The signature of substructure on gravitational lensing in the ΛCDM cosmological model

M. Bradač, P. Schneider, M. Lombardi, M. Steinmetz, L.V.E. Koopmans, and Julio F. Navarro

1 Institut für Astrophysik und Extraterrestrische Forschung, Auf dem Hügel 71, D-53121 Bonn, Germany
2 Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany
3 European Southern Observatory, Karl-Schwarzschild Str. 2, D-85748 Garching bei München, Germany
4 Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany
5 Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721, USA
6 Kapteyn Astronomical Institute, P.O.Box 800, 9700AV Groningen, The Netherlands
7 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
8 Department of Physics and Astronomy, University of Victoria, BC V8P 1A1, Canada

Submitted to A&A

Abstract. We present a study of the lens properties of quadruply imaged systems, lensed by numerically simulated galaxies. We investigate a simulated elliptical and disc galaxy drawn from high resolution simulations of galaxy formation in a concordance ΛCDM universe. The simulations include the effects of gas dynamics, star formation and feedback processes. Flux-ratio anomalies observed in strong gravitational lensing potentially provide an indicator for the presence of mass substructure in lens galaxies as predicted from CDM simulations. We particularly concentrate on the prediction that, for an ideal cusp caustic, the sum of the signed magnifications of the three highly magnified images should vanish when the source approaches the cusp. Strong violation of this cusp relation indicates the presence of substructure, regardless of the global, smooth mass model of the lens galaxy. We draw the following conclusions: (1) the level of substructure present in simulations produces violations of the cusp relation comparable to those observed, (2) higher-order catastrophes (e.g. swallowtails) can also cause changes of the order of 0.6 in the cusp relation as predicted by a smooth model, (3) the flux anomaly distribution depends on the image parity and flux and both the brightest minimum and saddle-point images are more affected by substructure than the fainter images. In addition, the brightest saddle point is demagnified w.r.t. the brightest minimum. Our results are fully numerical and properly include all mass scales, without making semi-analytic assumptions. They are ultimately limited by the mass resolution of single particles in the simulation determined by current computational limits, however show that our results are not affected by shot-noise due to the finite number of particles.

Key words. cosmology: dark matter – galaxies: structure – gravitational lensing

1. Introduction

Whereas the current Cold Dark Matter (CDM) paradigm for structure formation is widely accepted, two major problems for CDM remain. While simulations predict cuspy dark matter halos (e.g. Moore 1994), observed rotation curves of low surface brightness galaxies indicate that their dark matter halos have more shallow cores (Kravtsov et al. 1998, Swaters et al. 2000, van den Bosch & Swaters 2001, de Blok & Bosma 2002). The other is the apparent over-prediction of the small-scale power in CDM simulations. As was shown by Moore et al. (1999) and Klypin et al. (1999), the number of satellite halos seen in N-body simulations appears to far exceed the number of dwarf galaxies observed around the Milky Way. Particular discrepancies have been found for satellite masses \( \lesssim 10^9 M_\odot \).

Gravitational lensing is at present the only tool to investigate CDM substructure in galaxies outside the local group. As first noted by Mao & Schneider (1998), mass-substructure other than stars on scales less than the image separation can substantially affect the observed flux ratios in strong gravitational lens systems. Chiba (2002), Dalal & Kochanek (2002), Mao & Schneider (1998), Metcalf & Madau (2001), Metcalf & Zhao (2002), Metcalf et al. (2003), Keeton (2001), and Bradač et al. (2002) have all argued that substructure can provide the explanation for the flux anomalies in various sys-
tems. Dalal & Kochanek (2002) further conclude that the amount of substructure needed to explain the flux ratios of quadruply-imaged systems broadly agrees with the CDM predictions. At least for some systems the flux mismatches are probably not just an artifact of oversimplified macromodels of the main lens galaxy (see e.g. Evans & Witt 2003; Metcalf & Zhao 2002). As discussed by Keeton (2003) and Chen et al. (2003), fluxes can be further affected by clumps of matter at a redshift different from that of the lens, along the line of sight between the observer and the source; however, this effect is not dominant. It is also possible that the small scale structure does not consist of compact CDM clumps, also tidal streams or offset disc components can affect the flux ratios (see Möller et al. 2003; Quadri et al. 2003).

Keeton (2001) and Gaudi & Petters (2002) recently focused on the magnification relations that should be satisfied by particular four-image geometries (so called “fold” and “cusp” configurations). These relations are model-independent predictions for the magnifications of highly magnified images (Blandford & Narayan 1986; Blandford 1990; Schneider & Weiss 1992; Mao 1992). Strictly speaking, however, they hold only for ideal “fold” or “cusp” configurations and it is therefore in some cases hard to disentangle the effects of the source being further away from the cusp from the effect of substructure, purely by employing these relations.

The influence of substructure can not only be seen on image flux ratios, but also in the structure of multiple-imaged jets. The lens system B1152+199 consists of a doubly-imaged jet, one of which appears bent, whereas the other is not (Metcalf 2002). Alternative explanation is that an intrinsic bend in the jet is simply magnified in one image, and produces only a small effect in the other.

Microlensing can change the flux ratios not only in the optical (e.g. Woźniak et al. 2000), but also at radio wavelengths (Koopmans & de Bruyn 2000). Flux ratio anomalies can also be introduced by propagation effects in the interstellar medium (ISM) in the lens galaxy, by galactic scintillation, and scatter broadening (Koopmans et al. 2003). Fortunately, these effects are frequency dependent and one can distinguish them using multi-frequency observations. In addition, these electromagnetic phenomena are similar for images of different parities.

For substructure, on the other hand, Schecter & Wambsganss (2004) found that magnification perturbations should show a dependence on image parity. Microlensing simulations showed that the probability distributions for magnifications of individual images are not symmetric around the unperturbed magnification. The distribution depends on image parity and becomes highly skewed. The probability for the brightest saddle point image to be demagnified is increased.1

Observed lens systems also seem to show this image parity dependence (Kochanek & Dalal 2003), and this indicates that the flux ratio anomalies arise mainly from gravitational lensing, rather than propagation effects.

