Mon. Not. R. astr. Soc. (1988) **233**, 265–284

The focusing equation, caustics and the condition for multiple imaging by thick gravitational lenses

T. Padmanabhan★ *Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*

Kandaswamy Subramanian *Astronomy Centre, Physics Building, University of Sussex, Falmer, Brighton BN1 9QH*

Accepted 1987 December 17. Received 1987 November 30; in original form 1987 September 10

Summary. The conditions for the production of multiple images by an arbitrary (thick or thin) gravitational lens are studied. We show that the necessary and sufficient condition for the production of multiple images by a lens is the following: *The lens should produce a point conjugate to the observer, along some null geodesic, at an affine distance smaller than that of the source.* It is shown that previous results on multiple imaging by thin lenses can be obtained as a special case. We also show that a thick lens cannot be more efficient than a suitably designed thin lens for the production of multiple images.

1 Introduction

The light rays which reach us from distant sources (like quasars) would have been influenced by the gravitational field of all the matter distributed along the path. We ask the question: what general condition should the matter distribution satisfy if it has to act as a gravitational lens and produce multiple images of distant sources?

The answer to the above question is known in some special cases. The simplest situation corresponds to the case where a point mass acts as a lens; in this case, multiple images are always produced. This is due to the singularity in the density of the point-mass lens. Smooth lenses, however, need not be capable of producing multiple images in general. The conditions under which a smooth (and bounded) lens produces multiple images has been analysed only for the case when the lens is 'thin'. (A lens is considered 'thin' if the extent of the lens along the line-of-sight and the impact parameter distances of light rays are small compared to the distances to the observer and sources.) In this case it can be shown that (Subramanian & Cowling 1986): (i) Any

★On leave of absence from: Astrophysics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay, India 400 005.

thin lens with surface (mass) density Σ larger than a critical density

$$\Sigma_c = \frac{c^2}{4\pi G}\left(\frac{D_s}{D_l D_{ls}}\right) \tag{1}$$

(where D_s, D_l and D_{ls} are the angular diameter distances between observer–source, observer–lens and lens–source respectively) will produce multiple images. This is a sufficient condition. (ii) As one relaxes the symmetry of the lens, the Σ necessary for multiple imaging decreases. (iii) In general, there is no lower bound on the value of Σ necessary for multiple imaging.

However, a *general* mathematical criterion which is necessary and sufficient for multiple imaging by an arbitrary lens, is lacking. The lack of such a criterion prevents us from answering several important questions regarding thick gravitational lenses. Consider for example, the following questions: (i) Will multiple images be produced in an inhomogeneous universe, in which the inhomogeneities occur at scales comparable to cosmological length scales? (ii) Will such a thick lens be more efficient in producing multiple images compared to a thin lens? There are no clear-cut answers to such questions today. In general, thick gravitational lenses have proved to be tough nuts to crack, mainly because the usual approach to working out image configurations for thin lenses, the bending angle approach, proves inadequate for thick lenses.

In this paper we provide such a general criterion. We show that *if a bounded, smooth mass distribution can produce a conjugate point with respect to the observer along any ray, then such a lens will produce multiple images of sufficiently far away sources.* This condition is both necessary and sufficient. Thus the question of multiple imaging can be reduced to that of checking whether points conjugate to the observer exists. In principle, this can be achieved by solving the so-called focusing equation – if necessary – numerically. In addition, this provides an approach which is capable of handling several general properties of lenses, thick or thin.

The rest of the paper is organized as follows: In Section 2, we discuss some mathematical background and establish our general criterion for multiple imaging. Section 2.1 discusses in a pedagogical fashion the focusing equation and its solutions in a Robertson–Walker universe. Section 2.2 proves the result that the existence of conjugate points implies the existence of multiple images.

In Section 3 we discuss several applications of this result. In Section 3.1, we re-derive the results for the thin lens from the present approach. Section 3.2 applies this method to study multiple thin lenses. Section 3.3 discusses a thick lens with variable density and proves some general results. In particular we show that, in cases of astrophysical interest, whenever a thick lens can produce multiple images, a suitably placed thin lens can do it with lower surface density.

2 Mathematical background and the main result

2.1 FOCUSING EQUATION AND CONJUGATE POINTS

Consider a narrow beam of photons with a cross-sectional area $A \simeq \pi l^2$, travelling in a gravitational potential $\phi(\mathbf{x}, t)$. The difference in the gravitational acceleration across the cross-section of the beam will be $\sim\tilde{\nabla}(\tilde{\nabla}\phi)l \sim (\nabla^2\phi)l \sim 4\pi G\varrho_N l$ where ϱ_N is the Newtonian mass density producing ϕ. This differential acceleration can distort the cross-sectional area and we expect

$$\frac{d^2}{dt^2} l \simeq -4\pi G\varrho_N l \tag{2}$$

or

$$\frac{d^2\sqrt{A}}{dt^2} \simeq -4\pi G\varrho_N\sqrt{A}. \tag{3}$$

Though the argument given above is completely incorrect, rigorous, general relativistic analysis of null rays in a space-time does lead to an equation similar to (3)! The correct equation is (Sachs 1961; Kantowski 1969)

$$\frac{d^2\sqrt{A}}{d\lambda^2}=-\left(\frac{4\pi G}{c^2}\varrho_{\rm GR}+|\sigma|^2\right)\sqrt{A}. \tag{4}$$

In (4), A is the *proper* cross-sectional area of a family of null rays, infinitesimally separated from a fiducial null ray; $\varrho_{\rm GR}$ is related to the matter density and is given by

$$\varrho_{\rm GR}=T_{ab}k^a k^b \tag{5}$$

where T_{ab} is the energy momentum tensor for matter and k^a is the tangent vector to the fiducial null geodesic the quantity; σ represents the 'shear' of the family of null rays; λ is the affine parameter along the fiducial geodesic. In order to solve (4) we need to know the function $\sigma(\lambda)$. This is determined by the equation

$$\frac{d\sigma}{d\lambda}+\left(\frac{1}{A}\frac{dA}{d\lambda}\right)\sigma=C_{aibj}k^a k^b\bar{t}^i\bar{t}^j \tag{6}$$

where C_{ijkl} is the Weyl tensor for the given space-time. The complex vector t^i is chosen so that $k^a t_a=0$ and $\bar{t}^a t_a=-1$ and is parallel transported along the geodesic.

In any reasonable space-time $\varrho_{\rm GR}$ will be positive. Thus the right-hand side of (3) will (always) be negative tending to reduce the area A, and 'focus' the beam to a line or a point where A will vanish. (Whether such a focusing will actually succeed or not, of course, depends on the particular situation; but the *tendency* of matter and shear is to focus the beam.) Consider a beam of light emerging from a point source S which is made to converge (or focus) to a zero area at some point P. The area A vanishes at both S and P; $A'(dA/d\lambda)$ however, is positive at S (where the beam diverges) and is negative at P (where the beam converges). A pair of points (like P, S) along a null geodesic are called 'conjugate points'. More precisely a point P will be called conjugate to S if (i) P and S lie along some null geodesic; (ii) the area A vanishes at both P and S; (iii) A' has different signs at P and S.

In the later sections of the paper, we shall be concerned repeatedly with the conditions for the existence of points conjugate to a given point P (at which A vanishes and A' is positive). In settling such questions one can often ignore $|\sigma|^2$ in (4) and deal instead with the equation

$$\frac{d^2\sqrt{A}}{d\lambda^2}=-\frac{4\pi G}{c^2}\varrho_{\rm GR}\sqrt{A} \tag{7}$$

First this can be done if we are only interested in sufficient conditions for the existence of conjugate points. For we know that if a conjugate point exists with $|\sigma|^2=0$, it will certainly exist with $|\sigma|^2\neq 0$ since $|\sigma|^2$ always aids the focusing. The shear term also vanishes for propagation of light beams with initially circular cross-section, in either flat space-time or the conformally flat Robertson–Walker (RW) universe. Hence, here also, the solution of the focusing equations (4) and (6) reduces to solving equation (7).

It is convenient, at this stage, to introduce two more variables

$$\theta(\lambda)\equiv\frac{1}{\sqrt{A}}\frac{d\sqrt{A}}{d\lambda};\qquad D=\sqrt{A}. \tag{8}$$

The quantity θ measures the 'expansion' of the beam. At the source (where $A'>0$, $A=0$) $\theta=+\infty$ and at the converging focal point (where $A'<0$, $A=0$), $\theta=-\infty$. In general, $\theta<0$ denotes a converging beam while $\theta>0$ represents a diverging beam. In terms of θ, equation (7) can be written as

$$\theta'+\theta^2=-\frac{4\pi G}{c^2}\varrho_{GR}. \tag{9}$$

We will call D the 'diameter' of the beam, even though, in general, the cross-section will not be circular.

