kappa}(\theta_E)} - 1 \right) \hat{\kappa}'(\theta_E) \theta_E. \quad (20)$$

If we choose a singular isothermal sphere (SIS) as the original mass model, with  $\kappa = \theta_E/(2\theta)$ ,  $\alpha = \theta_E$ , and use the first three



**Fig. 1.** Axi-symmetric example of an SPT between a non-singular isothermal sphere (blue solid curve), with core  $\theta_c = 0.1\theta_E$ , and other mass profiles. The other curves are transformed mass profiles, using the SPT, with different deformation function  $f(\beta)$ , all satisfying the power-law condition (20), and with slope  $\nu = 0.5$  (flatter curves) and  $\nu = 1.4$  (steeper curves): The dashed black curves correspond to the polynomial  $f(\beta) = f_0 + f_2 \beta^2 / 2$ , the red curves to  $f(\beta) = f_0 + \beta_0^2 f_2 \beta^2 / [2(\beta^2 + \beta_0^2)]$ , with  $\beta_0 = 0.8\theta_E$ , and the blue dotted curves to  $f(\beta) = 2f_0 / \cosh(\beta/\beta_0) - f_0$ , with  $\beta_0 = \theta_E \sqrt{3(1-f_0)/(1+f_0)}$

relations of (17), the second equation of (20) then yields a condition for the second derivative of  $f$  at the origin,

$$\theta_E^2 f_2 = -\frac{2f_0(1+f_0)}{3(1-f_0)}, \quad (21)$$

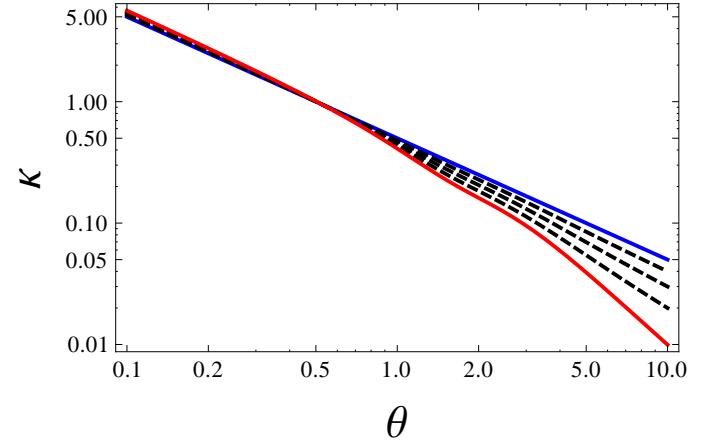
where  $f_0 \equiv f(0)$ ,  $f_2 \equiv f''(0)$ , and the local slope is  $\nu = (1+f_0)/(1-f_0)$ . In Fig. 1 we have plotted the original mass profile, where we have taken a non-singular isothermal sphere with core radius  $\theta_c = 0.1\theta_E$ ; the introduction of a small core does affect the foregoing relations only very little, as these are obtained by considering  $\kappa(\theta)$  at the Einstein radius. For two values of the slope  $\nu$  near the Einstein radius,  $\nu = 0.5$  and  $\nu = 1.4$ , we have plotted three different transformed mass profiles, where the corresponding functions  $f(\beta)$  are described in the figure caption. For all these cases, the local behavior near the Einstein radius is indeed well approximated by a power law. Some of the profiles become unphysical near  $\theta \sim 2$ , i.e.,  $\beta \sim 1$ , which is due to the fact that the first derivative of  $f$  which enters (16) becomes too large there for the corresponding deformation function. Nevertheless, these simple example already show the range of freedom the SPT offers to generate axi-symmetric mass profiles with identical strong lensing behavior. We also point out that we plotted the mass profiles only up to  $\theta = 2\theta_E$ , i.e., in the angular range where multiple images occur. For larger  $\theta$ , the mass profile can be chosen arbitrarily, without constraints.

We next consider the case of large  $\beta > \beta_c$ , for which no multiple images occur, so that the lens mapping becomes one-to-one there. Hence, for sufficiently large  $\theta$ , the lens equation defines a mapping  $\theta(\beta)$ . We then can rewrite (16) in the form

$$f'(\beta) + \frac{2[1-\kappa(\theta)]}{\theta \det \mathcal{A}(\theta)} f(\beta) = \frac{2[\kappa(\theta) - \hat{\kappa}(\theta)]}{\theta \det \mathcal{A}(\theta)}, \quad (22)$$

or by using the mapping  $\theta(\beta)$ ,

$$f'(\beta) + g(\beta) f(\beta) = h(\beta), \quad (23)$$



**Fig. 2.** Second example of an SPT, in the axi-symmetric case, between an SIS with mass distribution  $\kappa(\theta)$  and other mass distributions  $\hat{\kappa}(\theta)$ , characterized by a deformation function

$$f(\beta) = a \left[ \frac{\theta_E}{\sqrt{\beta^2 + \beta_0^2}} - \frac{2\theta_1^2 \ln(\sqrt{\beta^2 + \beta_0^2} + \theta_E)}{\sqrt{\beta^2 + \beta_0^2} (\sqrt{\beta^2 + \beta_0^2} + \theta_E)} \right],$$

where we set  $\theta_1 = \theta_E$ ,  $\beta_0 = 3\theta_E/2$ , and different values of  $a = 0, 1/4, 1/2, 3/4, 1$ . For  $a = 0$  (blue curve), the original SIS is obtained, for  $a = 1$  (red curve),  $\hat{\kappa}$  behaves like  $(\theta_1/\theta_E)^2$  for large radii

where  $g = 2(1-\kappa)/(\theta \det \mathcal{A})$  and  $h = 2(\kappa - \hat{\kappa})/(\theta \det \mathcal{A})$  are functions of  $\beta$ . For a given  $\kappa$  and a target  $\hat{\kappa}$ , the function  $f$  can be determined by solving this differential equation.

