Fermat potentials for nonperturbative gravitational lensing

Simonetta Frittelli
Department of Physics, Duquesne University, Pittsburgh, Pennsylvania 15282

Thomas P. Kling
Department of Physics, Astronomy and Geophysics, Connecticut College, New London, Connecticut 06320

Ezra T. Newman
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260
(Received 22 January 2002; published 11 June 2002)

The images of many distant galaxies are displaced, distorted and often multiplied by the presence of foreground massive galaxies near the line of sight; the foreground galaxies act as gravitational lenses. Commonly, the lens equation, which relates the placement and distortion of the images to the real source position in the thin-lens scenario, is obtained by extremizing the time of arrival among all the null paths from the source to the observer (Fermat's principle). We show that the construction of envelopes of certain families of null surfaces constitutes an alternative variational principle or version of Fermat's principle that leads naturally to a lens equation in a generic spacetime with any given metric. We illustrate the construction by deriving the lens equation for static asymptotically flat thin lens spacetimes. As an application of the approach, we find the bending angle for moving thin lenses in terms of the bending angle for the same deflector at rest. Finally we apply this construction to cosmological spacetimes (FRW) by using the fact they are all conformally related to Minkowski space.

DOI: 10.1103/PhysRevD.65.123007 PACS number(s): 98.62.Sb, 04.90.+e, 95.30.Sf

I. INTRODUCTION

It is the purpose of this paper to point out, study and apply an idealized construction of gravitational lens equations that is of potential use in many physical situations—from exact lensing to the weak-field thin-lens scenario—by means of an alternative version of the standard usage of Fermat's principle.

The fundamental aspect of gravitational lensing theory is the construction of the past light cone of an observer. This directly leads to the idea of the mapping from the space of images—the celestial sphere of the observer—to the space of the sources—usually to the ''source plane,'' though this specialization is by no means necessary. The mapping is carried out by following, backwards in time, the null geodesics of the light cone, from the observer to the source. In other words, by knowing where an image appears on the observer's celestial sphere and knowing the null geodesics that generate the past null cone, one could in principle follow the rays back to the source. In addition, it is often of considerable importance to know the transit times between the emission of light and its arrival at the observer. In fact, in view of Fermat's principle, the actual path taken by light is a local extremum of the transit time of all possible neighboring null paths, which leads to the formulation of gravitational lensing via Fermat's principle. Much of contemporary lensing theory is based on the construction, on a simple background (either Minkowski or a cosmological spacetime), for weak fields and with a thin-lens and small-angle approximation, of an appropriate transit time function. Then, by the local extremization of the time function, a lens equation is constructed. In the usual approach, the time function (referred to as a Fermat potential) represents the transit time along all possible null curves, not necessarily geodesic, that connect the source and the observer [1–3]. The extremization produces a suitable lens equation by selecting those curves that are geodesic, i.e., the variation of the travel time with respect to the paths, restricted to the condition that the paths be null, is equivalent to the geodesic equation with null tangents.

We wish to show that there is an attractive alternative (and significantly different) variational principle (an alternative Fermat's principle) for use in lensing that can be applied, at least in principle (and, with approximations, also in practice), in a generic situation. The basic framework is to begin with a general four-dimensional Lorentzian spacetime where the geometry (the metric) of the spacetime is to be considered as the ''gravitational lens.'' In other words, the detailed lens properties are to be coded directly into the metric tensor. We then consider a 2-parameter family of null (or characteristic) surfaces passing through an observer's world line at a given time. In fact, we assume that the family of surfaces is sufficiently generic for the null normals at the observer to be distinct and span the sphere of null directions (often just an open neighborhood of the sphere is sufficient). This family of null surfaces then contains all the points on the observer's light cone (or the open neighborhood), since at the intersection of the observer's world line with each surface the normal vector to the surface is null, geodesic, and lies on the surface. Each null geodesic that passes through the observer's world line on each of the null surfaces can then be followed into the past. These are the rays that an observer sees and constitute his *celestial* sphere (or an open neighborhood of it).

Now consider a point source of light moving along some given (source) timelike world line. We are interested in those null geodesics moving back in time from the observer to the

source, i.e., those geodesics that the observer ''sees'' as coming from the source. At *any given observer moment*, the observer will ''see'' a number of different light rays (or images), each of which, in general, will have intersected the source (or equivalently, will have been emitted by the source) at different source times. These countably few null geodesics lie on countably few of the null surfaces in the family that we are considering. Thus there are a number of surfaces in the 2-parameter family that intersect the world line of the source at a point that can be connected to the observer by a null geodesic on the surface. However, at *any observer moment*, all (or almost all) the other surfaces in the family intersect the source's world line as well, at varying times. In general, however, there are no curves lying on these surfaces connecting the source intersection point to the observer that are null geodesics. In fact, most of the curves that can be used to connect the source and the observer on each of these null surfaces are piecewise spacelike and/or null. For this reason, we prefer to drop any reference to paths, and keep the argument in terms of the null surfaces. The *source time t* at the intersection point is thus a function on the *sphere* of null surfaces intersecting the observer's world line at a particular observer's time T_0:

$$t = T(\theta, \varphi, T_0), \tag{1}$$

where (θ, φ) label the null surfaces. We now ask for the local extremes of T as a function of (θ, φ):

$$\partial_\theta T \equiv T_{,\theta}(\theta, \varphi, T_0) = 0, \tag{2a}$$

$$\partial_\varphi T \equiv T_{,\varphi}(\theta, \varphi, T_0) = 0. \tag{2b}$$

This operation picks out those null surfaces that possess curves from the observer traveling backwards to the source *which are null geodesics*. It constitutes our version of Fermat's principle. This version differs from the usual one not merely in form. If one thinks of it in terms of paths connecting the source and the observer, this version of Fermat's principle allows for curves that are neither null nor geodesic. It concentrates on and varies the null surfaces rather than the null curves.