For the sake of future use, we shall discuss the solution to the focusing equation in some special cases.

The simplest case corresponds to empty space where $\varrho_{GR}=0$. In this case D is proportional to λ or equivalently, $\theta(\lambda)$ is given by

$$\theta(\lambda)=\frac{1}{\lambda+\lambda_0}. \tag{10}$$

The constant λ_0, of course, has to be determined by the initial conditions.

The second case for which we will need the solution is that of a beam propagating in a RW universe with the metric

$$dS^2=dt^2-s^2(t)\left[\frac{dr^2}{1-kr^2}+r^2(d\tilde{\theta}^2+\sin^2\tilde{\theta}\ d\phi^2)\right]. \tag{11}$$

(We choose $c=1$ unit in the rest of the section.) Suppose a fiducial ray emitted at time t from $r=r(t)$ reaches $r=0$ at time t_0 travelling along $\tilde{\theta}=0$. Consider a bunch of neighbouring light rays travelling on a cone with vertical angle $\delta\tilde{\theta}$ around the fiducial ray. At time t the proper diameter of this beam will be

$$D'=S(t)r(t)\delta\tilde{\theta}. \tag{12}$$

Since $\delta\tilde{\theta}$ is a constant we see that the function

$$D=S(t)r(t) \tag{13}$$

must satisfy the focusing equation. This fact can also be verified directly. To do so, we need the relation between the affine parameter λ and t. The time component of the geodesic equation can be solved to give

$$\lambda=\int dt\, S(t). \tag{14}$$

It is now easy to verify that $r(t)S(t)$ actually satisfy the focusing equation. We have,

$$\frac{d}{d\lambda}\left(\frac{D}{S}\right)=\frac{1}{S}\frac{d}{dt}\left(\frac{D}{S}\right)=\frac{1}{S}\frac{dr}{dt}=\frac{\sqrt{1-kr^2}}{S^2} \tag{15}$$

or, equivalently, denoting differentiation with respect to λ by a prime,

$$SD'-DS'=\sqrt{1-kr^2}. \tag{16}$$

Differentiating again, we get

$$SD''-DS''=\frac{-2kr}{2\sqrt{1-kr^2}}\frac{1}{S}\frac{dr}{dt}=-\frac{kr}{S^2}=-\frac{kD}{S^3} \tag{17}$$

or

$$\frac{D''}{D}=\left(\frac{S''}{S}-\frac{k}{S^4}\right). \tag{18}$$

Now, in a Friedman universe dominated by dust of density ϱ,

$$\frac{S''}{S}=\frac{1}{S^2}\frac{d}{dt}\left(\frac{\dot{S}}{S}\right)=\frac{1}{S^2}\left[\frac{\ddot{S}}{S}-\frac{\dot{S}^2}{S^2}\right]=\frac{1}{S^2}\left[-\frac{4\pi G\varrho}{3}-\frac{8\pi G\varrho}{3}+\frac{k}{S^2}\right]=-4\pi G\left(\frac{\varrho}{S^2}\right)+\frac{k}{S^4} \tag{19}$$

(where we have used Einstein's field equations to eliminate $\dot{S}$, $\ddot{S}$ etc.) giving the final result

$$\frac{D''}{D}=\frac{S''}{S}-\frac{k}{S^4}=-4\pi G\left(\frac{\varrho}{S^2}\right)=-4\pi G\varrho_{\rm GR}. \tag{20}$$

Note that in such a FRW space-time, $T_{ik}=\varrho u_i u_k$ with $u^i=(1,\mathbf{0})$ and

$$k^i=\left(\frac{dt}{d\lambda},\frac{dr}{d\lambda}\right) \tag{21}$$

so that the $\varrho_{\rm GR}$ of equation (4) is actually

$$\varrho_{\rm GR}=T_{ab}k^a k^b=\varrho(u_a k^a)^2=\varrho S^{-2}. \tag{22}$$

The function $D(\lambda)$ is one particular solution to the focusing equation and satisfies the initial conditions

$$D(0)=0; \qquad \left.\frac{dD}{d\lambda}\right|_{\lambda=0}=\frac{1}{S(0)}. \tag{23}$$

In order to write down the most general solution to the focusing equation, we need an independent second solution, say, $C(\lambda)$, to the focusing equation. This solution is easy to find. We know, from the theory of differential equations, that C and D must satisfy the Wronskian condition:

$$CD'-DC'=1. \tag{24}$$

This equation can be integrated at once to give C in terms of D:

$$C(\lambda)=-D(\lambda)\int^{\lambda}\frac{d\lambda}{D^2(\lambda)}. \tag{25}$$

We can express this in a neater form by using (13) and the relations:

$$d\lambda=S\,dt=\frac{S^2\,dr}{\sqrt{1-kr^2}}. \tag{26}$$

We get

$$C(\lambda)=Sr\int\frac{dr'}{\sqrt{1-kr'^2}}\left(\frac{1}{r'^2}\right)=S\sqrt{1-kr^2}=S\sqrt{1-k\frac{D^2}{S^2}}. \tag{27}$$

This is our second solution. Note that $C(\lambda)$ satisfies the initial conditions

$$C(0)=S(0); \qquad \left.\frac{dC}{d\lambda}\right|_{\lambda=0}=0. \tag{28}$$

Also note that $C=S$ for $k=0$ while in general C, D and S are related by:

$$C^2+kD^2=S^2. \tag{29}$$

The angular diameter distance $D(\lambda)$ is defined with respect to an observer located at the origin with $\lambda=0$. It is often necessary to define a 'two-point angular distance' $D(\lambda;\lambda_0)$ which is the angular diameter distance an observer at $\lambda=\lambda_0$ will attribute to an object on the same null geodesic, at affine distance λ. This $D(\lambda, \lambda_0)$ can be easily constructed from $C(\lambda)$ and $D(\lambda)$. Since $D(\lambda, \lambda_0)\delta\theta$, by definition, is the width of a beam diverging from $\lambda=\lambda_0$, we can express $D(\lambda, \lambda_0)$ as an arbitrary superposition of C and D.

$$D(\lambda, \lambda_0)=aD(\lambda)+bC(\lambda). \tag{30}$$

To determine a and b we use the 'initial' conditions at $\lambda=\lambda_0$, namely

$$D(\lambda_0, \lambda_0)=0; \qquad \left.\frac{dD(\lambda, \lambda_0)}{d\lambda}\right|_{\lambda=\lambda_0}=\frac{1}{S(\lambda_0)}. \tag{31}$$

Combining (30) and (31) we get,

$$aD(\lambda_0)+bC(\lambda_0)=0, \qquad aD'(\lambda_0)+bC'(\lambda_0)=\frac{1}{S(\lambda_0)}. \tag{32}$$

These equations can be solved trivially since $CD'-DC'=1$. We get,

$$a=S^{-1}(\lambda_0)C(\lambda_0); \qquad b=-S^{-1}(\lambda_0)D(\lambda_0). \tag{33}$$

So that,

$$D(\lambda, \lambda_0)=\frac{1}{S(\lambda_0)}\,[C(\lambda_0)D(\lambda)-D(\lambda_0)C(\lambda)]. \tag{34}$$

By substituting the explicit forms (13), (27) it is easy to verify the physical meaning of $D(\lambda, \lambda_0)$. For example, in a $k=1$ model, using the substitution $r=\sin\chi$, we get,

$$D(\lambda, \lambda_0)=S(\lambda)[r(\lambda)\sqrt{1-r^2(\lambda)}-r(\lambda_0)\sqrt{1-r^2(\lambda)}]=S(\lambda)\sin[\chi(\lambda)-\chi(\lambda_0)]. \tag{35}$$

Clearly, $D(\lambda, \lambda_0)\delta\tilde{\theta}$ represents the proper arc length on the hypersphere $(\chi, \tilde{\theta})$, which is what one expects for properly defined angular diameter.

In the discussion above, we have tried not to introduce explicit formulae for D, C, etc. in terms of various observable cosmological parameters like Ω_0, z, and so on. In fact, we will never require such explicit forms.

There are other ways of expressing the second solution $C(\lambda)$. The reader can amuse himself by proving that,

$$C(\lambda)=(\text{constant})[S(\lambda)-\bar{a}S^2(\lambda)]$$

with

$$\bar{a}=\frac{2k}{\Omega_0}(\Omega_0-1)S_0^{-1}.$$

We will not need this result; but it is useful in deriving an explicit form for C.

2.2 CAUSTICS, CRITICAL CURVES AND MULTIPLE IMAGING

Suppose we have a smooth density distribution acting as a gravitational lens, and the density vanishes sufficiently rapidly at spatial infinity. We wish to know when such a smooth and bounded lens would be capable of producing multiple images. The case where the lens is thin (that is, confined to a very small range in redshift space and the impact parameter distances of light rays are small compared to the distance between the lens and the source or observer) has been considered in some detail by Subramanian & Cowling (1986). Here we do not necessarily want to restrict our consideration to the thin lens case. For this purpose, the focusing equation (4) turns out to provide a powerful tool.