We will give a simple example for this procedure. Assume  $\kappa$  to describe an SIS. Furthermore, let our target density profile be  $\hat{\kappa}(\theta) = (\theta_1/\theta)^2$ . This yields

$$g = \frac{2(1-\theta_E/2\theta)}{(\theta-\theta_E)} = \frac{2\beta+\theta_E}{\beta(\beta+\theta_E)}, \quad (24)$$

$$h = \frac{\theta_E/\theta - 2(\theta_1/\theta)^2}{\theta-\theta_E} = \frac{\theta_E(\beta+\theta_E) - 2\theta_1^2}{(\theta_E+\beta)^2\beta}, \quad (25)$$

for which the deformation function becomes

$$f(\beta) = \frac{\theta_E}{\beta} - \frac{2\theta_1^2 \ln(\beta+\theta_E)}{\beta(\beta+\theta_E)}. \quad (26)$$

This functional form of  $f$  is, however, only a valid description for large arguments; in particular,  $f$  is not an even function of  $\beta$ . We can now modify  $f$  such that it retains the form (26) for large  $\beta$ , but becomes a function of  $\beta^2$  only. A simple way to achieve this is to replace  $\beta$  in (26) by  $\sqrt{\beta^2 + \beta_0^2}$ , for a suitably chosen  $\beta_0$ . In Fig. 2, we illustrate this example, showing that by this choice of  $f$  one finds a mass distribution  $\hat{\kappa}$  which has a significantly different form from  $\kappa$ , nevertheless giving rise to exactly the same strong lensing properties for all source positions.

Note that the slope near the Einstein radius of the profiles shown in Figs. 1 and 2 are quite different, hence these mass models will give rise to very different predictions for the product of time delay and Hubble constant,  $\tau = H_0 \Delta t$ , which reinforces the point made in SS13.

## 4. The general case

We now drop the assumption of axi-symmetry; in this case, the matrix  $\hat{\mathcal{A}}$  in (8) will not be symmetric in general. This is due to the fact that  $\hat{\mathcal{A}}$  is symmetric only if the directions of the eigenvectors of  $\mathcal{A}$  and  $\mathcal{B}$  are the same. For source points which correspond to multiple images, even if we could arrange  $\mathcal{B}$  to have the same eigendirections as  $\mathcal{A}(\theta_1)$ , it will not have the same eigendirections as  $\mathcal{A}(\theta_i)$  for the other images  $\theta_i$ . The only possibility to keep  $\hat{\mathcal{A}}$  symmetric for a general lens is to have  $\mathcal{B}$  proportional to the unit matrix, which is the case for  $\hat{\beta} = \lambda\beta$ , which recovers the MST mentioned before.

The asymmetric nature of  $\hat{\mathcal{A}}$ , discussed in more detail in Sect. 4.1, implies that the deflection law  $\hat{\alpha}$ , which yields exactly the same strong lensing properties as  $\alpha$ , cannot be derived as the gradient of a deflection potential. Thus, in general there does not exist a surface mass density  $\hat{\kappa}$  which generates the same strong lensing properties for all source positions as  $\kappa$ . However, it may be possible that one can find a surface mass density which generates almost the same mapping for all sources. In Sect. 4.2 we obtain a crude estimate for the amplitude of the asymmetry of  $\hat{\mathcal{A}}$ , for the special case of a quadrupole lens. An illustrative example is discussed in more detail in Sect. 4.3, where we show explicitly that the impact of the asymmetry of  $\hat{\mathcal{A}}$  can be very small in realistic cases.

### 4.1. The transformed mass profile

The Jacobi matrix of the original lens mapping has the form

$$\mathcal{A} = [1 - \kappa(\theta)] \mathcal{I} - \gamma(\theta) \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}, \quad (27)$$

where  $\mathcal{I}$  is the unit matrix. We will write the matrix  $\mathcal{B}$  in a similar form,

$$\mathcal{B} = B_1 \mathcal{I} + B_2 \begin{pmatrix} \cos 2\eta & \sin 2\eta \\ \sin 2\eta & -\cos 2\eta \end{pmatrix}. \quad (28)$$

The angle  $\vartheta$  is the phase of the shear of the original lens mapping; similarly,  $\eta$  is the phase of the shear of the mapping  $\hat{\beta}(\beta)$  at the position  $\beta(\theta)$ . We then obtain

$$\begin{aligned} \hat{\mathcal{A}} &= B_1(1 - \kappa)\mathcal{I} + B_2(1 - \kappa) \begin{pmatrix} \cos 2\eta & \sin 2\eta \\ \sin 2\eta & -\cos 2\eta \end{pmatrix} \\ &- \gamma B_1 \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \\ &- \gamma B_2 \begin{pmatrix} \cos 2(\eta - \theta) & -\sin 2(\eta - \theta) \\ \sin 2(\eta - \theta) & \cos 2(\eta - \theta) \end{pmatrix}. \end{aligned} \quad (29)$$

The asymmetry of  $\hat{\mathcal{A}}$  is due to the fact that these two phases are different in general. As we will show below, in the axi-symmetric case, these two phases are the respective polar angles of  $\theta$  and  $\beta$ ; the fact that  $\beta$  and  $\theta$  are collinear in the axi-symmetric case guarantees that these two angles are the same (or differ by  $\pi$ ). This then implies the symmetry of  $\hat{\mathcal{A}}$ .