For this reason we think of our version of Fermat's principle as an alternative to the usual one. Correspondingly, we refer to the function T as a *generalized Fermat potential.* With respect to the underlying meaning of our version of Fermat's principle, one can see that Eqs. (2a), (2b) are equivalent to the construction of the envelope of the null surfaces passing through the observer, which in turn is the past light cone of the observer. Thus we arrive at the observer's light cone starting from surfaces in a way analogous to the usual approach, which arrives at the observer's light cone starting from null paths.

In Sec. II we describe this construction in greater detail and justify the claim that it does pick out the null surfaces so that null geodesics connect the observer with the source thus yielding the past light cone. In Sec. III we illustrate the construction for the trivial case of Minkowski space without a lens, while in Sec. IV we illustrate it for a Minkowski space background with a static thin lens, the conventional scenario. In Sec. V we apply the construction to thin lenses that are moving in order to obtain the corrections to the lens equation due to the motion of the lens. In Sec. VI these ideas are applied to the FRW cosmologies using the fact that they all are conformally related to Minkowski space. We conclude, in Sec. VII, with remarks and an outlook of the potential applications of the construction.

We find this method quite enlightening concerning the underlying ideas in gravitational lensing theory. And although it might be difficult to apply in many practical physical situations without the use of heavy approximations, preliminary calculations suggest that, with further development, it could well be of practical use.

II. IMPLICITLY DEFINED GENERALIZED FERMAT POTENTIALS

We assume that we have a four-dimensional Lorentzian manifold, (M, x^a) with a given Lorentzian metric, g_{ab}. We further assume that null geodesics and null surfaces (solutions to the eikonal equation) can be constructed, in some local coordinates, wherever needed. In general, null surfaces develop wave front singularities and, in principle, one must worry about them. However, the trouble caused by the singularities is easily bypassed if the surfaces are parametrized by the null geodesics that rule them. The fundamental objects for us will be *complete integrals of the eikonal equation:* two-parameter families of foliations of the spacetime by null surfaces, so that at any spacetime point there is a sphere's worth of null surfaces through that point. The two parameters are arbitrary, but for our purposes we specify them as points on the sphere. In the following, then, we adopt this particular choice of parameters to label the solutions of the eikonal equation in the complete integral, and we specify them by complex stereographic coordinates $(\zeta, \bar{\zeta})$, rather than the regular spherical coordinates (θ, φ). The complete integral is given analytically by an expression of the form

$$Z(x^a, \zeta, \bar{\zeta}) \tag{3}$$

such that for each fixed value of $(\zeta, \bar{\zeta})$ the level surfaces of Eq. (3) are null. Equivalently, Eq. (3) satisfies the eikonal equation

$$g^{ab}\partial_a Z \partial_b Z = 0, \tag{4}$$

with two further conditions: the four functions $(Z, \partial_\zeta Z, \partial_{\bar{\zeta}} Z, \partial_{\zeta\bar{\zeta}} Z)$ form a rank-four set with respect to x^a *almost everywhere*, whereas *globally* the three functions $(Z, \partial_\zeta Z, \partial_{\bar{\zeta}} Z)$ form a rank-three set. How these conditions are used is explained shortly below. In addition, the null covectors $l_a(x^a, \zeta, \bar{\zeta}) \equiv \partial_a Z(x^a, \zeta, \bar{\zeta})$ at fixed x^a sweep out all null directions as $(\zeta, \bar{\zeta})$ vary, thus reproducing the local light cone. We rewrite Eq. (3) as the two-point function

$$G(x^a, x_0^a, \zeta, \bar{\zeta}) \equiv Z(x^a, \zeta, \bar{\zeta}) - Z(x_0^a, \zeta, \bar{\zeta}) = 0. \tag{5}$$

Any two points (x^a, x_0^a) satisfying Eq. (5) for fixed $(\zeta, \bar{\zeta})$ lie on the same null surface. It is clear that the gradient of G

with respect to either point is a null covector. If x_0^a is fixed, however, then the points x^a that satisfy Eq. (5) for each $(\zeta,\bar{\zeta})$ lie on different null surfaces, but x_0^a belongs to all. So we think of x_0^a as a special point where all the null surfaces intersect.

The point x_0^a is chosen to lie on the observer's world line given by $x_0^a = x_0^a(T_0)$, T_0 being some observer time. Though it is not needed and is used only for convenience, we choose, in the neighborhood of the source, a special coordinate system where the source remains at a fixed spatial point, x^i, and only the source time $x^0 = t$ evolves. The equation $G(x^a, x_0^a, \zeta, \bar{\zeta}) = 0$ then becomes

$$G(x^i, t, x_0^a(T_0), \zeta, \bar{\zeta}) = 0 \tag{6}$$

or, rewritten (since $\partial_t G \neq 0$ by assumption) as

$$t = T(x^i, T_0, \zeta, \bar{\zeta}), \tag{7}$$

it gives rise to our time function, Eq. (1). As a matter of notation, by analogy with the usual theory of gravitational lensing it is natural for us to refer to T as a *generalized Fermat potential.* However, the function G arises more naturally from the starting point of the eikonal equation; therefore, with a slight abuse of terminology, we refer to it as the *implicit Fermat potential*, and for all practical purposes we use Eq. (6) instead of Eq. (7).