All these effects on flux ratios have placed some doubt as to whether the existence of substructure can be rigorously tested with strong lensing and what is expected signal. Several groups have tested the effects of substructure in strong lensing systems using a semi-analytic prescription for substructure (Metcalf & Madau 2001; Keeton et al. 2003; Dalal & Kochanek 2002; Kochanek & Dalal 2003). Recently Mao et al. (2004) showed using high-resolution numerical simulations that the fraction of surface mass density in substructures is lower than required by lensing; however both predictions are still uncertain.

In this paper we use different projections of two different galaxies obtained in N-body-gasdynamics simulations. Whereas a semi-analytic prescription overcomes the problem of shot-noise, and the problem of modelling becomes simpler (one has an analytic model for the underlying macro-distribution), by using the direct output of an N-body simulated galaxy, one does not make any simplifying assumption about the mass profiles of the macro model, or the substructure. Down to the resolution scales of the simulation we therefore believe we have a better comparison with a realistic galaxy.

Whereas higher-resolution DM-only simulations are available, the absence of baryons significantly affects lens properties. Those type of simulations are of limited use for the purpose of testing CDM substructure effect on strong lens systems. More precise, strong gravitational lensing is probing the galaxy potential on the inner 5 – 10 kpc (the typical size of the Einstein radius). It was shown by Treu & Koopmans (2004) that the range of projected baryonic fraction within the Einstein radius is $f_{bar} = 0.3$ – 0.6. It is therefore crucial to include the gravitational pull of baryons in our simulations.

This paper is structured as follows. In Sect. 2 we first give the main properties of the N-body simulations that we use. We also introduce an improved smoothing scheme compared to Bradač et al. (2002) and describe how to extract lensing properties from N-body simulations. In Sect. 3 we focus on the cusp relation for simulated lens systems. Sect. 4 describes the modelling of synthetic images and the phenomenon of suppressed saddle points. We conclude and give an outlook in Sect. 5.

1 Images form at extrema of the arrival time surface (i.e. Fermat’s principle). At saddle points, images have negative and at minima/maxima, they have positive parity. In quadruply-imaged system one observes two saddle point images, and two minima. The fifth image is a maximum, too faint to be observed.

2. Strong lensing by a simulated galaxy

N-body simulations can provide a powerful benchmark for testing the effects of substructure on strong lensing. One can simulate conditions in which propagation effects due to the ISM can be ignored and thus examine only the signature of substructure. The drawback of this method at
present lies in the resolution available for simulations that include dark matter, gas and star particles. This limits our analysis to mass clumps of $\gtrsim 10^8 M_\odot$. However, since the mass resolution is improving rapidly, this will soon be less of a problem.

As in Bradač et al. (2002), we used the numerical N-body simulations for several realizations of galaxies including gas-dynamics and star formation (Steinmetz & Navarro 2001). We investigate two different halos, each of them in three different projections. The simulations were performed using GRAPESPH, a code that combines the hardware N-body integrator GRAPE with the Smooth Particle Hydrodynamics (SPH) technique (Steinmetz 1990).

In Table 1, the properties of the halos are listed. In both cases the original simulated field contains approximately 300,000 particles. The simulation is contained within a sphere of diameter 32 Mpc which is split into a high-resolution sphere of diameter 2.5 Mpc centred around the galaxy and an outer low-resolution shell. Gasdynamics and star formation are restricted to the high-resolution sphere, while the dark matter particles of the low-resolution sphere sample the large scale matter distribution in order to appropriately reproduce the large scale tidal fields (see Navarro & Steinmetz 1997 and Steinmetz & Navarro 2001 for details on this simulation technique). From the original simulated field we use a cube of size $\sim 200^3 \text{kpc}^3$ centred on a single galaxy. This volume lies well within the high-resolution sphere and is void of any massive intruder particles from the low-resolution shell.

All simulations were performed in a ΛCDM cosmology ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_\Lambda = 0.019/h^2$, $\sigma_8 = 0.9$). They have a mass resolution of $1.26 \times 10^7 M_\odot$ and a spatial resolution of 0.5 kpc. A realistic resolution scale for an identified substructure is typically assumed to be $\sim 40$ particles which corresponds to $5 \times 10^4 M_\odot$. The quoted mass resolution holds for gas/stars. The high-resolution dark matter particles are about a factor of 7 (i.e. $\Omega_0/\Omega_b$) more massive. A detailed analysis of the photometric and dynamical properties of the simulated halos was carried out in Meza et al. (2003) for the elliptical and Abadi et al. (2003a,b) for the disc galaxy.

Early N-body simulations suffered from the problem of overmerging, i.e. most satellites were dissolved due to the tidal field of the host galaxy. Current N-body simulations demonstrate that quite often this tidal disruption is inefficient resulting in surviving substructures. While this qualitatively different result is usually accounted to the increased particle numbers of modern simulations, it is in fact rather caused by a more prudent choice of the numerical softening compared with a more rigorous limit on the numerical time stepping. Convergence studies using modern N-body simulations demonstrate that with a sufficiently small numerical softening length and sufficiently rigorous numerical time stepping, mass functions can be accurately produced down to clumps with only a very few tens of particles (see e.g. Springel et al. 2001).

2.1. Delaunay triangulation smoothing technique

From the irregularly sampled particle distribution in the simulation box, we reconstruct the density field. We apply a smoothing procedure and then project the resulting particle distribution to obtain the surface mass density $\kappa$. In Bradač et al. (2002) it was shown that a more sophisticated smoothing method should be employed for the data analysis than simply smoothing with a Gaussian kernel. A method is needed that adapts the kernel size in order to increase the signal to noise of the reconstructed field. For this purpose, we make use of the Delaunay tesselation technique from Schaap & van de Weygaert (2000).

We perform a three-dimensional Delaunay triangulation using the QHULL algorithm (Barber et al. 1996). The density estimator from Schaap & van de Weygaert (2000) is then evaluated at each vertex, and we interpolate values of the density at each three-dimensional grid point to obtain the $\kappa$ map. The resulting density field is then projected onto a two-dimensional grid.