We now define the modified surface mass density  $\hat{\kappa}(\theta)$  through the trace of  $\hat{\mathcal{A}}$ ,  $\text{tr}\hat{\mathcal{A}} = 2(1 - \hat{\kappa})$ , i.e., in the same way as if the Jacobian  $\hat{\mathcal{A}}$  was derived from a deflection potential; this yields

$$\hat{\kappa} = 1 + B_1(\kappa - 1) + B_2\gamma \cos 2(\eta - \theta), \quad (30)$$

where it should be kept in mind that  $\kappa$ ,  $\gamma$  and  $\vartheta$  depend on  $\theta$ , and the  $B_i$  and  $\eta$  depend on  $\beta(\theta)$ . In addition, we characterize the asymmetry of  $\mathcal{A}$  by

$$\hat{\kappa}_I = B_2\gamma \sin 2(\eta - \theta). \quad (31)$$

We can easily show that Eq. (30) for  $\hat{\kappa}$  reduces to our earlier result for an axi-symmetric mass profile. From (11) in the form  $\hat{\beta} = [1 + f(|\beta|)]\beta$ , we obtain

$$B_1 = 1 + f + \frac{\beta f'}{2}; \quad B_2 = \frac{\beta f'}{2}, \quad (32)$$

and the phase is  $\eta = \phi$ , where  $\phi$  is the polar angle of  $\beta$ . The shear of the lens is  $\gamma = \kappa - \bar{\kappa}$ , and its phase agrees with the polar angle of  $\theta$ ,  $\vartheta = \varphi$ . Here,  $\bar{\kappa}(\theta) = m(\theta)/\theta^2$  is the mean convergence within radius  $\theta$ . Since either  $\phi = \varphi$  or  $\phi = \varphi + \pi$ , the asymmetric term in (29) vanishes, and the modified convergence becomes

$$\begin{aligned} \hat{\kappa} &= 1 + (1 + f + \beta f'/2)(\kappa - 1) + (\kappa - \bar{\kappa})\beta f'/2 \\ &= \kappa + (\kappa - 1)f + (2\kappa - 1 - \bar{\kappa})\beta f'/2, \end{aligned} \quad (33)$$

which is seen to agree with (16), since  $\theta \det \mathcal{A} = (\theta - \alpha)(1 + \bar{\kappa} - 2\kappa)$ .

One can expect that the convergence  $\hat{\kappa}$  defined in (30) yields a deflection law which very closely resembles that of  $\hat{\alpha}$  if the asymmetry of  $\hat{\mathcal{A}}$ , as characterized by  $\hat{\kappa}_I$ , is small compared to  $\hat{\kappa}$ . This will be the case if the source plane deformation, which determines the amplitude of  $B_2$ , is sufficiently small, and/or if the misalignment between the shear  $\gamma$  of the lens mapping and that of the mapping  $\hat{\beta}(\beta)$  is small. The latter aspect depends on the kind of lens mapping one is dealing with.

### 4.2. Example: The quadrupole lens

As an illustrative case, we consider a quadrupole lens, i.e., an axi-symmetric matter distribution characterized by the convergence  $\kappa(|\theta|)$  plus some external shear, so that the lens equation becomes

$$\beta = [1 - \bar{\kappa}(|\theta|)]\theta - \begin{pmatrix} \gamma_p & 0 \\ 0 & -\gamma_p \end{pmatrix}\theta, \quad (34)$$

where  $\gamma_p$  is a constant shear caused by some external ‘perturbing’ large-scale mass distribution.<sup>4</sup> Denoting the shear of the main lens by  $\gamma_m(\theta) = \kappa(|\theta|) - \bar{\kappa}(|\theta|)$ , the Jacobian reads

$$\mathcal{A} = (1 - \kappa)\mathcal{I} - \begin{pmatrix} \gamma_p + \gamma_m \cos 2\varphi & \gamma_m \sin 2\varphi \\ \gamma_m \sin 2\varphi & -\gamma_p - \gamma_m \cos 2\varphi \end{pmatrix}, \quad (35)$$

which can be compared with the form (27) of  $\mathcal{A}$ ; this yields for the shear components of  $\mathcal{A}$

$$\begin{aligned} \cos 2\theta &= \frac{\gamma_p + \gamma_m \cos 2\varphi}{\gamma}; \quad \sin 2\theta = \frac{\gamma_m \sin 2\varphi}{\gamma}, \\ \gamma^2 &= \gamma_p^2 + \gamma_m^2 + 2\gamma_p\gamma_m \cos 2\varphi. \end{aligned} \quad (36)$$

For the SPT, we again consider a radial stretching,  $\hat{\beta} = [1 + f(|\beta|)]\beta$ , for which the coefficients of  $\mathcal{B}$  are given in (32),

<sup>4</sup> In general, such a large-scale perturber will also induce some external convergence. However, for simplicity we will neglect this effect here, since it can be easily scaled out by a mass-sheet transformation (see, e.g. Schneider 2006).

and for which the phase  $\eta$  equals the polar angle  $\phi$ . We calculate  $\beta$  from the lens equation (34), to find

$$\begin{aligned}\cos 2\phi &= \frac{\beta_1^2 - \beta_2^2}{\beta^2} = \frac{[(1 - \bar{\kappa})^2 + \gamma_p^2] \cos 2\varphi - 2\gamma_p(1 - \bar{\kappa})}{(\beta/\theta)^2}, \\ \sin 2\phi &= \frac{2\beta_1\beta_2}{\beta^2} = \frac{[(1 - \bar{\kappa})^2 - \gamma_p^2] \sin 2\varphi}{(\beta/\theta)^2}, \\ \beta^2 &= \theta^2 [(1 - \bar{\kappa})^2 + \gamma_p^2 - 2(1 - \bar{\kappa})\gamma_p \cos 2\varphi].\end{aligned}\quad (37)$$