The extremal values of T are calculated in terms of the implicit Fermat potential G by implicit differentiation of Eq. (6). Thus $\partial_\zeta T$ and $\partial_{\bar{\zeta}} T$ are obtained from

$$G_{,t}\partial_\zeta T + \partial_\zeta G = 0, \tag{8a}$$

$$G_{,t}\partial_{\bar{\zeta}} T + \partial_{\bar{\zeta}} G = 0, \tag{8b}$$

so that the vanishing of $\partial_\zeta T$ and $\partial_{\bar{\zeta}} T$ implies the vanishing of both $\partial_\zeta G$ and $\partial_{\bar{\zeta}} G$, and the reverse, i.e.,

$$(\partial_\zeta T, \partial_{\bar{\zeta}} T) = 0 \Leftrightarrow (\partial_\zeta G, \partial_{\bar{\zeta}} G) = 0. \tag{9}$$

Thus, setting $\partial_\zeta G = \partial_{\bar{\zeta}} G = 0$ gives us the extreme values of T. There is a deeper meaning to the extremization of T, in connection with the null surfaces in the family, as we subsequently show.

Returning to Eq. (5), with fixed x_0^a, we see that we have a two-parameter family of surfaces through a fixed point x_0^a. A new surface, with common tangent curves all passing through x_0^a, is the envelope of the family and is constructed by requiring that $\partial_\zeta G = \partial_{\bar{\zeta}} G = 0$. The triple

$$G(x^a, x_0^a, \zeta, \bar{\zeta}) = 0, \tag{10a}$$

$$\partial_\zeta G = 0, \tag{10b}$$

$$\partial_{\bar{\zeta}} G = 0, \tag{10c}$$

defines the envelope. For $x^a \neq x_0^a$ and in a region *without* wave front singularities [5,4], Eqs. (10b) and (10c) can be algebraically inverted so that

$$\zeta = \Upsilon(x^a, x_0^a), \tag{11a}$$

$$\bar{\zeta} = \bar{\Upsilon}(x^a, x_0^a). \tag{11b}$$

When they are substituted into Eqs. (10a)–(10c), we have the equation for a three-surface, namely the envelope:

$$G_{env}(x^a, x_0^a) \equiv G(x^a, x_0^a, \Upsilon(x^a, x_0^a), \bar{\Upsilon}(x^a, x_0^a)) = 0. \tag{12}$$

That $G_{env}(x^a, x_0^a) = 0$ is a null surface follows from both the fact that $\partial G / \partial x^a$ is a null vector (by assumption) and from Eqs. (10b) and (10c). That it is the null cone of the point x_0^a is shown by demonstrating that at $x^a = x_0^a$ the surface has a sphere's worth of tangents. This is seen by the following argument: in the limit as $x^a \to x_0^a$, Eqs. (10b), (10c) are identically satisfied for all values of $(\zeta,\bar{\zeta})$. This is a degenerate point where Eqs. (11a), (11b) *do not* hold. It then follows that the tangent vector, given by

$$\left.\frac{\partial G_{env}}{\partial x^a}\right|_{x=x_0} = \left.\frac{\partial}{\partial x^a} G(x^a, x_0^a, \zeta, \bar{\zeta})\right|_{x=x_0}, \tag{13}$$

has multiple values, yielding a null vector that sweeps out the tangent-space null cone at x_0^a, as $(\zeta,\bar{\zeta})$ moves over the sphere.

Our task is now to show that this extremization process picks out the particular surfaces that connect the source and observer by null geodesics; namely, the values of $(\zeta,\bar{\zeta})$ for each pair (x^a, x_0^a) given explicitly by Eqs. (11a), (11b)—or implicitly by Eqs. (10b), (10c)—label a geodesic null vector whose integral line passes through x_0^a and x^a. In order to show this, we notice that when Eqs. (10b), (10c) cannot be inverted, i.e., when the Jacobian

$$J = \begin{vmatrix} \partial_{\zeta\zeta} G & \partial_{\bar{\zeta}\zeta} G \\ \partial_{\zeta\bar{\zeta}} G & \partial_{\bar{\zeta}\bar{\zeta}} G \end{vmatrix} \tag{14}$$

vanishes, the null surfaces develop wave front singularities [5]. But by the assumption that $Z(x_0^a, x^a, \zeta, \bar{\zeta})$ is a complete integral and the rank conditions, in either case the three equations (10a)–(10c) can be solved for three (say x^i) of the four spacetime coordinates x^a, in terms of a fourth one (x^*) and $(\zeta,\bar{\zeta})$, namely

$$x^i = X^i(x_0^a, x^*, \zeta, \bar{\zeta}). \tag{15}$$

As a consequence of Eqs. (10a)–(10c), it turns out that the curves described by Eq. (15) by keeping $(\zeta,\bar{\zeta})$ fixed are null geodesics. This can be seen by implicit differentiation of Eqs. (10a)–(10c) with respect to x^*. Defining $t^a \equiv (1, \partial X^i / \partial x^*|_{\zeta,\bar{\zeta}})$, we obtain

$$G_{,a}(x^a, x_0^a, \zeta, \bar{\zeta}) t^a = 0, \tag{16a}$$

$$\partial_\zeta G_{,a} t^a = 0, \tag{16b}$$

$$\partial_{\bar{\zeta}} G_{,a} t^a = 0 \tag{16c}$$

where $(G_{,a}, \partial_\zeta G_{,a}, \partial_{\bar{\zeta}} G_{,a})$ are independent, by the rank condition, again. Equation (16a) implies that t^a has no component pointing out of the surface, so $t^a = \alpha g^{ab}(G_{,b} + \beta \partial_\zeta G_{,b} + \bar{\beta}\partial_{\bar{\zeta}} G_{,b})$. Using this in Eqs. (16b) and (16c) implies that $\beta = \bar{\beta} = 0$ and hence $t^a = \alpha g^{ab} G_{,b}$, and thus the tangent vector is proportional to $g^{ab}G_{,b}$ being therefore null, and the integral curves are null geodesics.