We defined the notion of conjugate points in the previous section. Note that in *flat* space-time in the absence of the lens there are no points conjugate to the observer (O). We will assume in what follows that this is also true in the RW universe, *for the range of source positions we are interested in.** We will then show the following: A smooth and bounded lens is capable of producing multiple images if and only if it leads to the existence of a point conjugate to the observer on some null ray.

For showing the above result we will proceed in the following steps. We assume the space to be Euclidean sufficiently close to the observer. In this Euclidean region we define a plane, say α, perpendicular to the line-of-sight, to a point l on the lens, and at a distance R from the observer. We set up a Cartesian coordinate system (x, y) on this plane with the origin being the point of intersection of the light ray OL with the plane α. Consider any other light ray coming to the observer along some direction. It will intersect the plane α at say (x, y). We then have a one-to-one correspondence between the directions of light rays reaching the observer and its (x, y) coordinates. We will refer to the (x, y) space as the image space.† We also define a 'source' space B as follows: consider all null geodesics reaching the observer O. Extend them backwards in time for an affine parameter distance $\lambda=\lambda_s$, say. The set of points on all these null geodesics at $\lambda=\lambda_s$ forms a two-dimensional surface which we denote by B_s. We also introduce a coordinate system (u, v) on this surface.

We now define a map ϕ_s from the image to the source space. Take any (x, y) in α. Consider the null geodesic which reaches the observer through (x, y). Trace it backward in time for an affine parameter distance λ_s. It will then reach a unique point say (u, v) on B_s. The map ϕ_s is defined as the transformation which takes (x, y) on α to (u, v) on B_s. By the map each fixed point on α will go to a unique point on B_s. But for each fixed point on B_s the inverse transformation of ϕ_s need not lead to a unique (x, y) on α. (This is due to the possibility of seeing multiple images of a single source.)

Now consider a bundle of light rays through an infinitesimal area $dx\,dy$ about (x_0, y_0). This will intersect an infinitesimal area $du\,dv$ about (u_0, v_0), where (u_0, v_0) is the image of (x_0, y_0) under the map ϕ_s. The 'Jacobian' of the map $J(x_0, y_0)$ is then given by $du\,dv/dx\,dy$. But the magnitude of this is also $A(x_0, y_0, \lambda_s)/A(x_0, y_0, R)$, where A (the cross-sectional area of a light beam traced backwards from O through (x_0, y_0) up to an affine distance λ_s), is defined in the previous section (equation 4). The sign of J is obtained by first taking a fixed orientation for the area $dx\,dy$ about (x_0, y_0) and then asking for the orientation of $du\,dv$, under the action of ϕ_s. If $du\,dv$ and $dx\,dy$ have the same orientation, J is positive and if the orientations of $du\,dv$ and $dx\,dy$ are different J is negative. (We will return to this point later on .) Let us consider the properties of J. First consider its behaviour at infinity, by taking a light beam through (x, y) such that $(x^2+y^2)^{1/2} \rightarrow \infty$. Since the lens is bounded there is no contribution to ϱ from the lens in the focusing equation for A. The shearing effects due to the Weyl curvature term of the lens will also fall off to zero at spatial infinity. The focusing equation then reduces to that which holds without the lens in this limit. The solution of the focusing equation in this case is then as in equations (10) or (13). A increases monotonically with λ in flat space and (by assumption) does not pass through a zero until the source λ in a FRW Universe. An infinitesimal oriented area $dx\,dy$ never gets mapped to a zero area by ϕ for any λ along the light ray through 'infinity' in (x, y). So the orientation of $du\,dv$ is then the same as that of $dx\,dy$ and $J(x, y)$ is then positive for $(x^2+y^2)^{1/2} \rightarrow \infty$.

Now we are ready to prove the main result of this section. We first show that if the lens leads to

*In other words, any conjugate point which exists in RW universe in the absence of the lens occurs at distances larger than that of the source λ.

†All we need of the image space α is that there should be a one-to-one relationship between directions on the sky and points on α. For this, it is not essential to have space Euclidean near the observer but only that (on hindsight based on the results which follow) no conjugate point to O should lie between the 'plane' α and the observer.

the existence of a conjugate point to O then there will be curves on α on which $J=0$. This will then be shown to imply the existence of a point on which J is negative. The existence of a point on which J is negative will then be seen to imply from Burke's theorem [Burke (1981) as stated in Subramanian & Cowling (1986)] that, such a lens is capable of producing multiple images. We will also establish the converse statements to complete the proof.

Suppose there exists a point O^1 conjugate to O lying on a null geodesic passing through (x_c, y_c) on α and $\lambda=\lambda_c$ surface, that is we have $\theta(x_c, y_c, \lambda_c) \to -\infty$ with $\theta(x_c, y_c, 0) \to \infty$. Since the lens is assumed to be sufficiently smooth, the condition $\theta(x, y, \lambda) \to -\infty$ (or $1/\theta=0$) will be satisfied in the generic case on a continuous, non-degenerate two-dimensional surface, if it is at all satisfied. (Here the two-dimensional surface is said to be non-degenerate if it is not a curve or a point.) We will refer to such a surface as the caustic surface. So if there exists a point O^1 conjugate to O, then in the generic case, this will imply the existence of a caustic surface. It is then clear there will exist at least one surface $\lambda=\lambda_s$ which will have a non-zero intersection with the caustic surface and whose points of intersection lie on non-degenerate continuous curves C_s. We will refer to the curves C_s as critical curves at λ_s. Let us now take a point $P_c(u_s, v_s)$ on C_s at λ_s. Since P_c is on a critical curve, there will exist at least one geodesic from O to P_c on which O and P_c are conjugate points. This will go through the plane α at the point $\bar{P}_c$ given by (x_s, y_s) say. Then the magnitude of $J(x_s, y_s)=A(x_s, y_s, \lambda_s)/A(x_s, y_s, R)=0$ since a bundle of light rays through an infinitesimal area around (x_s, y_s) will be focused to a zero area at P_c. Further P_c was just an arbitrary point on C_s. So the above result holds for the whole curve C_s. Then each point P_c on C_s will go to at least one point $\bar{P}_c$ on α where $J=0$. Since C_s are a set of continuous non-degenerate curves, the collection of these points $\bar{P}_c$s will generate a set of continuous curves $\bar{C}_s$ on which $J=0$. We will refer to $\bar{C}_s$ as image critical curves.

We now discuss some properties of the curves $J=0$. Whitney (1955), showed that for 'excellent' transformations of the plane into the plane the points $J=0$ lie on continuous curves across which J changes sign. He also showed that any sufficiently smooth transformation of the plane to the plane can be approximated arbitrarily closely by an 'excellent' transformation. This makes the properties of excellent transformations generic. For a smooth lens we can apply Whitney's results to ϕ_s. Now we saw earlier that J is positive at spatial infinity. So as we cross a $J=0$ image critical curve on α from spatial infinity J will change sign. So there will be at least one point in the (x, y) plane say P_- where J is negative. The existence of $J=0$ curves are then seen to imply the existence of at least one point P_- where J is negative.

Now let us put a source at $\phi_s(P_-)$ which say has the coordinates (u_-, v_-) on B_s. Burke (1981) showed that a smooth and bounded gravitational lens always produces an odd number of images. More specifically he showed that any such lens produces one image with positive parity, in the present context with $J>0$ and other images come in pairs, one with $J>0$ and the other with $J<0$. (The argument of Burke, adapted to the present context is outlined in Appendix A.) We already have one image at P_-, for a source kept at $\phi_s(P_-)$, which has $J<0$. This implies that there has to be at least two more images, located on α, for which $J>0$. For all these images at say (x_i, y_i), $u(x_i, y_i)=u_-$ and $v(x_i, y_i)=v_-$. So there will be at least three null geodesics through (x_i, y_i) which will connect the observer to the source, if $J(P_-)<0$. We have therefore that that if $J<0$ at any point P_-, the lens can produce multiple images for a source kept at $\phi_s(P_-)$. Combining all the above arguments one then has the result that if the action of a smooth bounded lens leads to the existence of a point conjugate to O, then the lens is capable of producing multiple images.

