Gravitational arcs as a perturbation of the perfect ring

C. Alard*

Institut d’Astrophysique de Paris, 98bis boulevard Arago, 75014 Paris, France

Accepted 2007 August 29. Received 2007 August 29; in original form 2007 July 5

ABSTRACT

The image of a point situated at the centre of a circularly symmetric potential is a perfect circle. The perturbative effect of non-symmetrical potential terms is to displace and break the perfect circle. These two effects, displacement and breaking, are directly related to the Taylor expansion of the perturbation at first order on the circle. The numerical accuracy of this perturbative approach is tested in the case of an elliptical potential with a core radius. The contour of the images and the caustic lines are well reproduced by the perturbative approach. These results suggest that the modelling of arcs, and in particular that of tangential arcs, may be simplified by using a general perturbative representation for points located on the circle. This linear perturbative approach is accurate when the gradient of the circular potential is almost linear; this constraint is satisfied when the potential is nearly isothermal.

Key words: gravitational lensing – methods: analytical.

1 INTRODUCTION

Since the discovery of gravitational arcs in clusters of galaxies by Lynds & Petrosian (1986) and subsequently by Soucail et al. (1987), the observation and study of arcs have developed considerably, and are now becoming an essential tool in astrophysics. Arcs provide a wealth of information about the mass distribution in clusters of galaxies (see for instance: Comerford et al. 2006; Broadhurst et al. 2005). However, the mass distribution of the astrophysical lenses is complex and involves a large parameter space which is difficult to explore. Thus, the derivation of a simplified perturbative theory that is able to reproduce the general features of gravitational arcs is an interesting tool to help one understand complex gravitational lenses.

2 BASIC IDEAS

Let us assume that the projected density of the lens $\Sigma$ is circularly symmetric and centred at the origin. Let us also assume that the lens is dense enough to reach critical density at a given radius $R_E$. Under such hypotheses, the image by the lens of a point source placed at the origin will be a perfect ring. We are now interested in small perturbations of this perfect ring. There are two types of perturbations to a point source perfectly aligned in a circularly symmetric potential: first, the source may not be perfectly at the centre, and secondly the potential may not be perfectly circular. Let us be more specific: in polar coordinates, the lens equation can be written as

$$ r_s = \left( r - \frac{\partial \phi}{\partial r} \right) u_r - \left( \frac{1}{r} \frac{\partial \phi}{\partial \theta} \right) u_\theta, $$

(1)

where $r_s$ is the source position, and $r, u_r$ and $u_\theta$ are the radial distance, radial direction and orthoradial direction in polar coordinates. In the unperturbed case, the equation reads

$$ r - \frac{d\phi_0}{dr} = 0, $$

(2)

where $\phi_0$ is a function of $r$ only. Let us now perturb this equation by introducing a small displacement of the source from the origin $r_s$, and a non-circular perturbation to the potential, $\psi$. Note that the perturbation on $r_s$ and the perturbation on $\phi$ are assumed to be of the same order. The perturbation may be described by the following formula:

$$ \begin{cases} r_s = \epsilon r_s, \\ \phi = \phi_0 + \epsilon \psi \end{cases} $$

(3)

Here $\epsilon$ is a small number: $\epsilon \ll 1$. Given a position $(r, \theta)$ for the source, the image positions $(r, \theta)$ can be obtained by solving equation (1). However, solving equation (1) directly may prove impossible in the general case. It is easier to find a perturbative solution by inserting equation (3) into equation (1). Assuming that $\epsilon$ is small, the perturbation will introduce a deviation from the perfect circle that will be of order $\epsilon$. An interesting feature of the unperturbed solution is that the image of a single point at the origin is a full circle, which covers the entire range of $\theta$. Thus whatever the position $\theta$ of the perturbed solution, there is always a point at the same $\theta$ in the unperturbed solution. However, the point in the unperturbed solution will be located at a slightly different radius $r$. The response to the perturbation on $r$ may be written as $r = R_E + \epsilon dr$. For convenience, it is always possible to re-scale the coordinate system, so that the Einstein radius is exactly equal to unity. In this case, the perturbation on $r$ is

$$ r = 1 + \epsilon dr $$

(4)

To summarize, we must solve equation (1) perturbatively, by expanding around the unperturbed solution (1, $\theta$) for small values of $\epsilon$. Note that this requires us to expand the potential at $r = 1$. Using