From these relations, we can calculate the products  $\gamma \cos 2(\eta - \vartheta) = \gamma \cos 2(\phi - \vartheta)$  and  $\gamma \sin 2(\phi - \vartheta)$ , which enter the quantities (30) and (31),

$$\gamma \cos 2(\phi - \vartheta) = \left(\frac{\theta}{\beta}\right)^2 \left\{ \gamma_m(1 - \bar{\kappa})^2 - 2\gamma_p(1 - \bar{\kappa}) \right. \quad (38)$$

$$\left. + \gamma_p [(1 - \bar{\kappa})^2 - 2\gamma_m(1 - \bar{\kappa}) + \gamma_p^2] \cos 2\varphi + \gamma_m \gamma_p^2 \cos 4\varphi \right\},$$

$$\begin{aligned}\gamma \sin 2(\phi - \vartheta) &= -\left(\frac{\theta}{\beta}\right)^2 \gamma_p \left\{ \gamma_m \gamma_p \sin(4\varphi) \right. \\ &\left. + [\gamma_p^2 - (1 - \bar{\kappa})(2\gamma_m + 1 - \bar{\kappa})] \sin 2\varphi \right\}.\end{aligned}\quad (39)$$

Combining (30), (32) and (38), we then obtain for the transformed surface mass density

$$\hat{\kappa} = (1 + f)\kappa - f + \frac{\beta f'}{2} [\kappa - 1 + \gamma \cos 2(\phi - \vartheta)]. \quad (40)$$

If we now approximate the function  $f(\beta)$  near the origin by  $f(\beta) = f_0 + f_2 \beta^2/2$ , where we accounted for the fact that  $f$  is an even function, we see that  $\beta f'/2 = f_2 \beta^2/2$ , so that

$$\begin{aligned}\hat{\kappa} &= (1 + f_0)\kappa - f_0 + \frac{f_2 \beta^2}{2} [2(\kappa - 1) + \gamma \cos 2(\phi - \vartheta)] \\ &= (1 + f_0)\kappa - f_0 + \frac{f_2 \theta^2}{2} \left\{ \gamma_m [2\gamma_p^2 + 3(1 - \bar{\kappa})^2] \right. \\ &\quad \left. - 2(1 - \bar{\kappa}) [(1 - \bar{\kappa})^2 + 2\gamma_p^2] \right\} \\ &\quad + [5\gamma_p(1 - \bar{\kappa})^2 - 6\gamma_p \gamma_m(1 - \bar{\kappa}) + \gamma_p^3] \cos 2\varphi + \gamma_m \gamma_p^2 \cos 4\varphi\end{aligned}\quad (41)$$

The terms independent of  $f_2$  present just the MST (1), with  $\lambda = 1 + f_0$ . The non-linear part of the SPT, here parametrized by  $f_2$ , both adds a monopole contribution to  $\hat{\kappa}$ , as well as contributions which depend on  $\varphi$ . The monopole contribution itself has two parts, one independent of  $\gamma_p$  – which can be shown to agree with (16) – and one  $\propto \gamma_p^2$ . This latter contribution is expected to be small, for reasonably small values of the external shear. The angle-dependent contributions to  $\hat{\kappa}$  vanish for  $\gamma_p = 0$ . At the tangential critical curve of the axi-symmetric lens, where  $\bar{\kappa} = 1$ , these angle-dependent terms are at least of order  $\gamma_p^2$ , again expected to be small. For locations away from the Einstein circle, the leading-order terms are  $\propto \gamma_p$  and thus typically a factor of  $\gamma_p$  smaller than the  $f_2$ -induced contributions to the monopole term.

With the same form of  $f(\beta)$ , the quantity (31) describing the asymmetry of  $\hat{\mathcal{A}}$  becomes

$$\begin{aligned}\hat{\kappa}_I &\approx -\frac{\gamma_p}{2} \theta^2 f_2 \\ &\times [\gamma_p^2 - (1 - \bar{\kappa})(2\gamma_m + 1 - \bar{\kappa}) + 2\gamma_m \gamma_p \cos(2\varphi)] \sin 2\varphi.\end{aligned}\quad (42)$$

From this result we see that the asymmetry vanishes if  $\gamma_p = 0$ . Furthermore, it vanishes on the axes,  $\varphi = n\pi/2$ ,  $n = 0, 1, 2, 3$ ,

since points  $\theta$  on the symmetry axes are mapped onto the corresponding axis in the source plane, so that the shear matrices are aligned in this case.

As expected,  $\hat{\kappa}_I$  is proportional to the curvature  $f_2$  of the distortion function at the origin, and further has as prefactor the external shear  $\gamma_p$ . Both of these are typically small; characteristic values for the external shear in strong lens systems are  $\gamma_p \lesssim 0.1$ . We have seen in Sect. 2 that distortion functions with too large curvature lead to unphysical mass distributions, so that  $f_2 \theta^2$  will be considerably smaller than unity in the multiple image region. Hence, the prefactor will be smaller than about  $10^{-2}$ . In fact, in the example considered in the following subsection, this prefactor is of order  $2 \times 10^{-4}$ . Furthermore, we also point out that the dominant term in the bracket (the middle one) will be small near the critical curve, since the latter will be close to the radius  $\theta_E$  of the Einstein circle of the original axi-symmetric lens model, where  $\bar{\kappa} = 1$ . Finally we note that the  $\varphi$ -average of  $\hat{\kappa}_I$  vanishes, due to the sine-factor, so that  $\hat{\kappa}_I$  will oscillate around zero as one considers circles of constant  $\theta$ . Overall, we thus expect  $\hat{\kappa}_I$  to be small, so that the difference between the deflection angle  $\hat{\alpha}(\theta)$  and the one derived from the modified surface mass density (30) will be small.

#### 4.3. An illustrative example

Whereas a more detailed investigation of this question will be deferred to a later publication, we will illustrate the SPT with a simple example. In SS13, we found that a quadrupole lens with a mass model  $\kappa(\theta)$  consisting of a Hernquist profile (representing the baryonic mass of a lens galaxy) and a generalized NFW-profile (to approximate the central part of the dark matter halo) is almost degenerate with a model  $\kappa_{PL}(\theta)$  corresponding to a power law with an inner core. Since the time-delay ratios do not simply scale by a constant factor, the transformation between these two mass models is not a MST. Instead, as we will show next, this transformation is an example of the SPT discussed here.