Conversely, suppose a lens is capable of producing multiple images for a source position, S on the $\lambda=\lambda_s$ surface. Consider the points $\phi_s^{-1}(S)$ on α. From the work of Burke (1981) at least one of these points, say $\bar{P}$ will have negative parity with $J(\bar{P})<0$. But $J>0$ at spatial infinity on α. Consider a curve from $\bar{P}$ and going to infinity. J has to become zero at some point on this curve say at $\bar{\bar{P}}$. Now take $\phi_s(\bar{\bar{P}})$ at $\lambda=\lambda_s$ surface. Then $A(\bar{\bar{P}}, \lambda_s)=J(\bar{\bar{P}})A(\bar{\bar{P}}, R)=0$. So $\phi_s(\bar{\bar{P}})$ is conjugate to

O. So if the lens can produce multiple images there will exist at least one point conjugate to O. Combining this with the result of the previous paragraph establishes the main result of this section. A smooth and bounded lens is capable of producing multiple images if and only if it leads to the existence of a point conjugate to the observer on some null geodesic.

It is also useful to establish a slight extension of the above result which we often appeal to in practical applications.

A smooth and bounded lens can create multiple images of a suitably placed source at a $\lambda=\lambda_s$ surface if it can produce a point conjugate to the observer on some null ray at $\lambda=\lambda_c<\lambda_s$ (i.e. it should produce a conjugate point in 'front of' the source).

In order to prove this it suffices to show that, for any λ_s, there are conjugate points on the $\lambda=\lambda_s$ surface iff there is a conjugate point on some null ray at $\lambda=\lambda_c<\lambda_s$. We shall outline the proof of this result in Appendix B. This would then imply that the caustic surface will cut the $\lambda=\lambda_s$ surface iff the lens produces a conjugate point at $\lambda=\lambda_c<\lambda_s$ and from the argument outlined previously the lens would be capable of producing multiple images of a source kept somewhere on a $\lambda=\lambda_s$ surface.

The above extension to our main result offers a simple and direct way of checking whether a given lens can, in principle, for some source position, on a given $\lambda=\lambda_s$ surface, produce multiple images. One has to check only whether the lens can form a point conjugate to the observer on some null geodesic, at an affine distance $\lambda_c<\lambda_s$ (in front of the source plane). This is the property that we exploit in the rest of the paper to study the ability of thick gravitational lenses as well as thin lenses, to produce multiple images.

3 Applications

3.1 SINGLE THIN LENS REVISITED

The main advantage of the present formalism lies in its ability to tackle thick lenses. We address this question in Section 3.3. Before we do that, however, we re-analyse the 'sheerless' thin lens using the present formalism in this section. This serves three purposes: (i) it shows how the concept of conjugate points operates in a familiar setting; (ii) it reproduces the previous results [(i) in Section 1] in a natural, simple manner; and (iii) it clarifies certain issues of principle related to the angular diameter distance.

We begin with the focusing equation, written in the context of RW cosmology:

$$\frac{d^2g}{d\lambda^2}=-4\pi G\left(\frac{\varrho}{S^2}\right)g. \tag{36}$$

The thin lens approximation involves splitting up the density ϱ into two parts ($\varrho=\varrho_{\mathrm{bg}}+\varrho_l$) where ϱ_{bg} corresponds to the background cosmological density and ϱ_l is the 'thin lens' localized as a delta function at some point $r=r_l$. In other words, we assume,

$$\varrho_l(r)=\frac{\Sigma\delta(r-r_l)}{\sqrt{g_{\mathrm{rr}}}}=\frac{\Sigma\sqrt{1-kr_l^2}}{S(t_l)}\,\delta(r-r_l). \tag{37}$$

This equation defines the proper surface density Σ, i.e. the integral of ϱ_l along the proper radial length $dl=\sqrt{g_{\mathrm{rr}}}\,dr$ will be equal to Σ. The event (t_l, r_l) is connected by a radial null geodesic to the origin. We arrange the coordinate system such that the observer is at $r=0$ and the affine distance λ is zero at $r=0$.

Since ϱ_l vanishes everywhere except at $r-r_l$, equation (36) reduces to the focusing equation in FRW universe everywhere except at $r=r_l$. Let λ be λ_l at $r=r_l$. The solution to (37) can be now written as,

$$g=\begin{cases} a_1 D+b_1 C & \text{for } \lambda<\lambda_l \quad (38)\\ a_2 D+b_2 C & \text{for } \lambda>\lambda_l \quad (39)\end{cases}$$

where D and C are the two independent solutions discussed in Section 2.1. At the origin (location of the observer) we choose the initial conditions to be

$$g=0; \qquad \frac{dg}{d\lambda}=\frac{1}{S}\frac{dg}{dt}=\frac{1}{S_0}. \tag{40}$$

Comparison of (40) with (23) and (28) gives us

$$a_1=1; \qquad b_1=0 \tag{41}$$

so that g is just the angular diameter distance D for $\lambda<\lambda_l$. To fix a_2 and b_2 we have to match solutions at $\lambda=\lambda_l$. From the standard theory of differential equations we know that $g(\lambda)$ should be continuous across $\lambda=\lambda_l$ while $g'(\lambda)$ will be discontinuous. Integrating (37) from $(\lambda_l-\varepsilon)$ to $(\lambda_l+\varepsilon)$ and noting that ϱ_{bg} is continuous, we get

$$g'(\lambda_l+\varepsilon)-g'(\lambda_l-\varepsilon)=-4\pi G\int_{\lambda_l-\varepsilon}^{\lambda_l+\varepsilon}\frac{\varrho_l}{S^2}g\,d\lambda. \tag{42}$$

Now [using (14) and (37)]

$$\int\frac{\varrho_l g}{S^2}\,d\lambda=\int\frac{\varrho_l g}{S^2}\,S\,dt=\int\varrho_l g\left(\frac{dt}{S}\right)=\int\varrho_l g\,\frac{dr}{\sqrt{1-kr^2}}$$

$$=\Sigma\int\frac{g}{S}\,\delta(r-r_l)\,dr. \tag{43}$$

So we get

$$g'(\lambda_l+\varepsilon)-g'(\lambda_l-\varepsilon)=-4\pi G\Sigma\,\frac{g(\lambda_l)}{S(\lambda_l)}. \tag{44}$$

Thus the matching conditions at $\lambda=\lambda_l$ are

$$a_2D'+b_2C'-D'=-4\pi G\Sigma\,\frac{D}{S} \tag{45}$$

$$a_2D+b_2C-D=0. \tag{46}$$

Since $CD'-DC'=1$, these equations are trivial to solve. We find that:

$$a_2=1-4\pi G\Sigma\left(\frac{CD}{S}\right)_l; \qquad b_2=4\pi G\Sigma\left(\frac{D^2}{S}\right)_l. \tag{47}$$

We now know the solution to the focusing equation for all λ. Let λ_c be the point conjugate to the origin, so that $g(\lambda_c)=0$. Using (39) and (47), we get

$$\frac{C(\lambda_c)}{D(\lambda_c)}=-\frac{a_2}{b_2}=-\left(\frac{1-4\pi G\Sigma C_lD_l/S_l}{4\pi G\Sigma D_l^2/S_l}\right). \tag{48}$$

Let the source be located at $\lambda=\lambda_s$. The condition for production of multiple images of this source is simply $\lambda_c<\lambda_s$ (i.e. the conjugate point should occur 'in front of' the source (see Section 2.2). From

(25) we see that, for $\lambda_c<\lambda_s$,

$$\frac{C(\lambda_c)}{D(\lambda_c)}-\frac{C(\lambda_s)}{D(\lambda_s)}=\int_{\lambda_c}^{\lambda_s}\frac{d\lambda}{D^2}>0. \tag{49}$$

In other words for $\lambda_c<\lambda_s$,

$$\frac{C(\lambda_c)}{D(\lambda_c)}>\frac{C(\lambda_s)}{D(\lambda_s)}. \tag{50}$$

Combining with (48) we finally arrive at the condition for multiple imaging by a thin lens:

$$\frac{C(\lambda_s)}{D(\lambda_s)}<\frac{C(\lambda_l)}{D(\lambda_l)}-\frac{S_l}{4\pi G\Sigma D_l^2}. \tag{51}$$

Rearranging and using the definition of $D(\lambda_1, \lambda_2)$ of (53), we can write this condition as

$$\Sigma>\Sigma_c \tag{52}$$

where

$$\Sigma_c=\frac{D(\lambda_s)c^2}{4\pi GD(\lambda_l)D(\lambda_s,\lambda_l)} \tag{53}$$

[where we have re-introduced the correct factor of c in (53)]. This is precisely the result proved by Subramanian & Cowling (1986). If the null geodesic connecting $\lambda=0$ and $\lambda=\lambda_s$ encounters anywhere in between a surface density larger than Σ_c then multiple images can be produced. (Note that the neglect of the shear term in the above is justified since we are only interested in a sufficient condition on Σ for multiple imaging.)

It is interesting to compare the role played by the angular diameter distance in the above derivation with the manner in which D is introduced in the conventional derivations. In conventional derivations, an impact parameter b is introduced, in terms of which the bending angle of light beam is given by

$$\tilde{\theta}=\frac{4GM}{c^2b}. \tag{54}$$

If we now *assume* that b should be identified with proper distance in the deflector plane, then it follows that angular diameter distances should be used to convert b to angles and vice versa. (Note that $D\delta\tilde{\theta}$ represents *proper* length at a distance D.)