Since the N-body simulations contain three independent classes of particles (gas, stars, and dark matter particles, each having different masses), we repeat the procedure described above for each group separately and the final $\kappa$ map is obtained by adding the contributions from all three classes.

The Delaunay tesselation method performs very well in comparison to the standard Gaussian smoothing technique used in Bradač et al. (2002), or an adaptive Gaussian smoothing technique (see also Schaap & van de Weygaert 2000). Since it is non-parametric, it adjusts the scale of the smoothing kernel such that regions of low noise (i.e. where the particle density is highest) are effectively smoothed less than regions with high noise. Also the shape of the kernel is self-adaptive. Hence, this method is very useful for the analysis of galaxies with high dynamic range and

| Halo | Elliptical | Disk |
|------|------------|------|
| $z_{\text{i}}$ | 0.81 | 0.33 |
| $z_{\text{s}}$ | 3.0 | 3.0 |
| $N_{\text{bar}}$ | 12 000 | 20 000 |
| $N_{\text{DM}}$ | 17 000 | 26 000 |
| $N_{\text{str}}$ | 70 000 | 110 000 |
| $M_{\text{tot}}$ | $1.5 \times 10^{12} M_\odot$ | $0.5 \times 10^{12} M_\odot$ |
significant structure in the mass distribution (e.g. mass clumps and spiral arms).

### 2.2. Estimating the noise properties

A drawback of using the Delaunay tesselation method is that the signal-to-noise evaluation for the final surface mass density map is non-trivial. For example, with a Gaussian kernel one can determine the noise by simply looking at the number of particles in a smoothing element (for a more detailed estimate, see Lombardi & Schneider 2002). When using the tesselation technique, such an approach is not viable.

One approach for estimating the error is to use bootstrapping (see e.g. Heyl et al. 1994). Normally we calculate physical properties by using all $n$ particles in the simulation. To create a bootstrap image one has to randomly select $n$ particles out of this simulation with replacement; i.e. some of the particles from the original simulation will be included more than once and some not at all. In other words, we randomly generate $n$ integers from 1 to $n$, representing the bootstrapped set of particles. If a particle is included $k$ times in a bootstrapped map, we put a particle at the same position with $k$ times its original mass. One can then make an error estimation using the ensemble of such images and calculating the desired physical quantity for each of them.

Whereas the tesselation itself is done very quickly, the interpolation of density on a grid is a process that takes a few CPU days on a regular PC. We therefore limit ourselves to 10 bootstrapped maps and the elliptical halo only.

For each pixel $i$ we calculate the associated error using

$$\sigma^2(\kappa_i) = \frac{1}{M-1} \sum_{m=1}^{M} (\kappa_i^m - \langle \kappa_i \rangle)^2$$  \hspace{1cm} (1)

where $\langle X_i \rangle$ is the average value of $X$ at pixel $i$ averaged over all $M$ bootstrapped maps; in our case $M = 10$. This procedure gives us an estimate for the error on $\kappa$. We find that within the critical curves, i.e. $\kappa \gtrsim 0.25$ the average noise $\sigma(\kappa_i)/\langle \kappa \rangle$ is of the order of $\lesssim 5\%$.

### 2.3. Strong lensing properties

Having obtained the $\kappa$-map, we then calculate the lens properties on the grid ($2048 \times 2048$ pixels). The Poisson equation for the lens potential $\psi$

$$\nabla^2 \psi(\theta) = 2 \kappa(\theta)$$  \hspace{1cm} (2)

is solved on the grid in Fourier space with a DFT (Discrete Fourier Transformation) method. We used the publicly available C library FFTW (“Fastest Fourier Transform in the West”) written by Frigo & Johnson (1998). To reduce the boundary effects, padding was introduced. In particular, the DFT was performed on a $4096 \times 4096$ grid, where one fourth of the grid contains the original $\kappa$ map and the rest is set to zero.

One can now proceed in two ways. Either, one can calculate the two components of the shear $\gamma_{1,2}$ and the deflection angle $\alpha$ by multiplying the potential in Fourier space by the appropriate kernel. However, since we calculate the Fourier transform of the potential on a finite...
grid, we filter out high spatial frequency modes. By multi-
plying the transform with different kernels for $\gamma_1,2$ and $\alpha$, these final maps do not correspond exactly to the same $\kappa$ map. The effect is small, but it shows up near the critical curves. Therefore, it is better to only calculate the lens potential $\psi$ using DFT and obtain the shear and deflection angle by finite differencing. The latter method is also less CPU-time and memory intensive.

The simulated galaxies are a field galaxy. However, most of the lenses in quadruple image systems are members of groups. To make our simulated galaxies more closely resemble a realistic system, we add an external shear to the Jacobi matrix (evaluated at each grid point). The shear components are the same for each projection and all halos. We use

$$\gamma_1^{\text{ext}} = -0.04, \quad \gamma_2^{\text{ext}} = -0.16;$$

corresponding to the shear of the best-fit singular-

isothermal ellipse model with external shear for the lens $B1422+231$ in [Bradač et al. 2002]. Figure 1 show the magnification maps of the elliptical and edge-on disk galaxy. The corresponding caustic curves (for a source at $z = 3$) were obtained by projecting the points with vanishing determinant of the Jacobian matrix from the image plane to the source plane. The critical curves are plotted in white in Fig. 1A for the elliptical and Fig. 1B for the edge-on disk galaxy. The corresponding caustic curves are plotted in Fig. 2A, B, and C respectively.

### 2.4. The importance of baryons

As mentioned above, the influence of baryons is very important in lens galaxies. Since we are interested in the effects of substructure, it would be desirable to use N-body simulations that have the highest possible dynamical range. At present this is achieved in high-resolution N-body simulations that only include dark matter. However, if baryons are not included, the central potential is more shallow than what we typically observe in lens galaxies. All quadruply-imaged systems for which the inner slope of the mass distribution has been measured, are well described by a total mass density profile $r^{-\gamma}$ with $\gamma \approx 2$ (Kochanek 1995, Cohn et al. 2001, Treu & Koopmans 2002, T. Koopmans & Treu 2003, Treu & Koopmans 2004), consistent with the combined mass distribution of dark matter and baryons seen in the simulations. Hence, dark matter only simulations do not accurately represent the overall properties of lens galaxies, and instead we need to use hydro-dynamical simulations.