We thus obtain the parametric description of all the null geodesics through x_0^a. Notice that having this parametric description of the light cone (in terms of the observer's celestial sphere) is entirely equivalent to having a lens equation relating the angular position of the source at a given distance to the angular position of the image on the celestial sphere, plus a time of arrival equation yielding the transit time of the light signal from the source to the observer [6,7]. To see this, the point x_0^a is taken to represent a point on the observer's world line. By treating $(\zeta, \bar{\zeta})$ as the celestial sphere of the observer, x^* as a measure of ''distance'' backwards along the null geodesic to a source and x^i as representing the ''time'' and ''angular position'' of the source at the moment of emission, we have constructed a lens equation together with the transit time equation. Thus one can see that, in a completely general context, our version of Fermat's principle, Eqs. (2a), (2b), leads immediately to the construction of the observer's past light cone, with associated lens and time of arrival equations.

When we extremize the function T we could obtain a maximum, a minimum or a saddle point. We wish to know how to rephrase the conditions for maximum, minimum or saddle in terms of our implicit Fermat potential, G. In principle, we need to look at the eigenvalues of the matrix containing the second derivatives of T evaluated at the extrema. We now switch to real variables (u,v) instead of $(\zeta, \bar{\zeta})$, via $\zeta \equiv u + iv$, and calculate the eigenvalues of the matrix

$$\delta^2 T \equiv \begin{pmatrix} T_{,uu} & T_{,uv} \\ T_{,uv} & T_{,vv} \end{pmatrix}. \tag{17}$$

Because of T being implicitly defined via Eq. (6), we just need to take second derivatives of Eq. (6). For instance, if we take two u derivatives of Eq. (6) [with Eq. (7) for t] we obtain

$$G_{,uu} + T_{,u}(2G_{,ut} + G_{,tt}T_{,u}) + G_{,t}T_{,uu} = 0. \tag{18}$$

Since we are interested only in the value of $T_{,uu}$ at the extremum, then $T_{,u} = 0$ and

$$T_{,uu} = -\frac{G_{,uu}}{G_{,t}}, \tag{19}$$

which presents no difficulty, since by assumption $G_{,t}$ must be nonvanishing. Likewise, the other components of the matrix come out proportional to the second derivatives of G:

$$\delta^2 T = -\frac{1}{G_{,t}} \begin{pmatrix} G_{,uu} & G_{,uv} \\ G_{,uv} & G_{,vv} \end{pmatrix}. \tag{20}$$

Thus, up to a factor of $-1/G_{,t}$ the eigenvalues are the same as those of $\delta^2 G$. In order to fix unnecessary sign freedom (since $-G$ for our purposes is just as good as G), we assume that $G_{,t} < 0$ at the extremum in question. It remains to rephrase our conditions for maximum, minimum and saddle in terms of the complex variable ζ. Since $\zeta = u + iv$, we have $\partial/\partial u = \partial/\partial\zeta + \partial/\partial\bar{\zeta}$ and $\partial/\partial v = i(\partial/\partial\zeta - \partial/\partial\bar{\zeta})$, thus

$$G_{,uu} = G_{,\zeta\zeta} + G_{,\bar{\zeta}\bar{\zeta}} + 2G_{,\zeta\bar{\zeta}}, \tag{21a}$$

$$G_{,vv} = -G_{,\zeta\zeta} - G_{,\bar{\zeta}\bar{\zeta}} + 2G_{,\zeta\bar{\zeta}}, \tag{21b}$$

$$G_{,uv} = i(G_{,\zeta\zeta} - G_{,\bar{\zeta}\bar{\zeta}}). \tag{21c}$$

If we think of G as spin-weight zero function on the sphere then, using the envelope condition, $G_{,\zeta} = G_{,\bar{\zeta}} = 0$ we have

$$\eth^2 G = (1 + \zeta\bar{\zeta})^2 G_{,\zeta\zeta}, \tag{22a}$$

$$\bar{\eth}^2 G = (1 + \zeta\bar{\zeta})^2 G_{,\bar{\zeta}\bar{\zeta}}, \tag{22b}$$

and hence

$$\delta^2 T = -\frac{2}{(1+\zeta\bar{\zeta})^2 G_{,t}} \times \begin{pmatrix} \eth\bar{\eth}G + \mathrm{Re}(\eth^2 G) & -\mathrm{Im}(\eth^2 G) \\ -\mathrm{Im}(\eth^2 G) & \eth\bar{\eth}G - \mathrm{Re}(\eth^2 G) \end{pmatrix}. \tag{23}$$

The eigenvalue equation for $\delta^2 T$ is given by

$$(\tilde{\lambda} - \eth\bar{\eth}G)^2 - \eth^2 G \bar{\eth}^2 G = 0, \tag{24}$$

where $\tilde{\lambda} = -(1+\zeta\bar{\zeta})^2 G_{,t}\lambda/2$ and λ is an eigenvalue of $\delta^2 T$. Since $G_{,t}$ is assumed to be negative, then the sign of λ is the same as the sign of $\tilde{\lambda}$. The solutions are

$$\tilde{\lambda}_\pm = \eth\bar{\eth}G \pm |\eth^2 G|. \tag{25}$$

Finally, the conditions are: If $|\eth\bar{\eth}G| < |\eth^2 G|$, the extremum is a saddle, since the eigenvalues have opposite signs. If $\eth\bar{\eth}G > |\eth^2 G|$, the extremum is a minimum, since both eigenvalues are positive. If $\eth\bar{\eth}G < -|\eth^2 G|$, the extremum is a maximum, since both eigenvalues are negative.