However, it is not quite clear that b in (54) should represent proper length in the deflector plane. In the Schwarzschild geometry the exact expression for $\tilde{\theta}$ is quite complicated. In particular $\tilde{\theta}^{-1}$ is not simply proportional to the *proper* length. One may argue that (54) is not supposed to be an exact result anyway, but valid only in the Newtonian limit of $b \gg 2M$. But in this limit the distinction between proper length and coordinate distance R vanishes and so we are no longer sure whether b is proper length or coordinate distance R. Thus the assumption that b is the proper length in the deflector plane seems to be non-trivial. It is thus gratifying to see that the angular diameter D enters the result in the present formalism in a natural manner (as a solution to the focusing equation), and need not be introduced by invoking the assumption of proper distance on the deflector plane.

3.2 MULTIPLE THIN LENSES

As a second example, we shall consider an intermediate case between thin and thick lenses – a set of well-separated thin lenses. One of the motivations of this subsection is also to introduce a simple geometrical – diagrammatic – method which proves useful in establishing the general results outlined in Section 3.3. For the sake of simplicity, we shall ignore the cosmological effects and assume the space-time to be flat between the lenses. (The cosmological effects can be easily incorporated by using D instead of flat space distance.) This allows us to use ordinary flat space distance x for λ.

We begin by giving a simple geometrical interpretation to the action of a single thin lens. This can be most easily done by using the θ-equation:

$$\frac{d\theta}{dx}=-\theta^2-\frac{4\pi G\varrho}{c^2}. \tag{55}$$

Suppose there is a source at $x=x_s$ and a lens at $x=x_l$. From (10) and the initial conditions (40) (which imply $\theta\to\infty$ at $x\to 0$) we know that the solution to (55) is

$$\theta(x)=\begin{cases} x^{-1}; & x<x_l \quad (56)\\ -(x_c-x)^{-1}; & x>x_l. \quad (57)\end{cases}$$

Clearly x_c is the point conjugate to the origin as $\theta\to-\infty$ at $x=x_c$. As in the previous section, we can determine the x_c by matching solutions across $x=x_l$. Integrating (55) across x_l we obtain

$$\frac{1}{x_c-x_l}+\frac{1}{x_l}=\frac{4\pi G\Sigma}{c^2}. \tag{58}$$

The condition for the multiple imaging ($x_c<x_s$) implies $(x_c-x_l)^{-1}>(x_s-x_l)^{-1}$, or, equivalently

$$\frac{1}{f}\geqslant\frac{1}{x_l}+\frac{1}{x_s-x_l} \tag{59}$$

where we have defined the 'focal length' of the lens f as $(4\pi G\Sigma/c^2)^{-1}$. This equation has a simple interpretation in terms of Fig. 1. Given the observer position ($x=0$) and a source position ($x=x_s$) we can draw two curves x^{-1} and $-(x_s-x)^{-1}$ shown in the figure. The effect of the lens is simply to pull down the value of θ by f^{-1} at the location of the lens. If the lens could pull down θ below – or, at least up to – the curve $-(x_s-x)^{-1}$ then we will have multiple images.

Several other interesting features can be easily seen from this figure: For a fixed value of x_s (fixed source distance) (i) The minimum distance ($4/x_s$) between the curves is at midway ($x_s/2$). A lens with $f^{-1}<4x_s^{-1}$ cannot produce multiple imaging wherever it is placed. (ii) For any $f^{-1}>4x_s^{-1}$, we can keep the lens anywhere in a strip of width $x_s(1-4fx_s^{-1})^{1/2}$ about the mid-point and still produce multiple images. (iii) Now suppose we consider x_l to be given and x_s to be variable. If $f^{-1}>x_l^{-1}$, the lens will make θ negative at $x_l+\varepsilon$ (i.e. it will pull θ below the axis). Once this happens there is bound to be a conjugate point at some $x-x_c$. If x_s is variable then for sufficiently large x_s (i.e. if $x_s>x_c$) we will have multiple images.

All the above concepts are easily generalized to the case with more than one lens. Fig. 2 shows the case of two lenses at x_1 and x_2 with focal lengths f_1 and f_2. It is easy to see that two lenses are less 'efficient' compared to a single lens of $\Sigma=\Sigma_1+\Sigma_2$, In other words if the lenses at x_1 and x_2 and with surface densities Σ_1 and Σ_2 respectively, act together to produce multiple images, then so can a single lens kept at an appropriate value of x (with $x_1\leqslant x\leqslant x_2$), but having a density $\Sigma<\Sigma_1+\Sigma_2$. Consider the case in which the first lens reduces θ from positive to negative values. (The other cases can be worked out similarly.) Where should the second lens be placed so that we can bring θ down to curve 2 with least Σ_2? Clearly the gap between curves 1 and 2 increases monotonically with x. Hence the most efficient way of lensing corresponds to moving the second lens to the position of the first one, $x_2=x_1$. Thus we have now reached a configuration with a single lens at $x=x_1$. But the process of moving the second lens towards the first decreases the value of its x

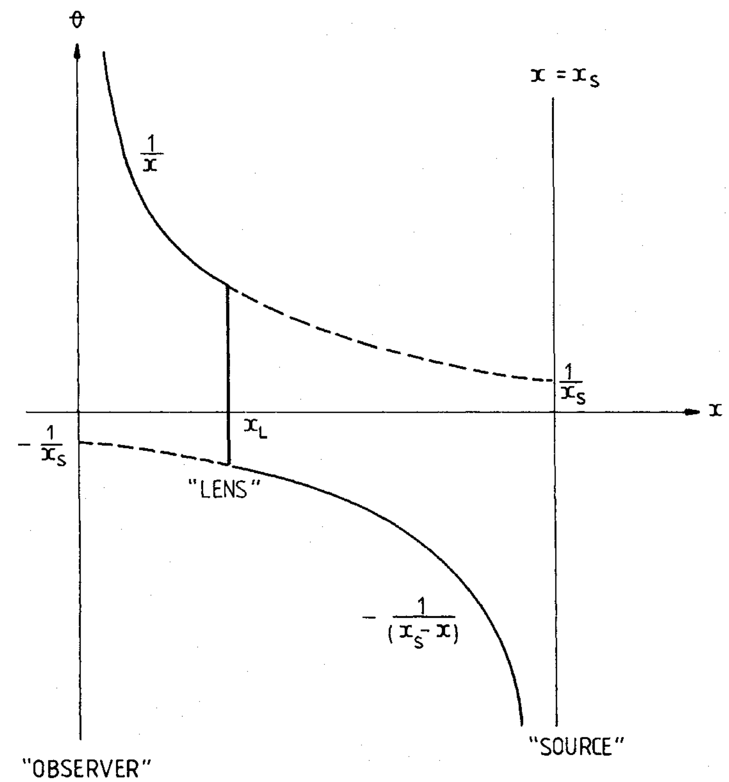

Figure 1. A schematic plot of θ versus x illustrating the action of a thin lens placed at x_L. If the lens can 'pull down' θ below or at least up to the curve $-(x_s-x)^{-1}$ then it will be capable of producing multiple images.

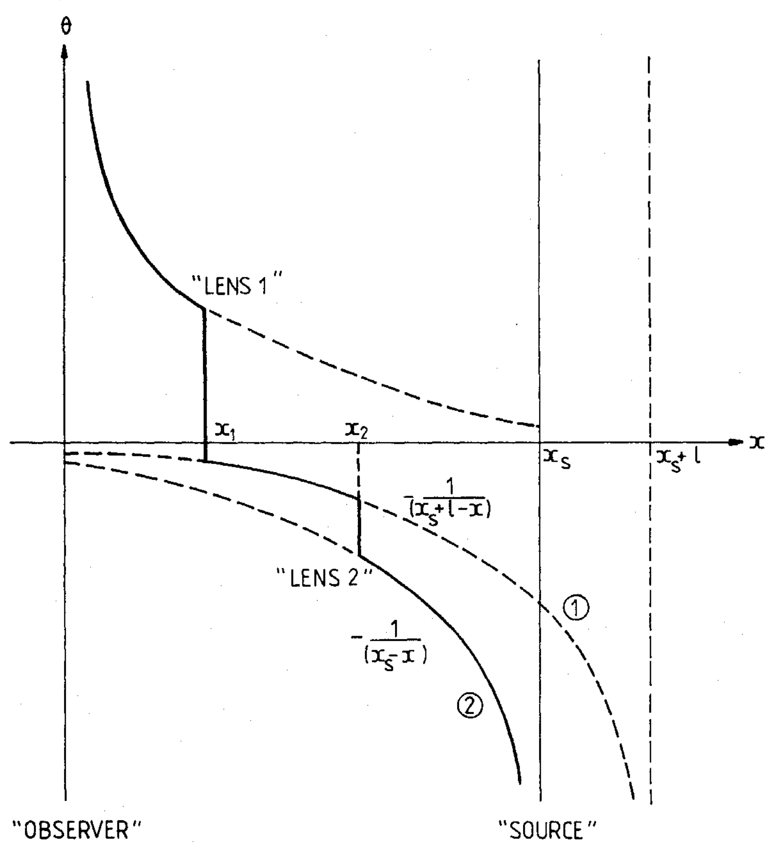

Figure 2. The corresponding schematic plot of θ versus x for two thin lenses at x_1 and x_2, as described in the text.