*E-mail: alard@iap.fr

© 2007 The Author. Journal compilation © 2007 RAS
equation (3), the Taylor expansion of $\phi$ may be written as

$$\phi = \phi_0 + \epsilon \psi = \sum_{n=0}^{\infty} [C_n + \epsilon f_n(\theta)] (r - 1)^n$$

where we define

$$C_n = \frac{1}{n!} \left[ \frac{d^n \phi_0}{dt^n} \right] \bigg|_{t=1}$$

and

$$f_n(\theta) = \frac{1}{n!} \left[ \frac{d^n \psi}{dr^n} \right] \bigg|_{r=1}$$

It is now possible to expand each side of equation (1) in a series of $\epsilon$, in the vicinity of the unperturbed solution. By inserting equations (4) and (5) into equation (1), the response $dr$ to the perturbation defined in equation (3) can be estimated to the first order in $\epsilon$.

$$dr = (1 + \epsilon dr - C_1 - \epsilon 2C_2 dr - \epsilon f_1)u_\theta - \epsilon \frac{d\phi_0}{d\theta}u_\theta$$

Note that equation (2) implies that $C_1 = 1$; consequently

$$r_\theta = (\kappa_2 dr - f_1)u_\theta - \frac{d\phi_0}{d\theta}u_\theta$$

with $\kappa_2 = 1 - 2C_2 = 2 - 2\kappa$. Here $\kappa$ represents the lensing convergence of the unperturbed symmetrical mass distribution.

3 RECONSTRUCTION OF IMAGES

3.1 Circular source contours

Let us consider a circular contour on a source with centre $r_0$, $0$. The equation for this contour is:

$$(r - r_0)^2 = R_0^2$$

Note that effect of the translation by the vector $r_0 = (x_0, y_0)$ can be taken into account by re-defining $f_0$, and $f_1$ in equation (12):

$$r = (\kappa_2 dr - f_1)u_\theta - \frac{d\phi_0}{d\theta}u_\theta$$

Solving equation (11), the following two solutions for $dr$ can be obtained:

$$dr = \frac{1}{\kappa_2} \sqrt{f_1^2 - R_0^2}$$

In the above equation the condition for image formation is $\Delta = R_0^2 - (d\phi_0/d\theta)^2 > 0$. The mean position of the two contour lines is $f_1/\kappa_2$, and the image width along the radial direction is $(2\sqrt{\Delta})/\kappa_2$. Note that for small distortions the function $f_1$ is directly related to the potential isophotes (see equations 17 and 20) and $f_0$ is related to the derivatives of the isophotes functional. The contributions to $f_1$ from elliptical distortions will be $\propto \cos (2\theta + \psi)$, and to $d\phi_0/d\theta$, $\propto \sin (2\theta + \psi)$. Higher order distortions, such as $\propto \cos(n\theta + \psi_0)$, where $n > 2$, should be due to substructures (Meneghetti, Bartleman & Mascodini 2003). The effect of these distortions is to displace the mean image position ($f_1$), change the image width, or even create sub-images ($d\phi_0/d\theta$).

3.2 Elliptical source contours

Using equation (10), the source Cartesian coordinates can be written as

$$\begin{cases}
x_i = (\kappa_2 dr - f_1) \cos \theta + \frac{d\phi_0}{d\theta} \sin \theta \\
y_i = (\kappa_2 dr - f_1) \sin \theta - \frac{d\phi_0}{d\theta} \cos \theta
\end{cases}$$

The equation for an elliptical contour aligned with its main axis aligned with the coordinate system is

$$(1 - \eta) x_i^2 + (1 + \eta) y_i^2 = R_0^2$$

Using equations (13) and (14), one obtains

$$dr = C_{\frac{1}{s}} \left[ f_1 + \frac{s}{r} \frac{d\phi_0}{d\theta} \pm \sqrt{r_0^2 - (1 - \eta^2) \left( \frac{d\phi_0}{d\theta} \right)^2} \right]$$

3.3 Conditions for the validity of the approximation

The perturbative approach developed in this Letter investigates small deviations from a circle of radius unity. The gradient of the potential is linearized near $r = 1$, and most of the errors are due to deviations from linearity. In particular, the gradient of the potential $\phi_0(r)$ must be close to linear in the range of image formation. Using equations (5) and (6) one can easily derive the condition for a nearly linear potential derivative:

$$\frac{d\phi_0}{dr} = 1 + 2C_2 dr + 3C_3 dr^2$$

$$3C_3 dr^2 < 1$$

The perturbed potential has non-circular isophotes near $r = 1$. Assuming that these isophotes are auto-similar they may be described by the equation

$$r = qf(\theta) = q[1 + \epsilon g(\theta)]$$

with $q$ a constant. At first order in $\epsilon$, equation (17) may be re-written: $r[1 - \epsilon g(\theta)] = q$.