On the other hand, the envelope develops singularities when

$$\eth\bar{\eth}G^2 = \eth^2 G \bar{\eth}^2 G \tag{26}$$

as anticipated earlier (the vanishing of the determinant of the Jacobian matrix). The type of extremum is of fundamental importance. For instance, in a very weak field, no singularities develop in the observer's light cone, there is no multiplicity of images and the lens equation represents a local

minimum of the Fermat potential. Increasing the strength of the field, in a situation where there are three images, one will be a minimum, one a maximum and the remaining one will be a saddle. The minimum corresponds to the primary image, produced by a ray that does not encounter a caustic in its path. The saddle corresponds to a ray that passes the caustic once. The maximum yields the ray that passes the caustic twice and yields the faintest image. The generalization of this interpretation to cases of more than three images is complicated [1,3] and lies beyond our present interest.

A. Parametric version of the implicit Fermat potential

We finish this section by presenting a very important method for describing implicit Fermat potentials. In many cases it is difficult or impossible (as we will see) to obtain a closed-form G for a family of surfaces $G(x^a,x_0^a,\zeta,\overline{\zeta})=0$, especially if the surfaces self-intersect and have singularities. In such cases, it is simpler to describe the family of surfaces in parametric form adapted to the null geodesics ruling the surface, namely, by specifying the map

$$x^a=\Gamma^a(s,r,q,\zeta,\overline{\zeta}) \tag{27}$$

where $(s,r,q)\equiv s^i$ are coordinates on the surface, given for each fixed value of $[\zeta,\overline{\zeta}]$. Since each member of the family [for fixed $(\zeta,\overline{\zeta})$] is a null surface, it is automatically ruled by null geodesics (except at singular points of the surface). Therefore the surface coordinates s^i can always be chosen as geodesic coordinates: (r,q) labeling the null geodesics ruling the surface, and s as an affine parameter along the geodesics. However, our interest is in surfaces that intersect at a given point x_0^a. Therefore, all the surfaces in the family must contain the common point x_0^a. We can make use of this fact by adapting the coordinates (s,r,q) to this point. Since s is a parameter along the geodesics, we fix it by demanding that the 2-surface $s=0$ be transverse (to the null geodesics in the null surface) and contain the point x_0^a. This is, then, the initial-data surface for the geodesics in the null surface. On this initial-data 2-surface, the origin of coordinates for r and q can be taken to be x_0^a. In this manner, all the null surfaces in the family will single out the observer's point as the point with $s=r=q=0$, namely

$$x_0^a=\Gamma^a(0,0,0,\zeta,\overline{\zeta}) \tag{28}$$

so that Γ^a depends on x_0^a, i.e.,

$$x^a=\Gamma^a(s,r,q,x_0^a,\zeta,\overline{\zeta}). \tag{29}$$

Often we take x_0^a as the observation time T_0 at the spatial origin and omit it from the equations.

In the case that the family of surfaces $G(x^a,x_0^a,\zeta,\overline{\zeta})=0$ is given parametrically via Eq. (27), the envelope construction proceeds in an implicit manner as follows. In the first place, the ruling of the null surface via null geodesics guarantees that the Jacobian matrix of the map (27), namely $\partial\Gamma^a/\partial s^i$, has rank three (except at singular points). Therefore, locally, from the equations (27) one can always choose 3 equations that can be solved for s^i. The fourth equation remains a function of s^i. So Eq. (27) can always be viewed in the form

$$\Gamma^i(s^i,\zeta,\overline{\zeta};x_0^a)-x^i=0, \tag{30a}$$

$$\Gamma(s^i,\zeta,\overline{\zeta};x_0^a)-x^4=0, \tag{30b}$$

$$\Gamma\equiv\Gamma^4 \tag{30c}$$

where Eq. (30b) is interpreted as the equation $G(x^a,x_0^a,\zeta,\overline{\zeta})=0$ if s^i is thought of as implicitly given in terms of x^i via Eq. (30a). In other words,

$$G(x^a,x_0^a,\zeta,\overline{\zeta})\equiv\Gamma(s^i(x^i,\zeta,\overline{\zeta};x_0^a),\zeta,\overline{\zeta};x_0^a)-x^4. \tag{31}$$

In the envelope construction we then have

$$0=\partial_\zeta G|_{x^a}=\partial_\zeta\Gamma|_{s^i}+\frac{\partial\Gamma}{\partial s^i}|_\zeta\partial_\zeta s^i|_{x^i}, \tag{32}$$

where $\partial_\zeta s^i|_{x^i}$ is determined by implicit differentiation of Eq. (30a). Both Eqs. (31) and (32) will be extensively used in the following sections.

This appears to be a convenient method for constructing implicit Fermat potentials and for obtaining lens equations, as demonstrated in the following sections.

III. THE IMPLICIT FERMAT POTENTIAL IN MINKOWSKI SPACETIME

The case of Minkowski spacetime is quite trivial but is helpful as an illustration of the main ideas underlying the concept of the implicit Fermat potential G. In Cartesian coordinates x^a, null surfaces in Minkowski spacetime are level surfaces of the solutions to the eikonal equation $\eta^{ab}G_{,a}G_{,b}=0$, with $\eta_{ab}=\mathrm{diag}(1,-1,-1,-1)$. There is a 2-parameter family of distinct foliations by null planes with the two parameters corresponding to a sphere's worth of null directions at any given spacetime point. Using complex stereographic coordinates, the null directions can be specified by $l^a(\zeta,\overline{\zeta})$ given by

$$l^a(\zeta,\overline{\zeta})\equiv\frac{1}{\sqrt{2}}\left(1,\frac{(\zeta+\overline{\zeta})}{1+\zeta\overline{\zeta}},-\frac{i(\zeta-\overline{\zeta})}{1+\zeta\overline{\zeta}},-\frac{1-\zeta\overline{\zeta}}{1+\zeta\overline{\zeta}}\right). \tag{33}$$

The null planes with null normal l^a containing the observer's location $x_0^a(T_0)$ are given by

$$G(x^a,x_0^a,\zeta,\overline{\zeta})\equiv(x^a-x_0^a)l_a(\zeta,\overline{\zeta})=0. \tag{34}$$