coordinate and hence the gap between the curves 1 and 2. In other words we require lesser and lesser surface density for the second lens as it is moved closer and closer to x_1 and the least density at $x=x_1$. (Note that the gap between the curves measures the value of the Σ required. The situation is quite analogous to the case of two ordinary optical convex lenses. We know that two lenses of focal length f_1 and f_2 separated by a distance D will act as an effective lens with focal length $f_{\rm eff}$ where $f_{\rm eff}^{-1}=f_1^{-1}+f_2^{-1}-D/f_1f_2$.) This also allows us to draw the following conclusion: We know from the above that if two lenses kept at different distances can produce multiple images so can an appropriately placed single lens. However, a single lens just cannot produce multiple images unless $c^2/\pi G x_s<\Sigma$. Since $\Sigma<\Sigma+\Sigma_2$, it follows that $(\Sigma_1+\Sigma_2)$ should exceed $(c^2/\pi G x_s)$ if multiple images have to be produced. These arguments can be easily generalized to N lenses. For the sake of completeness we give the solution $\theta(x)$ for N lenses of focal lengths $f_1, f_2 \ldots f_N$ located at $x_1, x_2 \ldots x_N$. The general solution consists of the set $(\theta_1, \theta_2 \ldots \theta_N)$, with θ_i applicable in the range $x_i<x<x_{i+1}$, and

$$\theta_i=\frac{1}{x-c_i} \tag{60}$$

with

$$c_i=x_i-\frac{f_i(x_i-c_{i-1})}{f_i-x_i+c_{i+1}}. \tag{61}$$

Equation (61) determines c_i inductively given $C_0=0$. The conclusions obtained earlier can be verified from (60), (61) as well, after somewhat lengthy algebra.

3.3 GENERAL FEATURES OF THICK LENSES

Realistic thick lenses differ from the idealized models discussed in the previous section in two crucial ways: (i) they will not have constant densities along every null geodesic (remember that a beam travelling at an angle will experience a different density profile compared to the beam travelling perpendicular to the lens when the lens is thick); (ii) cosmological effects can be significant when L is large and comparable to other distances. In this section we shall discuss some results which can be proved even if the density of the lens is not constant. The second problem requires knowledge about the cosmological solutions with large inhomogeneities, and we hope to discuss this in a future publication.

One general sufficiency condition for multiple images can be proved directly from the equation for θ. Consider an astrophysical situation in which: (i) Most of the lens density is confined between affine distances λ_1 and λ_2. (ii) Sources exist for arbitrarily large values of λ. (This assumption, of course, is unrealistic; some suitable approximation has to be resorted to while applying this condition.) In this case we only have to make sure that $\theta(\lambda_2)<0$. This would make $\theta(\lambda)$ to follow one of the curves of the form $-(\lambda_c-\lambda)^{-1}$ for $\lambda>\lambda_2$. Such a behaviour always leads to a conjugate point at some $\lambda=\lambda_c$. [Normally one has to ensure that $\lambda_s>\lambda_c$; here our assumption (ii) assures the validity of $\lambda_s>\lambda_c$ for at least some sources.]

Integrating the θ equation between λ_1 and λ_2 we get

$$\theta(\lambda_2)-\theta(\lambda_1)=-\int_{\lambda_1}^{\lambda_2}\theta^2\, d\lambda-\frac{4\pi G\Sigma}{c^2} \tag{62}$$

or

$$\theta(\lambda_2)=\theta(\lambda_1)-\frac{4\pi G\Sigma}{c^2}-\int_{\lambda_1}^{\lambda_2}\theta^2\, d\lambda \tag{63}$$

so the condition $\theta(\lambda_2)<0$ is ensured if $4\pi G\Sigma/c^2>\theta(\lambda_1)$ or, using $\theta(\theta_1)=1/\lambda_1$, if

$$\Sigma=\int_{\lambda_1}^{\lambda_2} \varrho(\lambda)\, d\lambda > \frac{c^2}{4\pi G\lambda_1}. \tag{64}$$

In many practical situations one would be interested in the multiple imaging of specific sources located at definite redshifts. It may not be possible to assume that sources exist for arbitrarily high values of λ. In such a case, we can still prove the following result: If a thin lens of surface density Σ cannot produce multiple images of a given source, then a thick lens of same Σ cannot produce multiple images, as long as the effects of shear are neglected. This result shows that a thick lens cannot be 'better' than a thin lens, even though it may have more mass. Note that when the effects of shear are neglected any thick lens may be looked at as effectively a set of N thin lenses with $\Sigma_i=\varrho(\lambda_i)\, d\lambda_i$, $d\lambda_i \to 0$, $N\to\infty$. One way of establishing the above result on the relative efficiencies of thin and thick lenses is to do this and use the results of Section 3.2. We give below a more elegant argument based directly on the focusing differential equation for a general $\varrho(\lambda)$ and using the geometrical method introduced in Section 3.2.

We plot the $(\theta-\lambda)$ diagram for a thick lens with some arbitrary $\varrho(\lambda)$ [see Fig. 3]. We have taken $\theta(\lambda)$ to be λ^{-1} for $\lambda<\lambda_1$ and $-(\lambda_s-\lambda)^{-1}$ for $\lambda>\lambda_2$, and the lens to extend from λ_1 to λ_2. We do not know the explicit form of $\theta(\lambda)$ for $\lambda_1<\lambda<\lambda_2$, but it will not be needed. It could be some smooth, monotonically decreasing [it is clear from (9) that θ should decrease monotonically for positive ϱ and from (63) that $\theta(\lambda_1+\varepsilon)<1/\lambda_1$ for small ε] function, as shown in the figure. Let this curve cut the $\theta=0$ axis at some $\lambda=\lambda_0$. We produce the λ^{-1} and $-(\lambda_s-\lambda)^{-1}$ curves to cut the $\lambda=\lambda_0$ line at C and

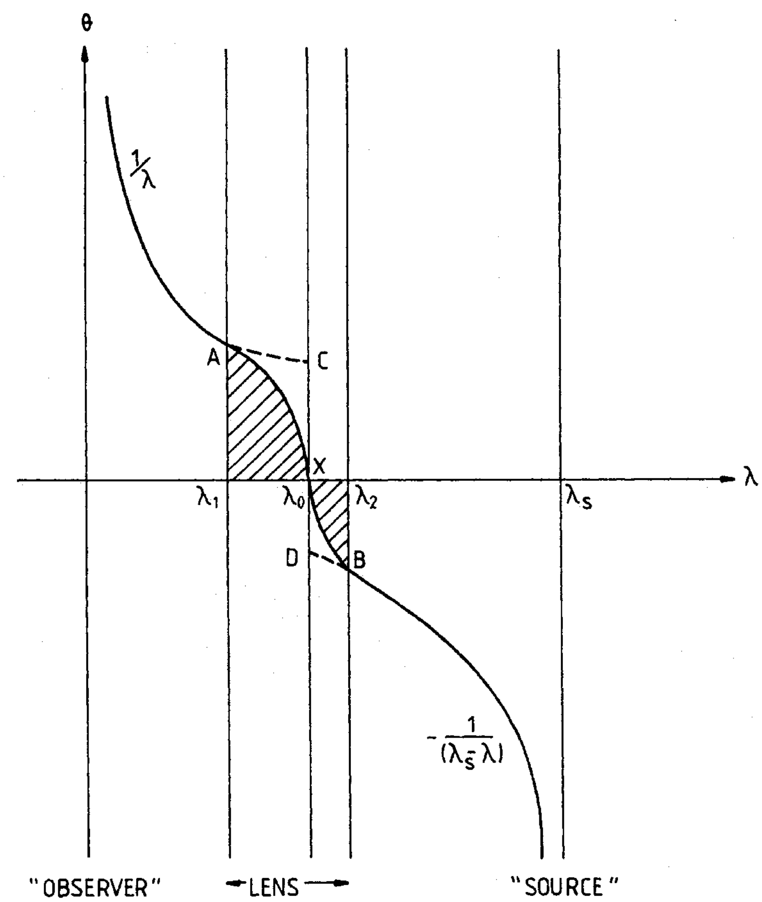

Figure 3. The corresponding schematic plot of θ versus λ for an arbitrary sheerless thick lens with non-vanishing density between λ_1 and λ_2.