The two mass models we study hereafter are the same as in Sect. 4.2 of SS13, namely a fiducial model composed of a spherically symmetric Hernquist+generalised NFW with an external shear  $(\gamma, \theta_\gamma) = (0.1, 90^\circ)$ , and a target power-law model with a core radius  $\theta_c = 0''.1$ , a logarithmic (three-dimensional) slope  $\gamma' = 2.24$  and an external shear  $(\gamma, \theta_\gamma) = (0.09, 90^\circ)$ . Hereafter, we note the lensing quantities associated to the power-law model with a *hat*, e.g.  $\kappa_{PL} = \hat{\kappa}$ .

Using our fiducial mass distribution and the public lens modeling code `lensmodel` (v1.99; Keeton 2001), we create mock images of a uniform grid of  $19 \times 19$  sources covering the first quadrant of the source plane from  $\beta = (\beta_x, \beta_y) = (0.025, 0.025)''$  to  $(\beta_x, \beta_y) = (0.975, 0.975)''$ . The sources and the mock lensed images are shown in Fig. 3. As we have shown in SS13, those lensed images are also reproduced to a good accuracy with the power-law model, provided that the sources are now located at positions  $\hat{\beta}$ . The right panel of Fig. 3 shows the offset (in arcsec) between the mock lensed images  $\theta$  and the lensed images  $\hat{\theta}$  corresponding to the source positions  $\hat{\beta}$ . We see that the two sets of images are almost identical. Assuming an astrometric accuracy of  $0''.004$  in the image plane, we have calculated an astrometric  $\chi^2$  for each source, as shown in the left panel<sup>5</sup> of Fig. 3. The two models are only *approximately degenerate* since  $\chi^2$  deviates significantly from 0 in the vicinity of the outer caustic. The figure

<sup>5</sup> Note that to ease legibility, we have shown only those sources with  $\chi^2 < 1$ .

reveals that the degeneracy between the two models is valid for  $|\beta| \lesssim 0.^{\circ}9$ .

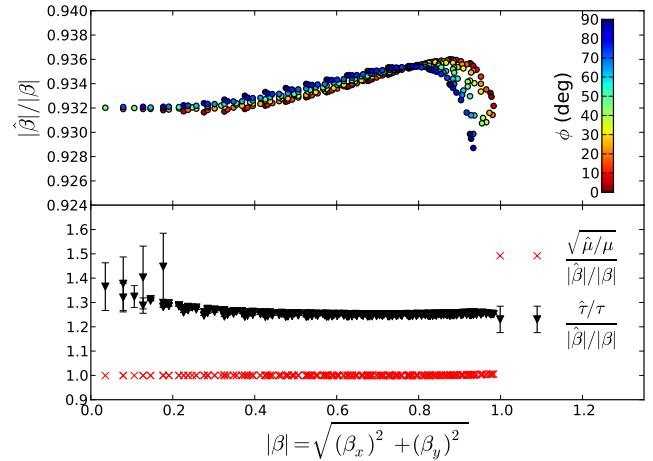
To characterize the nature of the degeneracy we first consider the relation between  $\beta$  and  $\hat{\beta}$ . In case of a MST,  $\hat{\beta} = \lambda\beta$ , and  $\lambda$  is constant. We show in the upper panel of Fig. 4 the variation of  $|\hat{\beta}|/|\beta|$  with  $|\beta|$  and with the polar angle of the source,  $\phi$ . We see that  $|\hat{\beta}|/|\beta| = 0.932$  at the origin, and it increases monotonically to  $|\hat{\beta}|/|\beta| = 0.936$  when  $|\beta| \sim 0.^{\circ}9$ . This implies that for this specific example, the product  $\theta^2 f_2$  occurring in (42) is about  $4 \times 10^{-3}$ , yielding an asymmetry of the Jacobian  $\hat{\mathcal{A}}$  of less than  $\sim 2 \times 10^{-4}$ .

The sharp decrease of the function  $(1+f)$  for  $|\beta| \gtrsim 0.^{\circ}9$  likely reflects the approximate character of the degeneracy between  $\kappa$  and  $\hat{\kappa}$ .<sup>6</sup> The change of  $|\hat{\beta}|/|\beta|$  with  $\beta$  demonstrates that the transformation is not a MST. On the other hand, the small dependence of  $|\hat{\beta}|/|\beta|$  on  $\phi$  shows that the mapping  $\beta \rightarrow \hat{\beta}$  is almost isotropic, but not exactly so.

We now compare the magnification of the lensed images and the corresponding total magnification  $|\mu|$  of the source. The spatial variation of the magnification ratio in the source plane is displayed in the third quadrant of Fig. 3. We see that the factor by which the magnifications are transformed is the square of the factor transforming the positions, as shown on the bottom panel of Fig. 4 which shows  $\sqrt{|\mu|}/|\hat{\mu}|$  normalised by  $|\beta|/|\hat{\beta}|$ . In other words, if  $\hat{\beta} = [1 + f(\beta)]\beta$ , the magnification has to be transformed as  $\hat{\mu} = [1 + f(\beta)]^2 \mu$ . This behaviour is similar to what is observed for the MST, except that there is a dependence on  $\beta$ .

Finally, we display also on Fig. 4 the change of the time delay  $\tau/\hat{\tau}$  normalised by  $|\beta|/|\hat{\beta}|$ . While the delays should scale like the source positions in case of a MST, they do not so for the SPT. The ratio  $(\tau/\hat{\tau})/(|\beta|/|\hat{\beta}|)$  is almost constant but differs significantly from 1. In addition, we see that for sources interior to the inner (astroid) caustic, the time delay ratios between lensed images (of the same source) are not conserved. The error bars in Fig. 4 show, for each source, the spread of  $\tau/\hat{\tau}$  among the lensed images. This spread may reach 10% for some peculiar source positions, but generally arises from the deviation of only one of the lensed images. Owing to accuracies on the time delays of a few percents currently achieved for quads (e.g. Fassnacht et al. 2002; Courbin et al. 2011; Tewes et al. 2012), time delay measurements in quadruply lensed systems could in the most favourable cases break the SPT independently of  $H_0$ .