The envelope of this 2-parameter family of null planes is the light cone of the point x_0^a and is obtained by taking the partial derivatives with respect to ζ and $\overline{\zeta}$ of the function $G(x^a,x_0^a,\zeta,\overline{\zeta})$ and setting these derivatives equal to zero. In this case,

$$\frac{\partial G}{\partial\zeta}=(x^a-x_0^a)\frac{\partial l_a}{\partial\zeta}=0, \tag{35a}$$

$$\frac{\partial G}{\partial \bar{\zeta}} = (x^a - x_0^a)\frac{\partial l_a}{\partial \bar{\zeta}} = 0. \tag{35b}$$

Equations (34) and (35a), (35b) define a 3-surface in the spacetime. One way to obtain a single equation defining such a surface is to solve Eqs. (35a), (35b) for ζ as a function of x^a, wherever possible, and substitute $\zeta(x^a)$ back into Eq. (34). Then the envelope is thus given, in principle, by

$$G_{env}(x^a) \equiv G(x^a, x_0^a, \zeta(x^a), \bar{\zeta}(x^a)) = 0. \tag{36}$$

In this case, this procedure can be carried out in closed form. Using the notation $\Delta x^a \equiv x^a - x_0^a$, Eq. (35b) is equivalent to

$$(\Delta x - i\Delta y)\zeta^2 - 2\Delta z\zeta - \Delta x - i\Delta y = 0 \tag{37}$$

with solution

$$\zeta(x^a) = \frac{\Delta z \pm \Delta r}{\Delta x - i\Delta y}, \tag{38a}$$

$$\Delta r \equiv \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}. \tag{38b}$$

[The other equation, Eq. (35a), is the complex conjugate of Eq. (35b) and yields no new information.] Substituting $\zeta(x^a)$ given by Eq. (38a) and its complex conjugate $\bar{\zeta}(x^a)$ into Eq. (34) yields

$$G_{env}(x^a, x_0^a) = G(x^a, x_0^a, \zeta(x^a), \bar{\zeta}(x^a))$$

$$= \frac{1}{\sqrt{2}}(\Delta t \mp \Delta r) = 0, \tag{39}$$

where the minus sign is for the future light cone and the plus sign is for the past light cone of the point x_0^a. Consequently, the set of three equations (37) and (34) are equivalent to the lens and time of arrival equations. Because there is no time occurrence in Eq. (37), then one can think of it as the lens equation since it relates the spatial position of source (x,y,z) to its corresponding image on the celestial sphere $(\zeta, \bar{\zeta})$. In principle, solving for two of (x,y,z) in terms of $(\zeta, \bar{\zeta})$ and substituting into Eq. (34) turns Eq. (34) into an equation relating the time of emission t from a source at (x,y,z) with image direction $(\zeta, \bar{\zeta})$ seen at the observer x_0^a.

IV. THE IMPLICIT FERMAT POTENTIAL FOR A STATIC ASYMPTOTICALLY FLAT THIN-LENS SPACETIME

In geometrical terms, a thin lens can be modeled as a spacetime with vanishing curvature everywhere except at a timelike surface, where the ''lens'' lives. Accordingly, the spacetime consists of two flat spacetimes, matched appropriately at the lens' worldsheet, with the observer lying in one flat spacetime and the source lying in the other one. We refer to these two flat spacetimes as the observer's side and the source's side of the spacetime.

It is clear that the families of null surfaces to be used on the observer's side of the spacetime are all the null planes that contain the observer's location, labeled by the direction of the null normal. On the source's side, however, the appropriate null surfaces to use are not null planes but the distortion of the null planes by the lens. For this reason they must, in general, be given in parametric form. We use the parametric method, via Eq. (27), explained in Sec. II, for the envelope construction.

For the purpose of clarity of presentation, we first develop our construction of the null surfaces in a lower-dimensional setting. We then generalize the argument to $3+1$ dimensions.

A. 2+1 static lens planes

The $2+1$ lensing setting is described in Cartesian coordinates $x^a = (t,x,y)$. The observer is at $x_0^a = (T_0,0,0)$, and the lens surface is a plane located at $y = y_l$. The metric of the spacetime is flat everywhere except at the lens plane:

$$ds^2 = c^2dt^2 - dx^2 - dy^2. \tag{40}$$

Our aim is to construct a 1-parameter family of null foliations which in principle can be given as the level surfaces of a function $G(x^a, \varphi)$ which replaces the $G(x^a, \zeta, \bar{\zeta})$ of the four dimensional lensing scenario. (The origin, x_0^a, is omitted.) This function will be entirely equivalent to the Fermat potential in this lensing scenario, and the envelope of the surfaces $G(x^a, \varphi) = 0$ will simultaneously give us the lens equation and time of arrival. In the approximation of small angles, the Fermat potential and the lens equation will reduce to the standard astrophysical lensing scenario.

Roughly our construction is as follows: we define a 1-parameter family of null surfaces (parametrized by φ) by imagining a straight line of photons traveling (backwards in time) parallel to each other (null rays) in a direction making an angle φ with the y axis. As the photons move, they trace a null plane in the spacetime. The moment each photon arrives at the lens plane, it is ''detained'' for a length of time $\mathcal{T}$ that depends only on the point at which the photon hits the lens plane. After this time, the photon is released in a new direction $\varphi' = \varphi - \alpha$, deviating by an angle α from its original direction. The bending angle α depends only on the point at which the photon hits the lens plane, and it should be entirely determined by the ''detention time'' $\mathcal{T}$. Because different photons arrive at the lens plane at different times and are detained for different lengths of time, the wave front they make up after leaving the lens plane is not necessarily plane.