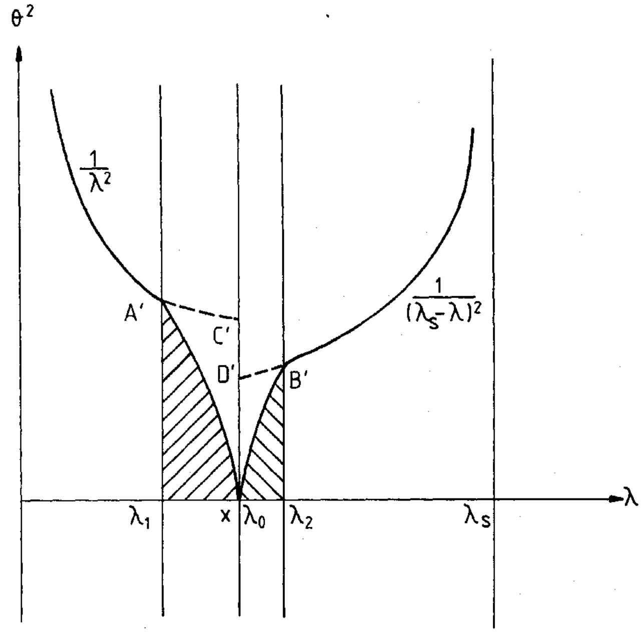

Figure 4. A schematic plot of θ^2 versus λ for the case shown in Fig. 3.

D respectively. The curve AXB describes the action of a thick lens while the path ACDB describes the action of a thin lens kept at $\lambda=\lambda_0$, so that these lenses both produce a conjugate point at $\lambda=\lambda_c=\lambda_s$. We are interested in comparing the surface densities of this thin and thick lens. We plot schematically $\theta^2(\lambda)$ in Fig. 4, for the $\theta(\lambda)$ given in Fig. 3. We notice from Fig. 4 that, in the range $\lambda_1<\lambda<\lambda_2$, θ^2 for any thick lens is less than θ^2 for a single thin lens kept at $\lambda=\lambda_0$. In other words,

$$\left(\int_{\lambda_1}^{\lambda_2}\theta^2\,d\lambda\right)_{\text{thick lens}}<\left(\int_{\lambda_1}^{\lambda_2}\theta^2\,d\lambda\right)_{\text{thin lens}}. \tag{65}$$

Since θ at λ_1 and λ_2 is the same for both lenses, we must have from equation (63)

$$\frac{4\pi G}{c^2}\,\Sigma_{\text{thin}}+\left(\int_{\lambda_1}^{\lambda_2}\theta^2\,d\lambda\right)_{\text{thin}}=\frac{4\pi G}{c^2}\,\Sigma_{\text{thick}}+\left(\int_{\lambda_1}^{\lambda_2}\theta^2\,d\lambda\right)_{\text{thick}} \tag{66}$$

we conclude

$$\Sigma_{\text{thin}}<\Sigma_{\text{thick}}. \tag{67}$$

In other words, we have established the following result: suppose some thick lens with Σ_{thick} produces multiple images of a source at $\lambda_s>\lambda_c$ then there will always exist some thin lens located between λ_1 and λ_2 which will produce multiple images of the same source with lower Σ_{thin} ($<\Sigma_{\text{thick}}$).

We can also now put bounds on Σ_{thick} using our known bounds on Σ_{thin}. (Since $\Sigma_{\text{thin}}>\Sigma_c$ implies $\Sigma_{\text{thick}}>\Sigma_c$.) Since the thin lens has to be located between λ_1 and λ_2, the bounds depend on the location of λ_1, λ_2. From our analysis of thin lenses in Section 3.2, we know that $\Sigma_{\text{thin}}>\Sigma_c'$ where

$$\Sigma_c'=\begin{cases} c^2[1/\lambda_2+1/(\lambda_s-\lambda_2)]/4\pi G & \lambda_2<\tfrac{1}{2}\lambda_s \\ c^2(4/\lambda_s)/4\pi G & \lambda_1<\lambda_s/2<\lambda_2 \\ c^2[1/\lambda_1+1/(\lambda_s-\lambda_1)]/4\pi G & \lambda_1<\tfrac{1}{2}\lambda_s. \end{cases} \tag{68}$$

Thus a thick lens located between λ_1 and λ_2 can produce multiple images only if its surface density Σ_{thick} is larger than the Σ_c' given above.

Of the three cases considered above, the case with $\lambda_1<\frac{1}{2}\lambda_s<\lambda_2$ is most frequently encountered. This is because in practical situations we will never be able to fix some λ_1 and λ_2 and be sure that $\varrho=0$ for $\lambda<\lambda_1$ and $\lambda>\lambda_2$. We have to assume that, effectively, ϱ exists everywhere between $0<\lambda<\lambda_s$. This gives a strict lower bound on Σ_{thick}.

$$\Sigma_{\text{thick}}>\frac{c^2}{\pi G\lambda_s}. \tag{69}$$

Two cautionary remarks are in order regarding the bounds derived above. First, note that these are obtained by neglecting shear terms. If the shear terms for a thick lens outweigh that of a thin lens, the above conditions can be violated. Secondly, the relationship between λ and geometrical distance can be fairly complicated for a thick lens. The Σ's we are dealing with are defined by integrating ϱ along the null geodesic rather than by the usual 'line-of-sight' integration. Condition (69), for example, only means that Σ_{thick} defined along *some* light ray has to be larger than Σ_{thin}.

4 Discussion and conclusions

In this paper we examine the general conditions for multiple imaging by an arbitrary (thick or thin), smooth and bounded gravitational lens. We establish a necessary and sufficient condition for a smooth and bounded lens to be capable of producing multiple images: The lens should produce a point conjugate to the observer along some null geodesic, at an affine distance smaller than that of the source. (We assume of course that before the introduction of the lens, the background space-time does not itself produce any conjugate points before the relevant source affine distances.) This general criterion is then used to decide if any given lens can produce multiple images. One has only to check if the lens leads to the existence of points conjugate to the observer and this can in turn be achieved by solving the focusing equation.

We first solved the focusing equation for a thin lens in a Robertson–Walker universe and reproduced the sufficient condition for multiple imaging proved by Subramanian & Cowling (1986): that if the projected surface density of a thin lens exceeds a critical value $c^2D_s/4\pi GD_lD_{ls}$ then the lens is capable of multiple imaging. We then went on to consider thick gravitational lenses. These have proved rather intractable in the past (see, however, Peacock 1986; Kovner 1987) and the power of the present approach can be seen in its ability to at least begin addressing general questions regarding thick gravitational lensing. We study the multiple thin lens before going on to consider general thick lens. We establish sufficient conditions for multiple imaging by this lens. We also show that, if the effects of shear are neglected, a thick gravitational lens with an arbitrary density profile $\varrho(\lambda)$ is less efficient than a thin lens with the same projected surface density: More specifically, if a thick lens with $\varrho(\lambda)$ non-zero for $\lambda_1<\lambda<\lambda_2$ can produce multiple images of some source, then so can a thin lens located somewhere between λ_1 and λ_2, but with

$$\Sigma<\int_{\lambda_1}^{\lambda_2}\varrho(\lambda)\,d\lambda$$

(i.e. Σ_{thin} less than Σ_{thick}).

One of the motivations for the present work was to examine whether large-scale inhomogeneities in the Universe (on scales larger than 50–100 Mpc, say) can be probed by using their multiple imaging property. Such inhomogeneities may be indicated by present observations of the space distribution of galaxies and by observations of large-scale streaming motions (*cf.* Lapparent, Geller & Huchra 1986; Burstein *et al.* (1986). The considerations of this work indicates that the projected surface density of such inhomogeneities (which act as thick gravitational lenses) has to exceed a density of order Σ_c (of equation 1) for this purpose. So a very large-scale inhomogeneity of mass $\sim 10^{17} M_{\odot}$ at cosmological distances has to have a core$\lesssim$6 Mpc to be capable of producing multiple images of distant sources. It is not clear at present whether such inhomogeneities are plausible.

In all the above applications of our general criterion for multiple imaging to specific lenses, we have ignored the effects of 'shear'. As we mentioned earlier this is justified if we are interested in deriving sufficient conditions for multiple imaging, since shear always aids the focusing of light beams. However, to derive necessary conditions for multiple imaging [like those derived in Subramanian & Cowling (1986) for thin lenses] and to strengthen the claims about the relative efficiency of thick and thin lenses for multiple imaging, the effects of shear have to be included. We hope to address this issue in a later publication.