To test the importance of baryons we have simulated an elliptical halo using only dark matter particles and performed the same lensing analysis as described above. We project the density approximately along the long axis of the halo, thus maximising the central density. In Fig. 2 we plot the corresponding caustic curves for the halo simulations; in panel (a) the simulation including baryons (DM+B), for panel (b) we use the DM-only simulation (DM). The radial critical curve in the DM+B halo is located close to the galaxy centre and is therefore not well resolved (but also irrelevant for our purposes) - the corresponding caustic is therefore not plotted.

The two caustic curves are very different, indicating very different overall strong lensing properties of these two simulations. Whereas the DM+B simulation has a steep inner mass profile (very close to singular isothermal), the DM simulation has a caustic configuration typical for a lens with a shallow density profile (see e.g. [Wallington & Narayan 1993]). The radial critical curve has become smaller compared to the DM+B halo and the corresponding caustic curve is almost entirely enclosed by the asteroid caustic. The prominent naked cusp region is a three-image region. This configuration is extremely rare among the observed lensed systems. For a source located in such a region one would observe three highly magnified images. There is only one possible example of a triply imaged quasar out of $\sim 50$ doublets and triplets (Evans & Hunter 2002), namely APM 08275+5255 (Ibata et al. 1999). Further, the vast majority of systems similar to B1422+231 cannot form in such a potential. We therefore conclude that the effects of baryons have to be included, and we will use only DM+B simulations from now on, discussing their limits where necessary.

### 3. The cusp relation

We generate different four-image systems using each simulated galaxy. For each projection, regions in the source plane where five images form are determined. The image plane is projected back to the source plane using the magnification and deflection angle maps. We used the grid search method from Blandford & Kochanek (1987) to find the pixels enclosed by the asteroid caustic and approximate image positions. Then the MNEWT routine from Press et al. (1992) is applied and the lens equation is solved for the image positions. For this step, we interpolate the deflection angle between the grid points. We use bilinear and bicubic spline interpolation, and both methods give comparable results. Once we have the image positions, their magnifications are calculated and the four brightest images are chosen. These represent the “observable” images; the fifth image is usually too faint (except in the regions where more than five images are formed) and therefore likely to escape observation.

We can now investigate the basic properties of the synthetic four-image systems. There are three basic configurations: the fold, cusp, and cross. They correspond to a source located inside the asteroid caustic, close to a fold, a cusp or near the centre, respectively (see e.g. Keeton 2001). All configurations have been observed, and even though one would naively think that fold and cusp images are rare among observed lenses, they are in fact frequently observed due to the large magnification bias. In this section we will mainly concentrate on cusp image configurations.
Fig. 2. The caustic curves of two simulated galaxies. Panel (a) represents the simulation which includes baryonic and dark matter particles, whereas for panel (b) we use a simulation with dark matter particles only. The radial caustic for the dark matter only simulation is almost entirely enclosed within the asteroid caustic, prohibiting formation of cusp images in quadruply-imaged systems.

The behaviour of gravitational lens mapping near a cusp was first studied by Blandford & Narayan (1986), Blandford (1990), Schneider & Weiss (1992), and Mao (1992), who investigated the magnification properties of the cusp images and concluded that the sum of the signed magnification factors of three merging images approaches zero as the source moves towards the cusp. In other words,

$$R_{\text{cusp}} = \frac{|\mu_A + \mu_B + \mu_C|}{|\mu_A| + |\mu_B| + |\mu_C|} \to 0, \quad \text{for } \mu_{\text{tot}} \to \infty,$$

where $\mu_{\text{tot}}$ is the unsigned sum of magnifications of all four images, and $A$, $B$ and $C$ is the triplet of images forming the smallest opening angle (Keeton 2001). The opening angle is measured from the galaxy centre and is spanned by the two images of equal parity. The third image lies inside this opening angle.

3.1. The cusp relation of an N-body simulated elliptical galaxy

The cusp relation is an asymptotic relation and holds when the source approaches the cusp from inside the asteroid. One can derive the properties of lens mapping close to critical curves using a Taylor expansion of the Fermat potential around a critical point (see e.g. Schneider & Weiss 1992). Such calculations are very cumbersome and therefore it is difficult (if not impossible) to explore the influence of arbitrary substructure analytically. In practice, we can calculate $R_{\text{cusp}}$ for the N-body simulated systems. Smoothing the original $\kappa$-map on different scales then gives an indication of the influence of substructure on $R_{\text{cusp}}$.

The three cusp images [designated as A, B and C in Fig. 3] are chosen according to the image geometry. Since we know the lens position, this procedure is straightforward and foolproof. We have identified the triplet of images belonging to the smallest opening angle (described above). Since we know the image parities and magnifications, one is tempted to identify the three brightest images as the cusp images and assign different parity to the brightest one than to the other two (e.g. as in Möller et al. 2003). However, due to the presence of shear and substructure this could lead to misidentifications.

Figure 3a shows the caustic curve in the source plane for the simulated elliptical at a redshift of $z_l = 0.81$. The source is at a redshift of $z_s = 3$. Approximately 30 000 lens systems are generated with source positions inside the asteroid caustic. $R_{\text{cusp}}$ is plotted in gray-scale. The apparent discontinuities originate from different image identification. In the very centre of the caustic the meaning of “cusp image” is ill defined. As the source moves in the direction of the minor or major axes we chose different subsets of three cusp images and therefore the discontinuity arises.