The functional $\phi_0(q)$ with $q$ defined in equation (18) represents the general expression of the perturbed potential. Thus, $\phi(r, \theta) = \phi_0(q)$ and to first order in $\epsilon$, the gradient of the potential reads

$$\frac{d\phi_0}{dr} = \phi_0 - (g\phi_0 + \phi_0' q) \epsilon$$

$$\frac{d\phi_0}{dr} = -\epsilon \phi_0'(q)$$

Note that if the derivative of the unperturbed potential is linear in $r$, equation (19) implies that the gradient of the perturbed potential is also linear. Thus the condition on the linearity of the potential gradient is reduced to a condition on the linearity of the background potential derivative, which is also the essential condition for the validity of the perturbative approach (equation 16).

3.3.1 Errors on arcs

Arcs form near the critical lines in the lens plane. The equations of the critical lines are given by equation (30). From equation (19) we
infer that \( f_0 \) and \( f_1 \) are

\[
\begin{align*}
\frac{d\phi}{d\theta} &= \frac{dg}{d\theta} \\
f_1 &= -(1 + 2C_2)g.
\end{align*}
\]  
(20)

Thus, at critical radius,

\[
dr = \frac{1}{\kappa_2} \left[ -(1 + 2C_2)g + \frac{d^2g}{d\theta^2} \right].
\]  
(21)

Note that, in general, for potentials close enough to the isothermal case \( C_2 \) is not large and will not have a large influence on the estimation on the critical radius or the relevant errors. For simplicity we will neglect \( C_2 \) in the continuation. Because the potential is a functional of \( q = r(1 - \epsilon g) \), the effective parameter \( dr_c \) is

\[
dr_c \simeq -\frac{d^2g}{d\theta^2} - 2g.
\]  
(22)

Note that for some particular \( g(\theta) \) equation (22) implies that \( dr_c \) may be zero. In general, however, it is not; in the case of an elliptical potential, \( g \simeq \eta/2 \cos 2\theta \), for instance, which gives \( dr_c \simeq \eta \cos 2\theta \). Thus the deviation from the circle, \( dr_c \), is maximum at \( \theta = 0, \pi, \) and the amplitude of the deviation is \( \eta \). Considering equation (16), the maximum error due to the non-linearity of the gradient is

\[
D = 3C_1\eta^2.
\]  
(23)

In general, potentials may be represented locally by power laws or, more generally, power-laws with varying exponents. Near \( r = 1 \), one may write the potential as

\[
\phi = \frac{1}{\alpha} r^{\alpha(1 + \beta(s - 1))}.
\]  
(24)

Expanding equation (24), near the isothermal case (\( \alpha = 1, \beta = 0 \)), one obtains \( C_3 \simeq -1/6 (\alpha - 1) + \beta/2 \) and, finally, the deviation from linearity:

\[
D \simeq \left( \frac{1 - \alpha}{2} + \frac{3\beta}{2} \right) \eta^2.
\]  
(25)

The perturbative approach requires \( D \ll 1 \), which is always true in the isothermal case, but \( D \) may be of the order of \( \eta^2 \) in some cases, for instance when the critical radius is close to the core radius of the potential.

3.3.2 Error due to the impact parameter

Let us assume that the source is on the \( x \)-axis, and consider the following situation: the impact parameter is large with respect to the asymmetry of the potential (\( \eta \ll x_0 \)). In this case the mean position of the image is given by

\[
dr \simeq \frac{1}{\kappa_2} x_0 \cos \theta \simeq x_0 \cos \theta.
\]  
(26)

The maximum value of \( dr \) is \( x_0 \), and with the help of equations (16) and (25) we find that an upper bound for the error in this case is close to

\[
D \simeq \left( \frac{1 - \alpha}{2} + \frac{3\beta}{2} \right) x_0^2.
\]  
(27)

3.4 Numerical testing

The numerical investigation will be conducted with the NFW profile (Navarro, Frenk & White 1997). The lensing potential for the NFW profile has been studied by Meneghetti et al. (2003) and Bartelmann (1996); for a potential with elliptical isophotes we have

\[
\begin{align*}
\phi &\propto \frac{1}{z} \log^2 \left( \frac{1}{z} \right) - 2\arctanh^2 \left( \sqrt{\frac{1-\epsilon}{1+\epsilon}} \right) \\
x &= u_0 \sqrt{1 - \eta \cos 2\theta}.
\end{align*}
\]  
(28)