The scheme that we have just described has a corresponding analog in the wave front distortion produced by a thin sheet of glass with a variable index of refraction. A light ray impinging nearly normally on the sheet will bend inside the glass due to the component of the gradient of the index of refraction perpendicular to its path, exiting the glass with a finite deviation angle. A plane incident wave front then comes out of the glass distorted. The local index of refraction at the point of incidence multiplied by the thickness of the glass is a measure of the time spent by each light ray inside the glass, or equivalently, the ''detention'' time of light inside

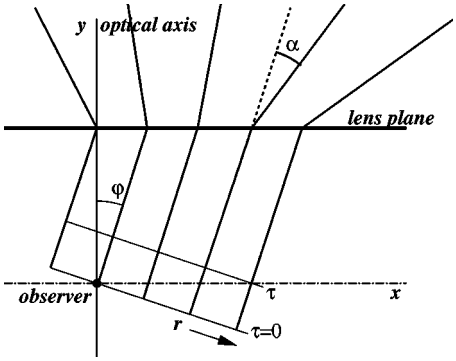

FIG. 1. Parametric construction of the null surfaces. On the observer side, the surface is a null plane ruled by parallel light rays making an angle φ with the y axis. The light rays are labeled by their distance r away from the observer, and are parametrized by the distance $c\tau$ away from the wave front that passes by the observer. On the other side of the lens, the light rays bend through an unspecified angle α that depends on the position at which the light rays reach the lens plane.

the glass. For later purposes, it is important that we make the following point with respect to the limit in which the thickness of the glass tends to zero. The ''detention time'' in the region occupied by the glass defined in this way has two contributions: the transit time in vacuum through the region [the ''vacuum'' time: $c^{-1}\times$ (thickness of the glass)] plus the amount of time that light is delayed by the slowing down in the presence of the medium, with respect to vacuum. In the limit in which the thickness of the glass tends to zero, the first (vacuum) term vanishes and the ''detention time'' coincides with the time delay with respect to vacuum.

In the case of interest to us, the detention time $\mathcal{T}$ is determined by the gravitational field of a deflector on the lens plane and it coincides with the gravitational time delay with respect to vacuum:

$$\mathcal{T}=\frac{-2}{c^3}\int U dl \tag{41}$$

where U is the Newtonian potential of an isolated mass distribution and the integration takes place along the null ray or line of sight. The situation we are contemplating is the limiting case in which the support of U tends to zero, while the integral of U remains constant, namely, U is a distribution along each null ray.

More precisely, *on the observer side* of the spacetime we consider the parallel light rays traced by a plane wave front, moving (backwards in time) at an angle φ with the optical axis (the y axis) as defining our null surface. See Fig. 1. The wave front passes by the observer at time T_0. The (perpendicular) distance from the observer to the wave front at later times divided by c, which is denoted by τ, can be used as a parameter along each light ray. The individual light rays at $\tau=0$, can be labeled by their distance r from the observer. In terms of these parameters, the time coordinate of the light rays is given by

$$t=T_0-\tau, \tag{42}$$

whereas the space coordinates of the light rays are given by

$$x=c\tau\sin\varphi+r\cos\varphi, \tag{43a}$$

$$y=c\tau\cos\varphi-r\sin\varphi. \tag{43b}$$

(Note that as τ increases time is going backwards.) It is clear that $r=\tau=0$ defines the observer at $(x,y)=(0,0)$. Each light ray (for fixed φ) reaches the lens plane at $\tau=\tau_l$ when $y=y_l$, hitting a lens point $x=x_l$:

$$x_l=c\tau_l\sin\varphi+r\cos\varphi, \tag{44a}$$

$$y_l=c\tau_l\cos\varphi-r\sin\varphi. \tag{44b}$$

These, in turn, can be solved for τ_l and r and yield

$$c\tau_l=x_l\sin\varphi+y_l\cos\varphi, \tag{45a}$$

$$r=x_l\cos\varphi-y_l\sin\varphi. \tag{45b}$$

Equations (43a), (43b) hold while $\tau<\tau_l$.

On the other side of the lens line, the wave fronts will be determined by two things: the bending angle $\alpha(x_l)$ and the gravitational time delay imposed on the light rays at the lens line, which we here denote by $\mathcal{T}(x_l)$. The bending angle is to be determined in terms of $\mathcal{T}(x_l)$. The bending angle and the time delay can be incorporated into the parametric equations for these light rays in the following way:

$$x=c(\tau-\tau_l-\mathcal{T})\sin(\varphi-\alpha)+x_l, \tag{46a}$$

$$y=c(\tau-\tau_l-\mathcal{T})\cos(\varphi-\alpha)+y_l. \tag{46b}$$

These equations hold for $\tau>\tau_l+\mathcal{T}$. For times in between, the light rays are ''delayed'' at the lens plane, so we have

$$x=x_l, \tag{47a}$$

$$y=y_l, \tag{47b}$$

for $\tau_l<\tau<\tau_l+\mathcal{T}$. Equations (43a), (43b), (46a), (46b) and (47a), (47b) represent the parametric expression of a one-parameter family of null surfaces which in principle can also be expressed by a single equation of the form $G(x,y,t,\varphi)=0$ away from caustics.

On the observer side this can be done explicitly by eliminating r from Eqs. (43a), (43b) yielding

$$G(x,y,\tau,\varphi)\equiv y\cos\varphi+x\tan\varphi-\tau=0. \tag{48}$$

On the source side, however, this cannot be done explicitly. The parameter that labels the null geodesics is x_l, as opposed to r on the observer side. They are related by Eq. (45b). The idea would be to eliminate x_l between Eqs. (46a) and (46b) but this is impossible since both $\mathcal{T}$ and α are functions of x_l. We thus need to define x_l as a function of x implicitly by Eq. (46a). We take Eq. (46b) as the equation that defines the function G where x_l is the implicit function of x given by Eq. (46a):

$$G(x,y,\tau,\varphi)=y-y_l-(c\tau-x_l\sin\varphi-y_l\cos\varphi-c\mathcal{T}(x_l))\times\cos(\varphi-\alpha(x_l))=0. \tag{49}$$

We have used Eq. (45a) to eliminate τ_l in terms of x_l and y_l.