The remaining panels of Fig. 3 show the effect of smoothing the small-scale in the surface mass density $\kappa$ map with Gaussian kernels characterised by standard deviation $\sigma_G$. The values for $\sigma_G$ were chosen to be $\sigma_G \sim 1, 2, 5$ kpc for panels (b), (c) and (d) respectively. Note that we do not smooth the $\kappa$-map directly. First we obtain the smooth model $\kappa_{\text{smooth}}$ for the $\kappa$-map by fitting elliptical contours to the original map using IRAF.STSDAS package ellipse. We subtract $\kappa_{\text{smooth}}$ from $\kappa$ and smooth the difference using different Gaussian kernels. We then add the resulting map back to the $\kappa_{\text{smooth}}$. In this way
Fig. 3. The cusp relation $R_{\text{cusp}}$ for the N-body simulated elliptical galaxy at a redshift of $z_l = 0.81$. The source was put at a redshift of $z_s = 3$ and approx. 30,000 systems were generated that lie inside the asteroid caustic. $R_{\text{cusp}}$ is plotted in gray-scale, for sources close to the cusp the smooth models would predict $R_{\text{cusp}} \sim 0$ (i.e. white - red). The deviations are due to the substructure. Due to magnification bias most of the observed lenses correspond to the fold and cusp configurations. Discontinuities in the maps arise when the source moves in the direction of the minor or major axes, since we chose different subsets of three cusp images. On top we plot the caustic curve. Panel (a) shows $R_{\text{cusp}}$ for the original mass distribution, whereas panels (b)–(d) show the cusp relation for the models where we additionally smoothed the substructure (see text) with a Gaussian kernel characterised by standard deviation $\sigma_G \sim 1 \text{ kpc}$ (b), $\sigma_G \sim 2 \text{ kpc}$ (c) and $\sigma_G \sim 5 \text{ kpc}$ (d).

The effect of smoothing on the cusp relation is clearly visible. In Fig. 3b one sees that the substructure is completely washed out when smoothing on scales of $\sigma_G \sim 5 \text{ kpc}$ is applied. As we go to smaller smoothing scales, the effects of substructure become clearly visible. In the extensions of swallowtails there is a region where the cusp relation is strongly violated (with $R_{\text{cusp}} \sim 0.6$, where the smooth model predicts $R_{\text{cusp}} \lesssim 0.1$). However, further out, a swallowtail can cause the cusp relation to change the trend and go to zero (due to high-magnification systems being formed in such region).

Finally, the cusp relation behaves differently for the source on the major or the minor axis (see especially Fig. 3d). This is a generic feature for smooth elliptical
models and can easily be calculated for e.g. an elliptical isopotential model (see e.g. Schneider et al. 1992). We use this model since it is analytically tractable for source positions along the major and minor axis. In Fig. 4 we plot the cusp relation for the source moving along the major (minor) axis as a thick (thin) solid line for the elliptical isopotential model with $\epsilon = 0.15$. As the source approaches the cusp, $R_{\text{cusp}} \rightarrow 0$ for both source positions, however the slope is different. We also plot the total magnification factor of the three cusp images, i.e. $\mu_{A+B+C} = |\mu_A| + |\mu_B| + |\mu_C|$ as a thick (thin) dashed line.

**Fig. 4.** $R_{\text{cusp}}$ for a simple elliptical isopotential lens model with $\epsilon = 0.15$. $R_{\text{cusp}}$ is plotted as a thick (thin) solid line for sources along the major (minor) axis. $\theta_{\text{open}}$ represents the angle measured from the position of the galaxy, spanned by two “outer” cusp images (A and C). The opening angle $\tau$ means that the source is located at the centre (images A, C and the galaxy lie on the same line). When the source approaches the cusp, $\theta_{\text{open}} \rightarrow 0$. The total magnification for the three cusp images $\mu_{A+B+C}$ is plotted as a thick (thin) dashed line for sources along the major (minor) axis.

### 3.2. The cusp relation in an N-body simulated disk galaxy

The procedure described above was applied to three different projections of the elliptical halo and very similar conclusions can be drawn. However, the question is, how much these conclusions depend on the specific morphological type of the galaxy. To investigate this we follow a similar procedure for a simulated disk galaxy.

In this case, however, we did not look at the effects of additional smoothing. It is difficult to subtract a smooth model, since the galaxy consists of a bulge, warped disk and extended halo, which can not simply be fitted by ellipses. If we were to smooth the edge-on projection with a

**Fig. 5.** The cusp relation for the N-body simulated disk galaxy at a redshift of $z_l = 0.33$. The source was put at a redshift of $z_s = 3$ and approx. 10 000 systems were generated which lie inside the asteroid caustic. $R_{\text{cusp}}$ is plotted in grey-scale. On top we plot the caustic curve. Panels (a), (b), and (c) show three different orthogonal projections of the halo, (a) corresponds to the face-on, (b) and (c) to the edge-on projection. Panel (c) corresponds to the magnification map in Fig. 4.
Gaussian kernel, we would also wash out the (in our case warped) disc. We only include the analysis of the cusp relation for this halo in order to show the effect of the disc on $R_{\text{cusp}}$.

Figure 3 shows the cusp relation of an N-body simulated disk galaxy in three projections. The disc clearly plays a role for the edge-on projection (see Fig. 3c), whereas the face-on projection is similar to the elliptical galaxy. Especially in Fig. 3d, where the disc extent in projection is smaller than the typical image separation, the cusp relation in the upper-right and lower-left cusp is strongly violated. This direction corresponds to the orientation of the disc.

### 3.3. Observed cusp relation

In this section we compare predictions of $R_{\text{cusp}}$ from our simulated galaxies with the values from observed lens system. Unfortunately the number of observed systems is not large; seventeen four-image systems have been published (see Keeton et al. 2003 for a summary), and four of them are believed to show a typical cusp geometry (see Table 2).

Note that relation (3.3) is model independent and can only be broken in the presence of substructure on scales smaller than the image separation. Hence, if the smooth model provides an adequate description of the lens galaxy, one would expect $R_{\text{cusp}} \sim 0$ for these lenses. This is clearly not the case and for this reason it is difficult to explain their flux ratios using simple, smooth models.

If we make a comparison with simulations, one can see that the pattern in Fig. 3 clearly does not explain the observed $R_{\text{cusp}}$ of these four lenses. On the other hand, substructure on few kpc scales and below provides enough perturbations to $R_{\text{cusp}}$ to explain the observed values.