In summary, we have shown that the degeneracy between a composite and a power-law model with finite core identified in SS13 is an example of an approximate SPT in the case of a quadrupole lens. The non-uniform rescaling of the source morphology implies also a rescaling of its surface brightness. Although this means that in principle an examination of the source should allow one to exclude some inadequate degenerate models, it is likely that, as for the example shown here, the source corresponding to the two different models will both have similarly plausible morphologies and surface brightnesses. Again, time delay ratio measurement may play a critical role to break the degeneracy between SPT-generated models. First, because the time delay ratios are not perfectly conserved when 3 time-



**Fig. 4.** Properties of an approximate SPT  $\beta \rightarrow \hat{\beta}$  (11) for a quadrupole lens. This SPT transforms a surface mass density  $\kappa$  for our fiducial model into a surface mass density  $\hat{\kappa}$  of a power law with a finite core. *Top:* Proxy to the factor  $1 + f(\beta)$  which scales the positions, given by the ratio of source position modulus obtained for the power law and the fiducial models, i.e.  $|\hat{\beta}|/|\beta|$ . The color coding of this quantity as a function of the polar angle of the source  $\phi$  (in degrees) shows the azimuthal change of  $f(\beta)$ , or in other words, the slight anisotropy of the SPT. *Bottom:* The curves show how the SPT modifies the magnification  $\mu \rightarrow \hat{\mu}$  and the time delay  $\tau \rightarrow \hat{\tau}$ . The red crosses show  $\sqrt{\mu/\hat{\mu}}/(|\beta|/|\hat{\beta}|)$  and the black triangles show  $(\hat{\tau}/\tau)/(|\beta|/|\hat{\beta}|)$  for each source position. The error bars give the standard deviation of the quantity as obtained for the different images of the same source. The standard deviation is negligible for the magnification but can reach 10% for the time delay

delays of the same source are observed. Second, because the time delay is not invariant under an SPT, and so, if the value of the Hubble constant is assumed to be known from other observations, the degeneracy may be at least partially broken. Finally, we also note that the amplitude of the shear also gets modified by the SPT.

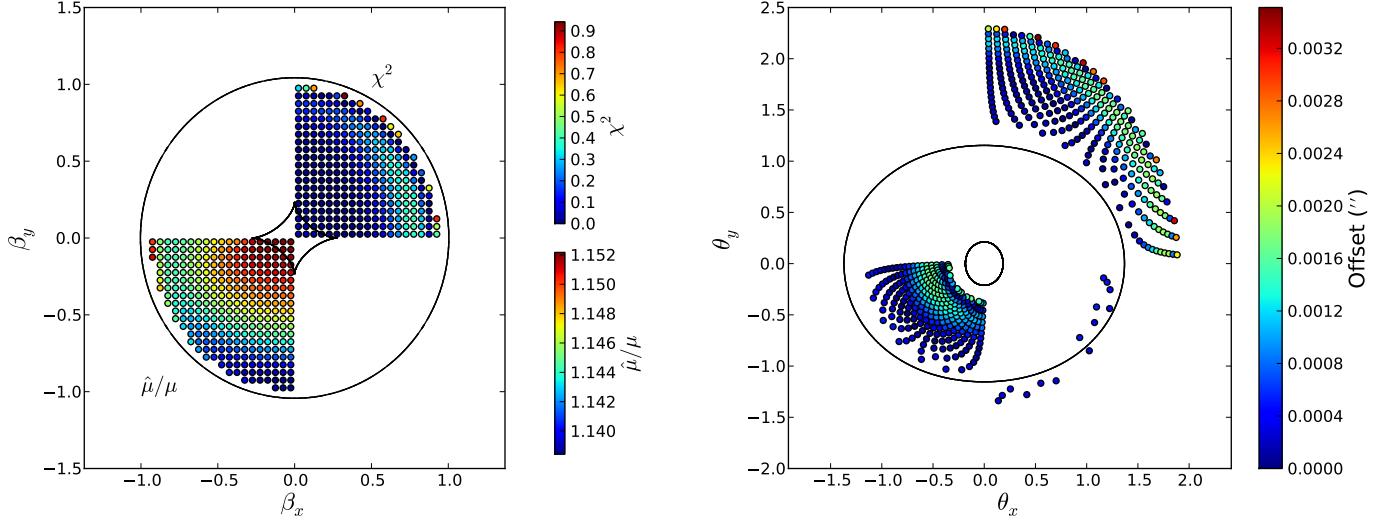
## 5. Discussion and conclusions

We have shown that there exists a transformation of the deflection angle in gravitational lens systems which (1) leaves all strong lensing observables invariant, except the time delay, and (2) is much more general than the well-known mass-sheet transformation. For the axi-symmetric case, this new source-position transformation is exact, in the sense that there exists a surface mass density  $\hat{\kappa}(\theta)$  which corresponds to the transformed deflection law  $\hat{\alpha}(\theta)$ . In the general case, the transformed deflection  $\hat{\alpha}$  has a curl component, which causes the resulting Jacobi matrix to attain an asymmetric contribution.

We have then defined the convergence  $\hat{\kappa}$  of the transformed lens in terms of the trace of the transformed Jacobi matrix  $\hat{\mathcal{A}}$ . The deflection angle obtained from this convergence is expected to closely approximate the transformed deflection angle  $\hat{\alpha}$ , provided the asymmetry of  $\hat{\mathcal{A}}$  is sufficiently small. Hence, in this case the mass distribution  $\hat{\kappa}$  will yield almost the same strong lensing predictions as the original mass profile  $\kappa$ .