Acknowledgments

TP and KS were supported by SERC, UK. We thank Dr Rajaram Nityananda for many helpful discussions.

References

Burke, W. L., 1981. *Astrophys. J.*, **195**, L1.
Burstein, D., Davies, R. L., Dressler, A., Faber, S. M., Lynden-Bell, D., Terlevich, R. & Wegner, G., 1986. In: *Galaxies Distances and Deviations from Universal Expansion*, p. 123, eds Madore, B. F. & Tully, R. B., Reidel, Dordrecht, Holland.
Guillemen, V. & Pollock, A., 1974. *Differential Topology*, Prentice-Hall, Englewood Cliffs, N.J.
Kantowski, R., 1969. *Astrophys. J.*, **155**, 89.
Kovner, I., 1987. *Astrophys. J.*, **316**, 52.
Lapparent, V. D., Geller, M. J. & Huchra, J. P., 1986. *Astrophys. J.*, **302**, L1.
Peacock, J. A., 1986. *Mon. Not. R. astr. Soc.*, **223**, 113.
Sachs, R., 1961. *Proc. R. Soc. London*, **264**, 309.
Subramanian, K. & Cowling, S. A., 1986. *Mon. Not. R. astr. Soc.*, **219**, 333.
Whitney, H., 1955. *Ann. Math.*, **62**, 374.

Appendix A

Consider the vector field $\mathbf{X}=(u-u_-)\mathbf{i}+(v-v_-)\mathbf{j}$ on α. (Here $\mathbf{i}$ and $\mathbf{j}$ are unit vectors in the x and y directions.) The zeroes of this vector field (x_i, y_i) (the points on α where $\mathbf{X}=0$) give the location of the images on α for a source located at (u_-, v_-) on B_s. From the definition of the map ϕ_s we also have $u=c_1x$ and $v=c_2y$ as x and y tend to infinity where c_1 and c_2 are positive constants (since at spatial infinity there is no bending due to the lens). Consider a circle C of radius $R\to\infty$ on the x–y plane. The index of the vector field $\mathbf{X}$ on the curve is $+1$. (The index of a vector field on a curve C is defined as the number of rotations the vector on the curve makes in units of 2π as we go once around the curve in an anticlockwise direction). From standard results in differential topology (see Guillimen & Pollock 1974) we know that the index of $\mathbf{X}$ around C is the sum of the indices around the zeroes of $\mathbf{X}$ (the points on α where $\mathbf{X}=0$). The indices of the zeroes are also generically either $+1$ or -1. Since the index on C is $+1$, the total number of zeroes of $\mathbf{X}$ inside C has to be odd, with at least one zero of index $+1$ and other zeroes appearing in pairs with indices of $+1$ and -1.

The index of a zero of $\mathbf{X}$ can also be related to the sign of the Jacobian of the map ϕ_s at (x_i, y_i) as follows. Consider an infinitesimal circle of radius ε, C_ε, around a zero at $P_0(x_0, y_0)$. The index of the zero is evaluated by calculating how much $\mathbf{X}$ rotates as one moves around C_ε in units of 2π and is then taking the limit $\varepsilon\to 0$. In the limit $\varepsilon\to 0$, the vector field $\mathbf{X}$ on C_ε tends to the vector

$$\mathbf{X}=\left[\frac{\partial u}{\partial x}\bigg|_{P_0}(x-x_0)+\frac{\partial u}{\partial y}\bigg|_{P_0}(y-y_0)\right]\mathbf{i}+\left[\frac{\partial v}{\partial x}\bigg|_{P_0}(x-x_0)+\frac{\partial v}{\partial y}\bigg|_{P_0}(y-y_0)\right]\mathbf{j} \quad \text{(A1)}$$

where we need to retain only the leading term in the Taylor expansion of **X** about the zero. The index around the zero i_0 is then given by

$$i_0 = \frac{1}{2\pi}\int_0^{2\pi} \frac{d\bar{g}}{d\phi}\, d\phi \tag{A2}$$

where

$$\bar{g} = \frac{\left.\dfrac{\partial u}{\partial x}\right|_{P_0} + \left.\dfrac{\partial u}{\partial y}\right|_{P_0} \tan\phi}{\left.\dfrac{\partial v}{\partial x}\right|_{P_0} + \left.\dfrac{\partial u}{\partial y}\right|_{P_0} \tan\phi}. \tag{A3}$$

Note that the integral in (A2) is not $[\bar{g}(2\pi)-\bar{g}(0)]/2\pi$ since **g** is not a continuous function. A straightforward but lengthy calculation gives $i_0=\text{sign}\,(J_{P_0})$, where

$$J_{P_0} = \left.\frac{\partial u}{\partial x}\right|_{P_0} \left.\frac{\partial v}{\partial y}\right|_{P_0} - \left.\frac{\partial u}{\partial y}\right|_{P_0} \left.\frac{\partial v}{\partial x}\right|_{P_0}.$$

So the index of a zero of **X** at (x_i, y_i) is $+1$ or -1 if and only if the Jacobian at the image located at (x_i, y_i) is positive or negative respectively. So we have the result that a smooth and bounded lens produces at least one image with $J>0$ (index $+1$) and other images come in pairs, one with $J>0$, and the other with $J<0$ (index -1).

Appendix B

Before outlining the proof one must first note the following: (i) In general at a conjugate point the bundle of light rays from the observer do not focus to a point but to a line. Focusing to a point is non-generic and requires high symmetry. We will assume in what follows that all conjugate points are 'generic'. (ii) Suppose we consider a small *oriented* cross-section $\boldsymbol{\delta}$ of a bundle of light rays from the observer. Then as the fiducial ray passes through a 'generic' conjugate point the orientation of $\boldsymbol{\delta}$ reverses (or changes sign). Note that if the conjugate point is 'non-generic', and the rays focus to a point, the orientation of $\boldsymbol{\delta}$ is preserved (though $\boldsymbol{\delta}$ is 'inverted' inside out).

Now suppose a lens produces points conjugate to the observer at $\lambda=\lambda_c<\lambda_s$ on some null rays. Consider any such ray going through (p, q) on α. It can have an odd or even number of conjugate points on it before reaching a $\lambda=\lambda_s$ surface: First consider the case when there are an odd number of conjugate points on this ray. Then the orientation δ about this ray changes sign an odd number of times. So the sign of the Jacobian J of the map ϕ_s is negative at (p, q). But J is positive at infinity on α. So J has to be zero on some curve around (p, q). The image of this curve under ϕ_s is clearly a critical curve on $\lambda=\lambda_s$ surface. So the caustic surface has to cut the $\lambda=\lambda_s$ surface.

On the other hand let there be an even number of conjugate points on the ray through (p, q) extended to $\lambda=\lambda_s$. Take any curve C which passes through (p, q) and extends to infinity in two different directions. Consider the family of null rays which pass through the points on this curve. The conjugate points on these rays will trace out a set of curves say $\tilde{C}$ on the caustic surface. Now on a null ray through 'infinity' there are no conjugate points. (So the number of conjugate points on a null ray has to decrease from an even number to zero as one takes the ray along the family from (p, q) to 'infinity'.) Since ϕ_s is smooth, the curves $\tilde{C}$ can behave in only one of two ways (i) a curve in $\tilde{C}$ has to cross the $\lambda=\lambda_s$ surface, at say R and go to larger values of λ. (And for null rays on 'either side' of R the number of conjugate points differs from one.) In this case R is a conjugate point to O. (ii) The other possibility is that the curves $\tilde{C}$ are a set of closed bounded curves which never reach a $\lambda=\lambda_s$ surface. We now show that this possibility is non-generic: because in this case there will be a null ray cutting α at say S somewhere between (p, q) and infinity, which will be tangent to a curve in $\tilde{C}$ at say $\bar{S}$ and for which $\bar{S}$ will be the only conjugate point. A null ray on one side of S on C will have two conjugate points, while on the other side of S will have none. So the sign of J will remain the same (and positive) as we pass through S along C, and J is positive at S. But the ray through S to $\lambda=\lambda_s$ has only one conjugate point. So J at S can be positive if and only if this conjugate point is non-generic (a bundle of rays focus to point). In fact, since this situation obtains for any curve through (p, q) to infinity, one will then have a curve of non-generic conjugate points on the caustic surface. Such a situation is not stable to arbitrarily small perturbations of the lens, and is not generic. So we have that any generic lens cannot produce $\tilde{C}$ which are closed and bounded and which do not reach $\lambda=\lambda_s$ surface.

The converse is also easily proved. Suppose a lens can produce conjugate points to O on a $\lambda=\lambda_s$ surface. Then there is a non-degenerate critical curve on the $\lambda=\lambda_s$ surface. Since the caustic surface is also non-degenerate, it will have to extend to $\lambda<\lambda_s$. So there will then be a conjugate point at $\lambda=\lambda_c<\lambda_s$.