The question arises, however, whether from the value of $R_{\text{cusp}}$ we can conclude that we indeed see the effects of CDM substructure. Keeton et al. (2003) indeed argue that the cusp relation alone does not reveal anomalous flux ratios in B1422+231. Still, detailed modelling for this system shows that the flux ratios are anomalous. The difficulties in modelling B1422+231 are not only a consequence of a violation in cusp relation, but also that image D is a fainter than predicted from smooth models. On the other hand, the simulated disk galaxy shows violation of $R_{\text{cusp}}$ even though there are no clear mass clumps present in the region where images are formed. Hence one has to be cautious when making conclusions about the presence of CDM substructure based on the value of $R_{\text{cusp}}$ alone.

### 3.4. The influence of noise on the cusp relation

The simulated cusp-relation can be reliably compared with observations only if we know the noise properties of our simulations. The effects of noise and physical substructures need to be disentangled through a detailed analysis.

From the bootstrapping procedure (Sect. 2.2) we also get an estimate of the error in the cusp relation. We estimated the noise on $R_{\text{cusp}}$ using a similar technique as for $\kappa$ in (1). Moreover, we do not perform the analysis directly in the source plane by subtracting the maps pixel-by-pixel. The problem is that bootstrapping somewhat changes the shape of the caustic curve (see also Fig. 7). Since we never observe the source plane directly, in reality we can not distinguish between two shifted, but identical caustics. We therefore have to match different bootstrap maps in the image plane. We compare the image positions generated by each source in the original frame with those generated by bootstrapped lenses. Thus for each source position in the original map (see Fig. 3), we search for the source position in the bootstrapped map such that the four image positions from both maps differ as little as possible.

In principle, one can redo the ray-tracing and calculate image fluxes for the position in the original map using the bootstrapped-lens properties. However this is not necessary, since our grid in the source plane is fine enough and we only need an approximate error estimate. Figure 6 shows the estimated absolute error $\sigma(R_{\text{cusp},i})$ for the elliptical halo from Fig. 3h. As described above, each value of $\sigma(R_{\text{cusp},i})$, plotted in grey-scale, does not correspond to the error for a source position, but refers to the error of the system with image positions similar to the original ones.

The absolute error becomes slightly larger in the regions of the swallow tails. This is, however, not the effect of substructure vanishing in individual bootstrapped realisations. It is rather the effect of slight position changes of individual clumps. If one looks at the individual caustic curves of bootstrapped halos (see Fig. 7), the swallow-tails are present in all realisations, although they change their positions slightly. Since this hardly affects the image positions we cannot perfectly match the source positions $i'$ with the original source position $i$ in these regions; thus we are overestimating the true error.

We conclude that the values of $R_{\text{cusp}}$ in the close proximity of the cusp can be as high as $R_{\text{cusp}} = 0.6$, with the error of $\pm 0.1$ as determined from the bootstrapping. Smoothing the substructure on scales as large as $\sim 1$ kpc does not remove this effect, but reduces it. This is expected, since smoothing changes the profile of substructures.

### Table 2

The values for $R_{\text{cusp}}$ taken from Keeton et al. (2003) for four lens systems showing a typical cusp geometry.

| Lens                | $R_{\text{cusp}}$  |
|---------------------|---------------------|
| B2045+265          | 0.52 ± 0.04         |
| B0712+472          | 0.26 ± 0.06         |
| RX J0911+0551      | 0.22 ± 0.06         |
| B1422+231          | 0.18 ± 0.06         |
Finally, we investigate how well we can sample the smooth mass distribution given the number of particles in the original simulation. We take an ellipsoidal power-law density profile following $\rho(r) \propto r^{-2.9}$. The power-law index was chosen, to closely reproduce the surface mass density $\kappa$ of the simulation which follows $\kappa(\theta) \propto \theta^{-1.9}$ in the vicinity of the critical curve (see Sect. 4.1 for more details). The number of particles we use is the same as in the original N-body simulation. To each particle we assign the average mass of the original sample, leaving the total mass of the lens galaxy unchanged. We sample the density profile using a rejection method (e.g. Press et al. 1992). The resulting particle distribution was again adaptively smoothed using Delaunay tessellation and we perform the lensing analysis as in all previous cases.

The resulting cusp relation is given in Fig. 8a. We have chosen the profile such that the particle densities around the outer critical curves are similar in both cases. Only in that case can we reliably compare the noise properties of $R_{\text{cusp}}$. The fact that the caustic is larger than in the case of the N-body simulated elliptical halo (compare to Fig. 3a) is here of lesser importance; it arises due the difference in the central profile, far from the critical curve.

The absence of strong violations of $R_{\text{cusp}}$ close to the caustic in Fig. 8a as compared to Fig. 3a confirms that deviation of $R_{\text{cusp}}$ from zero is not dominated by the shot noise of the particles, but is due to genuine substructure in the N-body simulation. For a more quantitative comparison we show the probability distribution of $R_{\text{cusp}}$ for systems with $\mu_{\text{tot}} > 20$ in Fig. 5b as a solid line for the sampled smooth halo and as a dashed line for the original. The much tighter distribution for the sampled halo confirms the absence of strong violations of $R_{\text{cusp}}$ in the sampled smooth halo compared to the original simulation.

This analysis also shows the advantage of smoothing with Delaunay tesselation. If we only use a Gaussian kernel (of the same size as in Fig. 3b), the deviations in the cusp relation for the sampled particle distribution are much larger.

4. Lens models for synthetic image systems

To explore the saddle point demagnification phenomena we fit the synthetic image systems using a singular isothermal ellipse model with external shear (SIE+SH) (Kormann et al. 1994). We do not include the flux ratios in the fit, since they are affected by the substructure. Furthermore, we keep the lens position fixed. Using only image positions we thus have 7 parameters and 8 constraints (the parameters that we used are lens strength, two components of the ellipticity of the lens, two external shear components and the source position. The 4 image positions provide 8 constraints).

4.1. Surface mass density profile

We find that the average unsigned magnifications predicted by the SIE+SH model are higher than those from N-body simulations. This is true for all systems generated with the same halo; a consequence of the mass profile being steeper than isothermal around the typical position where the images are formed.