Such an approximate agreement between the original and the transformed lensing properties is sufficient for the typical strong lens systems. These are usually modeled by ‘simple’ mass profiles with a small number of parameters which fit the observed

<sup>6</sup> We point out here that there was no fine-tuning involved in constructing this special example, i.e., this example from SS13 turned out to be almost degenerate with a cored power-law mass profile. The fact that it almost corresponds to an SPT is accidental. Some fine-tuning of the mass profiles, equivalent to a fine-tuning of the function  $\hat{\beta}(\beta)$ , would enable an even better agreement between the lensing properties of the two models.



**Fig. 3.** *Right:* Set of mock lensed images produced by our fiducial lens model (using a grid of source positions shown in the first quadrant of the left panel). The color coding shows the offset in arcseconds, between the original image positions and those found with the power-law model. *Left:* First quadrant (upper right): Grid of source positions lensed with our fiducial model to produce the images shown in the right panel. The color coding indicates the  $\chi^2$  associated to each source, assuming an astrometric uncertainty on the image position of  $\sigma_{\theta_{x,y}} = 0.004''$ . In the third quadrant (lower left), the sources do not have corresponding lensed images on the right panel, and are colored as a function of their total magnification as derived with the power-law and the fiducial model (i.e.  $\mu_{\text{PL}}/\mu_{\text{fid}} = \hat{\mu}/\mu$ )

image configuration remarkably well. Exact fitting is not required, though, for at least two reasons: first, the observed image positions (and the observed brightness profile for extended images) have an observational uncertainty, which for the best optical imaging available (with HST) is of the order of  $\sim 1/10$  of a pixel, i.e.,  $\sim 5$  mas, corresponding to  $\Delta\theta/\theta_E \sim 5 \times 10^{-3}$  for galaxies as lenses. Higher accuracy in lens modeling is thus currently not required. Second, real mass distributions are not really smooth, but contain substructures; these substructures are almost certainly responsible for the mismatch between observed flux ratios of images, and the magnification ratios obtained from simple (i.e., smooth) lens models (e.g., Mao & Schneider 1998; Kochanek & Dalal 2004; Bradač et al. 2004; Xu et al. 2010). In addition, low-mass halos along the line-of-sight to the source may change magnifications (e.g., Metcalf 2005; Xu et al. 2012). For these reasons, flux ratios are usually not used as constraints in lens modeling. The upper mass end of the substructure can in addition cause positional shifts of individual images; such astrometric distortions have been observed in several lens systems where the substructure was indeed identified (e.g., MG 0414+0534 – e.g., Trotter et al. 2000; Ros et al. 2000 – and MG 2016+112 – e.g., Koopmans et al. 2002). Thus, independent of the accuracy of the observed image positions, for physical reasons one may not expect to reproduce the observed position to better than a few milliarcseconds with a smooth mass model. Hence, as long as the transformed deflection angle  $\hat{\alpha}$  (which yields exactly the same strong lensing properties as the original lens) and the deflection obtained from  $\hat{\kappa}$  differ by less than the smallest angular scale on which modeling by a smooth mass distribution is still meaningful, this difference is of no practical relevance.

In the near future, images of structured Einstein rings may be observed at higher resolution (i.e.  $0''.001$ ) with ALMA (Hezaveh et al. 2013), the Extremely Large Telescope (ELT) or the James Webb Space Telescope (JWST). It still has to be investigated how critical could the SPT be for sources observed at

those resolutions if substructures are not explicitly included in the lens models and/or if reasonable freedom is allowed regarding the angular structure of the lens (e.g. Evans & Witt 2003; Saha & Williams 2006). Nevertheless, the expected substructure and line-of-sight inhomogeneities mentioned above will put a lower limit to the positional accuracy at which lens systems can be modeled with smooth matter distributions. In any case, since the SPT is more general than the MST, it will limit the accuracy of some applications of strong gravitational lensing such as the detection of substructures via the time-delay method (Keeton & Moustakas 2009), the estimate of  $H_0$  for systems where only one time delay is measured (e.g. Vuissoz et al. 2007; Suyu 2012), or the determination of the profile of the dark matter distribution based only on strong lensing (e.g. Cohn et al. 2001; Suyu et al. 2012; Eichner et al. 2012). Free-form lens modeling might already give a hint of the impact of the SPT on some of these applications (Coles 2008; Paraficz & Hjorth 2010; Leier et al. 2011).

It should also be pointed out that ‘simple’ mass models of strong lensing clusters typically fail to reproduce the location of multiple images at the level of  $\sim 1''$ , even if the contributions of the cluster galaxies are explicitly accounted for. This level of mismatch is typically not considered to be a problem, since one expects the total mass distribution of clusters to be more complicated than that of one or a few large-scale mass components plus the mass profiles of cluster galaxies (for which simple scalings between luminosity and mass properties are employed). Here, the relative mismatch is  $\Delta\theta/\theta_E \sim 1/20$ , i.e., larger than for galaxies, and correspondingly, the required agreement between  $\hat{\alpha}$  and the deflection from  $\hat{\kappa}$  is less stringent.

Free-form lens models developed for studying strongly lensed sources by galaxies or clusters (e.g. Saha & Williams 2004; Diego et al. 2005; Liesenborgs et al. 2006; Coe et al. 2008) fit perfectly the lensing observables and derive ensemble of models reproducing existing data, for sets of individual multiply-imaged sources. Those techniques effectively explore degenera-

cies between lens models (Saha & Williams 2001, 2006), but with the drawback that many non-physical models (e.g., dynamically unstable or with arbitrary substructure) can be obtained. As emphasized by Coe et al. (2008), there is no unambiguous set of criteria to define a priori whether a model is physical or not. Hence, free-form modeling may sometimes worsen the impact of degeneracies by exploring a parameter space which is too large. More work is surely needed to develop schemes allowing to explore lens degeneracies such as the SPT in a controlled way, maybe implying to combine model-based and model-free approaches, or include perturbations around local solutions (Alard 2009).