The potential is normalized so that at \( r = 1 \), which is the critical radius, \( (d\phi_0/dr)_{\theta=\pi} = 1 \). Note that in equation (28), a parameter \( u_0 \) has been introduced; this is a scaling parameter that control the relative position of the critical radius with respect to the NFW scale parameter. A small \( u_0 \) means that the critical radius is near the centre of the profile. A comparison of the prediction of the perturbative approach on the arc position to a direct numerical integration of the full equation for the NFW profile will give a good estimate of the errors. The previous section shows that for the arcs the error on the perturbative estimation is maximum at \( \theta = 0 \). Using equations (20) and (31) we estimate that the corresponding impact parameters of a source near the caustic must be approximately \( x_c \simeq 2\eta, y_c = 0 \). We have now to solve numerically the lensing equation (equation 1) with impact parameters \( x_c \simeq 2\eta, y_c = 0 \) for the NFW profile and by the perturbative approach. The results will be compared for different values of the ellipticity \( \eta \) and of the scaling parameter \( u_0 \). The results are presented in Fig. 1. Note that errors are only of a few per cent for \( u_0 = 1 \) or \( u_0 = 1.5 \) even for large ellipticities. The situation is quite different for smaller values of \( u_0 \) when arcs form in areas where the projected density slope is not constant and not isothermal. As predicted from equation (25), the error in the worst case scenario is of the order of \( \eta^2 \).

3.5 Comparison with ray tracing

A more detailed comparison of the perturbative approach with numerical integration of the NFW lens equation can be performed by using the ray-tracing technique with a circular source. Using the
be corrected using an iterative approach. As good for the inner image, although as mentioned in Section 3.3, this may be due to the outer contour of the ray-tracing solution for the tangential image. Due to the red line on the ray-tracing solution. The red contour is close to the image edges. The result is visible in Fig. 2. Note that the large tangential arc is well reproduced by the approximation, while there is some mismatch for the inner image. The problem of the inner image can be corrected using an iterative approach: the perturbative method gives a first guess of the image position, at this initial position one may carry a local Taylor expansion of the potential in order to find a better solution, and the procedure may be iterated again.

3.6 Inverse modelling

The perturbative method has the advantage of offering a linear non-parametric (not model-dependent) approach of strong gravitational lensing. It is clear that in this approach the inversion problem is greatly simplified and requires only the reconstruction of the fields \( f_0(\theta) \) and \( f_1(\theta) \). Let us define the arc system as a set of contours, with one contour per image. For each contour, a radial line of direction \( \theta \) intersects the contour at two points, \( r_1 \) and \( r_2 \) (provided that \( \theta \) is in a suitable range). It is simple to relate these two functionals to the fields, \( f_0 \) and \( f_0 \), of the perturbative approach, in the case of a source with circular contour (equation 12):

\[
\begin{align*}
\bar{r}_1 &= \frac{\kappa}{2}(r_1 + r_2) + C \\
\frac{df_0}{d\theta} &= \sqrt{R_0^2 - \frac{\kappa^2}{4}(r_2 - r_1)^2},
\end{align*}
\]

where \( C \) is a constant term. Note that \( \kappa_2 \) is unknown; actually there is a fundamental degeneracy in \( \kappa_2 \), and only \( f_1/\kappa_2, f_0/\kappa_2, \) and \( R_0/\kappa_2 \) may be determined. This is the usual mass-sheet degeneracy. By using inner or radial images in the iterative approach mentioned in Section 3.3, in some cases it might be possible to break the degeneracy of \( \kappa_2 \). Note that the accuracy of the reconstruction will depend on the linearity of the background field \( df_0/dr \). One can always choose a background field that is well-behaved in the perturbative approach and perform a reconstruction using equation (29). There is no particular requirement on the background field at this level, this degeneracy was already noticed by Bartelmann & Meneghetti (2004). It is also clear that instead of circular contours, one could have considered an elliptical one, and equation (25) may have been used. In such cases, it is just a matter of where to put the complexity, either in the source or the lens, and obviously the judging criteria should be that, as a whole, the complexity is minimal.