We have calculated the $\kappa$ profile for the N-body simulated elliptical by fitting the $\kappa$ with elliptical contours using the IRAF.STSDAS package ellipse. For the inner
The cusp relation for the smooth ellipsoidal model sampled with the same amount of particles as present in the original N-body simulated elliptical galaxy. The redshifts of the source and the lens were kept at $z_l = 0.81$ and $z_s = 3$. $R_{\text{cusp}}$ is plotted in gray-scale. On top we plot the corresponding caustic curve. The absence of large fluctuation in $R_{\text{cusp}}$ as compared to Fig. 3a shows that the signal is not dominated by the shot noise of the particles.

(b) The probability distribution of $R_{\text{cusp}}$ for systems with $\mu_{\text{tot}} > 20$. The solid line gives the probability distribution for the sampled smooth ellipsoidal model, while the dashed line corresponds to the original N-body simulation (cf. Fig. 3a).

\[ \sim 500 \text{ pc} \] the profile is very close to isothermal with slope $-1.0 \pm 0.2$, the profile becomes steeper than isothermal with slope $-1.9 \pm 0.2$. The break radius, where the profile becomes steeper than isothermal, is smaller than the radius where images form.

We therefore over-predict the magnifications using an isothermal profile. In order to deal with this problem, one would preferably use a power-law profile with the index mentioned above for the lens modelling, instead of SIE. However, these profiles require the evaluations of hypergeometric functions and/or numerical integrations (Grogin & Narayan [1996] Barkana [1998]. Modelling is therefore computationally too expensive to apply to several 10 000 sources. Moreover, this is not critical, since we are not searching for the best fit macro-model, but rather pretending that we observe the systems and try to fit them with the model used for most observed lenses. Since we only observe the flux ratios and not directly the magnifications, it is impossible to compare magnification factors in practice. Thus one cannot spot the difference in profiles using only this consideration when dealing with real lens systems.

4.2. Suppressed saddle points

It was first noted by Witt et al. [1995] that the expected flux changes due to stellar-mass perturbers differ between saddle points and minima images. This was further investigated by Schechter & Wambsganss [2002] who conclude that for fold images a stellar-mass component added to a smooth mass distribution can cause a substantial probability for the saddle point image to be demagnified w.r.t. the minimum.

Recently, Kochanek & Dalal [2003] and Keeton [2003] investigated the effect of singular isothermal sphere (SIS) mass clumps on the flux perturbations for images of different types. Their conclusions are similar; SIS perturbers also cause the brightest saddle point image to behave statistically differently compared to those at minima.

Knowing whether the flux anomalies depend upon the image parity for observed lenses would be a major step forward in identifying flux anomalies either with substructure, or with propagation effects. Namely, if the observed flux anomalies depend upon the image parity and its magnification we can set limits on the influence of the ISM (see Kochanek & Dalal [2003]).

To test whether saddle point images are also suppressed in our simulation we proceed as follows. Following Kochanek & Dalal [2003] we define $C (\leq \ln(\mu_{\text{obs}}/\mu_{\text{mod}}))$ to be the cumulative probability distribution of flux residuals, where $\mu_{\text{obs}}$ is the “observed” magnification of a simulated image and $\mu_{\text{mod}}$ is the magnification predicted by the best-fitting smooth SIE+SH model (as described above). In Fig 4, we plot $C (\leq \ln(\mu_{\text{obs}}/\mu_{\text{mod}}))$ for the systems with “observed” total unsigned magnifications $\mu_{\text{tot}} > 20$ for the simulated elliptical halo (see also Fig. 3a). We chose systems that have highly magnified images, because they are affected by substructure at most. We repeat the whole procedure with the same halo, but smoothed on a scale of $\sigma_G \sim 5 \text{ kpc}$; in this case essentially no difference is seen among images of different parity, as expected.

For the original halo the cumulative distribution is much broader for the brighter minimum and saddle point. This is in accordance with the conclusions from...
Mao & Schneider (1998); the higher the magnification, the more is the image flux affected by substructure. Among the two most magnified images, the saddle point is on average more demagnified compared to the brighter minimum. We have examined two other orthogonal projections of the mass density of the same halo and we find that qualitatively the results do not change with projection.

The effect, however, is not as pronounced as in Kochanek & Dalal (2003). The reason is two-fold. First, our simulations have a resolution of $\sim 10^8 M_\odot$ and structure at this scale and below is suppressed when using Delaunay tessellation. Kochanek & Dalal (2003), however, used SIS clumps with masses of $10^6 M_\odot$. Since these are more numerous, they can enhance the effect. Further, fitting SIE to the global mass profile is not fully justified. The mass profile is known only for a handful of observed lenses. Whereas the lens galaxies in MG1654+134, MG2016+112, 0047−281, and B1933+503 have a nearly isothermal profile, the one of PG1115+080 seems to be steeper (see Kochanek 1995; Cohn et al. 2001; Treu & Koopmans 2002a,b; Koopmans & Treu 2003.

Besides the absence of substructure on scales $\lesssim 10^7 M_\odot$, our synthetic systems and their modelled quantities closely resemble the properties of realistic lenses and the way in which these are modelled.

In the analysis of Kochanek & Dalal (2003) the synthetic images were generated using an SIE macromodel with SIS substructure. This simplifies the model fitting and explains why they get a transition of cumulative distribution exactly at $\ln(\mu_{\text{obs}}/\mu_{\text{mod}}) \sim 0$.