We have illustrated the behavior of the SPT for a quadrupole lens. As an example, we have shown that in the presence of an external shear  $\gamma \sim 0.1$ , a mass model constituted of a baryonic component (modeled as a Hernquist profile) and of a dark matter component (modeled as a generalized NFW) could be transformed into a power-law model with a finite core. In fact, as mentioned before, this example was not constructed as an SPT, but the (almost perfect) degeneracy between these two mass models was found ‘experimentally’ in SS13. We have shown here that this degeneracy nevertheless can be traced back to the SPT; in particular, we saw that the SPT corresponding to that transformation is a spatially varying (at less than a percent level) and nearly isotropic contraction of the source plane positions, i.e.  $\hat{\beta} = [1 + f(\beta)]\beta$ . We have shown that the magnifications are transformed such that  $\hat{\mu} = [1 + f(\beta)]^2 \mu$ , while the time delay transforms differently. In addition, the time delay ratios are not conserved when four lensed images are formed.

Throughout this paper we have considered strong lensing only. The weak lensing properties are not invariant under the SPT, since it changes the weak lensing observable, i.e., the reduced shear  $\gamma/(1 - \kappa)$ , except for the special case of a MST Schneider & Seitz (1995). Thus, the SPT is not an invariance transformation for weak lensing studies. In the strong lensing regime, if magnification information can be obtained from observations, the invariance with respect to the SPT can also be broken. The latter most likely is more relevant for clusters, where magnification can be estimated from the observed number density of background sources, though these estimates have a considerable uncertainty within the strong lensing regime of clusters.

In a future work, we aim to quantify the consequences of the asymmetry of  $\hat{\mathcal{A}}$  in more detail, and thus to find criteria to derive which kind of SPTs are allowed for a given tolerance in the changes of image positions.

*Acknowledgements.* Part of this work was supported by the German *Deutsche Forschungsgemeinschaft*, DFG project number SL172/1-1.

## References

- Alard, C. 2009, A&A, 506, 609  
 Bartelmann, M. 2010, Classical and Quantum Gravity, 27, 233001  
 Bradač, M., Schneider, P., Lombardi, M., et al. 2004, A&A, 423, 797  
 Burke, W. L. 1981, ApJ, 244, L1  
 Coe, D., Fuselier, E., Benítez, N., et al. 2008, ApJ, 681, 814  
 Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R. 2001, ApJ, 554, 1216  
 Coles, J. 2008, ApJ, 679, 17  
 Courbin, F., Chantry, V., Revaz, Y., et al. 2011, A&A, 536, A53  
 Diego, J. M., Sandvik, H. B., Protopapas, P., et al. 2005, MNRAS, 362, 1247  
 Eichner, T., Seitz, S., & Bauer, A. 2012, MNRAS, 427, 1918  
 Evans, N. W. & Witt, H. J. 2003, MNRAS, 345, 1351  
 Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1  
 Fassnacht, C. D., Xanthopoulos, E., Koopmans, L. V. E., & Rusin, D. 2002, ApJ, 581, 823  
 Hezaveh, Y., Dalal, N., Holder, G., et al. 2013, ApJ, 767, 9  
 Keeton, C. R. 2001, astro-ph/0102340  
 Keeton, C. R. & Moustakas, L. A. 2009, ApJ, 699, 1720  
 Kochanek, C. S. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 91–268  
 Kochanek, C. S. & Dalal, N. 2004, ApJ, 610, 69  
 Koopmans, L. V. E., Garrett, M. A., Blandford, R. D., et al. 2002, MNRAS, 334, 39  
 Leier, D., Ferreras, I., Saha, P., & Falco, E. E. 2011, ApJ, 740, 97  
 Liesenborgs, J. & De Rijcke, S. 2012, MNRAS, 425, 1772  
 Liesenborgs, J., De Rijcke, S., & De Jonghe, H. 2006, MNRAS, 367, 1209  
 Mao, S. & Schneider, P. 1998, MNRAS, 295, 587  
 Metcalf, R. B. 2005, ApJ, 629, 673  
 Paraficz, D. & Hjorth, J. 2010, ApJ, 712, 1378  
 Read, J. I., Saha, P., & Macciò, A. V. 2007, ApJ, 667, 645  
 Ros, E., Guirado, J. C., Marcaide, J. M., et al. 2000, A&A, 362, 845  
 Saha, P. & Williams, L. L. R. 1997, MNRAS, 292, 148  
 Saha, P. & Williams, L. L. R. 2001, AJ, 122, 585  
 Saha, P. & Williams, L. L. R. 2004, AJ, 127, 2604  
 Saha, P. & Williams, L. L. R. 2006, ApJ, 653, 936  
 Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 1–89  
 Schneider, P. & Seitz, C. 1995, A&A, 294, 411  
 Schneider, P. & Sluse, D. 2013, arXiv-1306.0901  
 Suyu, S. H. 2012, MNRAS, 426, 868  
 Suyu, S. H., Hensel, S. W., McKean, J. P., et al. 2012, ApJ, 750, 10  
 Tewes, M., Courbin, F., Meylan, G., et al. 2012, arXiv-1208.6009  
 Trotter, C. S., Winn, J. N., & Hewitt, J. N. 2000, ApJ, 535, 671  
 Vuissoz, C., Courbin, F., Sluse, D., et al. 2007, A&A, 464, 845  
 Xu, D. D., Mao, S., Cooper, A. P., et al. 2012, MNRAS, 421, 2553  
 Xu, D. D., Mao, S., Cooper, A. P., et al. 2010, MNRAS, 408, 1721