4 CAUSTICS IN THE PERTURBATIVE APPROACH

Caustics are singularities, which are defined by the simple property that the determinant of the Jacobian matrix \( J \) is zero on the caustic lines:

\[
J = \frac{\partial x}{\partial r} \frac{\partial y}{\partial \theta} - \frac{\partial x}{\partial \theta} \frac{\partial y}{\partial r} = 0.
\]

The calculation of the Jacobian is straightforward from equation (13); it follows that

\[
dr = \frac{1}{\kappa_2} \left[ f_1 + \frac{d^2 f_0}{d\theta^2} \right] \cdot (30).
\]

Note that in this case a shift by a vector \( r_0 \) does not have to be introduced, thus \( f_0 = f_0 \) and \( f_1 = f_1 \). Equation (30) defines the critical lines. The caustics in the source plane can be obtained by inserting equation (30) in equation (13):

\[
\begin{align*}
x_s &= \frac{d f_0}{d\theta} \cos \theta + \frac{df_0}{d\theta} \sin \theta \\
y_s &= \frac{d f_0}{d\theta} \sin \theta - \frac{df_0}{d\theta} \cos \theta.
\end{align*}
\]

Note that in equation (31) the caustic line depends only on \( f_0 \); this equation is similar to equation (3.8) from Blandford & Kovner (1988). However, Blandford & Kovner use a very different method for the reconstruction of the images; they start from the caustics equation and try to derive the images’ position using a complicated geometric method. Actually, Blandford & Kovner (1988) derived only part of the perturbative theory; their work misses equations (8), (12) and (25), which are essential for image reconstruction and inversion. The functional \( f_0 \) is directly related to the multipole expansion of the perturbative potential on the circle. The multipole expansion also has the advantage of directly relating \( f_0 \) to the density by the means of the coefficients (\( \alpha_n \)) of the multipole expansion at \( r = 1 \):

\[
\psi(r = 1) = \sum_n \alpha_n \cos n\theta + b_n \sin n\theta.
\]

Turning now to a numerical application, by estimating \( f_0 \) using equation (28) and evaluating equation (31) a parametric equation of the caustic line is obtained, and the result is presented in Fig. 3. The perturbative calculation of the caustic curve is accurate, which suggests that this approximation may be used to derive general results on caustics. A simple result is that

\[
r^2 = \left( \frac{df_0}{d\theta} \right)^2 + \left( \frac{d^2 f_0}{d\theta^2} \right)^2.
\]
Figure 3. Caustics of the NFW potential for $u_0 = 0.5$; the dotted line represents the perturbative solution, the solid line is the numerical solution of the system of equations without approximations. Note that, despite the non-linearity of the background potential in this example, the mean distance between the two solutions is only $0.03R_E$.

By combining equations (32) and (33) and integrating in the interval $0 < \theta < 2\pi$, it is possible to obtain the average value of $r_s^2$, which gives an estimate of the caustic size:

$$\langle r_s^2 \rangle = \sum_n (a_n^2 + b_n^2) \frac{(n^2 + n^4)}{2}.$$  \hspace{1cm} (34)

This calculation demonstrates that the size of the caustic is strongly influenced by higher order multipole terms; equation (34) shows the caustic size varies approximately as $n^2$. Considering that substructures contribute significantly to higher order multipoles (Meneghetti et al. 2003), it is clear that, due to the $n^2$ factor, the size of the caustic will be strongly influenced by substructures. This result is confirmed by the numerical analysis of Meneghetti et al. (2007), who evaluate the effect of the different classes of deviations from the circular symmetry of the potential observed in astronomy asymmetries on the formation of gravitational arcs.

ACKNOWLEDGMENTS

The author would like to thank J. P. Beaulieu, S. Colombi, S. Prunet, K. Benabed, Y. Mellier and M. Alard for some helpful comments.

REFERENCES

Bartelmann M., 1996, A&A, 313, 697
Bartelmann M., Meneghetti M., 2004, A&A, 418, 413
Blandford R. D., Kovner I., 1988, Phys. Rev. A, 38, 4028
Broadhurst T. et al., 2005, ApJ, 621, 53
Comerford J., Meneghetti M., Bartelmann M., Schirmer M., 2006, ApJ, 642, 39
Kochanek C. S., 1991, ApJ, 373, 354
Lynds R., Petrosian V., 1986, BAAS, 18, 1014
Meneghetti M., Bartleman M., Moscardini L., 2003, MNRAS, 340, 105
Meneghetti M., Argazzi R., Pace F., Moscardini L., Dolag K., Bartelmann M., Li G., Oguri M., 2007, A&A, 461, 25
Navarro J., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Soucail G., Fort B., Mellier Y., Picat J. P., 1987, A&A, 172, L14

This paper has been typeset from a TeX/LaTeX file prepared by the author.