In addition, we have looked at the cumulative flux mismatch distribution in the bootstrapped maps, to investigate the significance of our results. In the bootstrapped images we confirm the broader distribution for bright minimum and saddle point images. In Fig 9a we plot the difference of the cumulative distribution of flux residuals $\Delta C(<\ln(\mu_{\text{obs}}/\mu_{\text{mod}}))$ between the brightest saddle point and the brightest minimum images for the original halo (solid line with dots), bootstrapped images (solid lines) and halo when additionally smoothed on a scale of $\sigma_G \sim 5$ kpc (dashed line). Positive values of $\Delta C$ thus denote the saddle point demagnification. In all bootstrapped images $\Delta C$ is positive, except for few points (corresponding to $C(<\ln(\mu_{\text{obs}}/\mu_{\text{mod}})) \sim 1$). This confirms that the effect of the saddle point demagnification is present and comparable to the original halo, whereas in the smoothed halo this effect is not seen. We conclude that substructure on mass scales $\gtrsim 10^8 M_\odot$ significantly contributes to the saddle point demagnification; possibly going a long way in explaining the observed saddle point demagnification.

5. Conclusion

In this work we have studied strong gravitational lensing properties of N-body simulated galaxies. In particular, we concentrated on the influence of substructure on flux ratios on highly magnified images. Such an analysis is crucial in order to fully understand the lensing signal that we observe in realistic lenses and to disentangle the influence on lensed-image fluxes due to propagation effects in the ISM and mass substructure.

We have examined two strong lensing signatures of substructure, i.e. the broken cusp relation observed in images that show a typical cusp configuration and the saddle point suppression. The saddle point suppression has been previously studied using semi-analytic prescrip-
tions of substructure (e.g., Schecter & Wambsganss 2002; Kochanek & Dalal 2003). The effect of substructure on the cusp relation, however, has up to now not been studied in detail.

In order to determine the magnitude of both effects we use N-body simulated galaxies. The difference compared to the works of Schecter & Wambsganss (2002) and Kochanek & Dalal (2003) is that we are using a representation of substructure that is as realistic as possible and do not make any assumptions on the mass function and abundance of sub-halos. The drawback, however, is the resolution of the simulations. We are therefore not able to extrapolate the analysis to masses \(< 10^8 M_\odot\). Still, the signatures of both effects are clearly present, and in the future we plan to use higher-resolution N-body plus gas-dynamical simulations to explore their effects in greater detail.

The main question when dealing with N-body simulations is how much are the magnification factors, that we use for synthetic image systems, affected by noise which can mimic substructure of \(< 10^8 M_\odot\). We show that the average relative noise in the surface mass density \(\sigma(\kappa)/\kappa\) lies below the \(\sim 5\%\) level for \(\kappa \geq 0.25\). Second, in Sect. 3.4 we show that the results for \(R_{\text{cusp}}\) are not significantly affected by the noise, and are dominated by physical substructure. The signal is dominated by several resolved mass clumps, which in projection lie close to the Einstein radius. Similarly, in the case of the phenomenon of suppressed saddle points, the bootstrap analysis shows that the signal also here is not dominated by noise.

The behaviour of \(R_{\text{cusp}}\) for sources close to a cusp is a very promising tool to detect substructure. Its main advantage is that it makes definite, model-independent predictions for the image magnifications. These predictions can only be broken in the presence of structure in the potential on scales smaller than the image separation. In Fig. 5 we show that the results for \(R_{\text{cusp}}\) are not significantly affected by the noise, and are dominated by physical substructure. The signal is dominated by several resolved mass clumps, which in projection lie close to the Einstein radius. Similarly, in the case of the phenomenon of suppressed saddle points, the bootstrap analysis shows that the signal also here is not dominated by noise.

The behaviour of \(R_{\text{cusp}}\) for sources close to a cusp is a very promising tool to detect substructure. Its main advantage is that it makes definite, model-independent predictions for the image magnifications. These predictions can only be broken in the presence of structure in the potential on scales smaller than the image separation. In Fig. 5 we show that the results for \(R_{\text{cusp}}\) are not significantly affected by the noise, and are dominated by physical substructure. The signal is dominated by several resolved mass clumps, which in projection lie close to the Einstein radius. Similarly, in the case of the phenomenon of suppressed saddle points, the bootstrap analysis shows that the signal also here is not dominated by noise.

Finally, our analysis shows that the two brighter images are more affected by substructure than the two fainter ones. In addition, we confirm that the brightest saddle point image in N-body simulated systems has a higher probability to be demagnified, in accordance with predictions from microlensing and from semi-analytic work by Kochanek & Dalal (2003). It is therefore necessary that all mass scales are properly accounted for, in order to compare observations with theoretical predictions in detail.

For future work, we plan to look for jet curvature as seen from N-body simulations. At present, there is only one case of a curved jet observed that is likely the cause of gravitational lensing (Metcalfe 2002). It will be interesting, however, to investigate the probability of more of these occurrences. We plan to investigate the signal one expects on average for multiple-imaged jets; this signal is also less affected by noise in the simulation and low-mass substructure.

In summary, gravitational lensing remains a very powerful tool for testing the existence of CDM substructure. N-body simulated galaxies do seem to produce the same effects as seen in observed lens systems. In addition, systematics on flux anomalies (scatter broadening, scintillation, microlensing) can be efficiently ruled out by multi-frequency and higher-frequency observations of lenses. Furthermore, the statistical analysis of large samples of lenses can directly probe the properties of CDM substructure in galaxies to a redshift of \(z \sim 1\). This provides a unique tool to measure the evolution of these structures with cosmic time, as predicted in the hierarchical structure formation scenario.

Acknowledgements. We would like to thank to our referee for his constructive comments, Oleg Gnedin, Chris Kochanek and Oliver Czoske for many useful discussions that helped improve the paper, Vincent Eke for the initial conditions of the simulations, and Richard Porcas for careful reading of the manuscript. This work was supported by the International Max Planck Research School for Radio and Infrared Astronomy at the University of Bonn, by the Bonn International Graduate School, by the Deutsche Forschungsgemeinschaft under the project SCHN 342/3–1, and by the NASA ATP program under grant NAG 5-10827.

References
Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003a, ApJ, 591, 499
—. 2003b, ApJ, 597, 21
Barber, B. C., Dobkin, D. P., & Huhdanpaa, H. 1996, ACM Transactions on Mathematical Software, 22, 469
Barkana, R. 1998, ApJ, 502, 531
Blandford, R. & Narayan, R. 1986, ApJ, 310, 568
Blandford, R. D. 1990, QJRAS, 31, 305