The envelope is obtained by setting $\partial G(x,y,t,\varphi)/\partial\varphi=0$, where the partial derivative is taken at constant values of (x,y,t). On the observer side, i.e., for $y<y_l$, taking $\partial/\partial\varphi$ on Eq. (48) we obtain

$$-y\sin\varphi+x\cos\varphi=0 \tag{50}$$

or

$$\tan\varphi=\frac{x}{y}, \tag{51}$$

an equation for φ as a function of x^a on the observer's side. Substituting Eq. (51) back into Eq. (48) we have (for $y<y_l$),

$$\sqrt{x^2+y^2}+c(t-T_0)=0 \tag{52}$$

which is the observer's past light cone, as expected. Notice that when Eq. (51) is evaluated at the lens plane it yields

$$\tan\varphi=\frac{x_l}{y_l}. \tag{53}$$

From Eq. (45b), we see that this implies that $r=0$, i.e., that these rays come from the observer. In other words, of all the light rays at angle φ, the envelope condition picks the null ray that passes through the observer. This light ray hits the lens plane at $x=x_l$ and subsequently bends through an as yet undetermined angle $\alpha(x_l)$. In order to determine the bending angle α we use the envelope condition on the source's side, imposing on it the condition (53).

The envelope is obtained by setting to zero the implicit differentiation of $G(x,y,\tau,\varphi)$ from Eq. (49) with respect to φ keeping x^a fixed. We obtain

$$0=\sin(\varphi-\alpha)\left[\frac{\partial x_l}{\partial\varphi}\frac{\partial\alpha}{\partial x_l}-1\right]\times(cT_0-ct-x_l\sin\varphi-y_l\cos\varphi-c\mathcal{T})+\cos(\varphi-\alpha)\times\left[y_l\sin\varphi-x_l\cos\varphi-\frac{\partial x_l}{\partial\varphi}\left(c\frac{\partial\mathcal{T}}{\partial x_l}+\sin\varphi\right)\right]. \tag{54}$$

The quantity $\partial x_l/\partial\varphi$ is calculated by implicit differentiation of Eq. (46a) with respect to φ, keeping x^a fixed:

$$\frac{\partial x_l}{\partial\varphi}=-\frac{A}{B}, \tag{55}$$

with

$$A=\cos(\varphi-\alpha)(cT_0-ct-x_l\sin\varphi-y_l\cos\varphi-c\mathcal{T})+\sin(\varphi-\alpha)(y_l\sin\varphi-x_l\cos\varphi) \tag{56a}$$

$$B=1-\cos(\varphi-\alpha)(cT_0-ct-x_l\sin\varphi-y_l\cos\varphi-c\mathcal{T})\frac{\partial\alpha}{\partial x_l}-\sin(\varphi-\alpha)\left(c\frac{\partial\mathcal{T}}{\partial x_l}+\sin\varphi\right). \tag{56b}$$

From Eq. (55), and some algebra, Eq. (54) becomes

$$0=-\sin(\varphi-\alpha)+c\frac{\partial\mathcal{T}}{\partial x_l}+\sin\varphi+(y_l\sin\varphi-x_l\cos\varphi)\times\left[\frac{\cos(\varphi-\alpha)}{(cT_0-ct-x_l\sin\varphi-y_l\cos\varphi-c\mathcal{T})}-\frac{\partial\alpha}{\partial x_l}\right]. \tag{57}$$

Using Eq. (53), the first parenthesis in the last term of Eq. (57) vanishes, so that Eq. (57) reduces to

$$0=\sin\varphi-\sin(\varphi-\alpha)+c\frac{\partial\mathcal{T}}{\partial x_l}. \tag{58}$$

This is the basic relationship between the bending angle and the gravitational time delay that holds for large as well as small angles. Using the regime of small angles, Eq. (58) reduces to

$$\alpha=-c\frac{\partial\mathcal{T}}{\partial x_l}, \tag{59}$$

the standard relationship in the astrophysical approach to lensing in a two-dimensional scenario.

The standard lens equation is obtained in the following manner. Elimination of $\tau-\tau_l-\mathcal{T}$ in Eqs. (46a), (46b) implies

$$(x-x_l)\cos(\varphi-\alpha)-(y-y_l)\sin(\varphi-\alpha)=0. \tag{60}$$

With small angles, and with $y_l=D_l$ and $y=D_s$ (the source distance), this becomes

$$x-x_l-(D_s-D_l)(\varphi-\alpha)=0. \tag{61}$$

From the envelope condition, Eq. (53), with the small-angle approximation ($\varphi=\tan\varphi=x_l/D_l$), we finally have

$$\frac{x}{D_s}=\frac{x_l}{D_l}-\frac{(D_s-D_l)}{D_s}\alpha, \tag{62}$$

the standard lens equation, which is understood as a map from x_l on the lens line to x on the source line (i.e., $x_l\to x$), with fixed values of D_l and D_s.

As an illustration, Fig. 2 shows how the envelope construction works for the case of a bending angle that does not depend on the point on the lens plane: $\alpha(x_l)=\text{const}$ and $c\mathcal{T}=-\alpha x_l$. Panel (a) shows a spacetime picture of one member of the 1-parameter family of null surfaces that is used to construct the envelope. In this case, two null planes in the directions φ and $\varphi-\alpha$ are matched at the lens' worldsheet. The vertical separation between the intersections of the two null planes with the lens' worldsheet represents the time delay $\mathcal{T}$. A spatial projection of the resulting null surface shows the bending of light rays, whereas the time delay is observed in the shift of the wave fronts in passing through

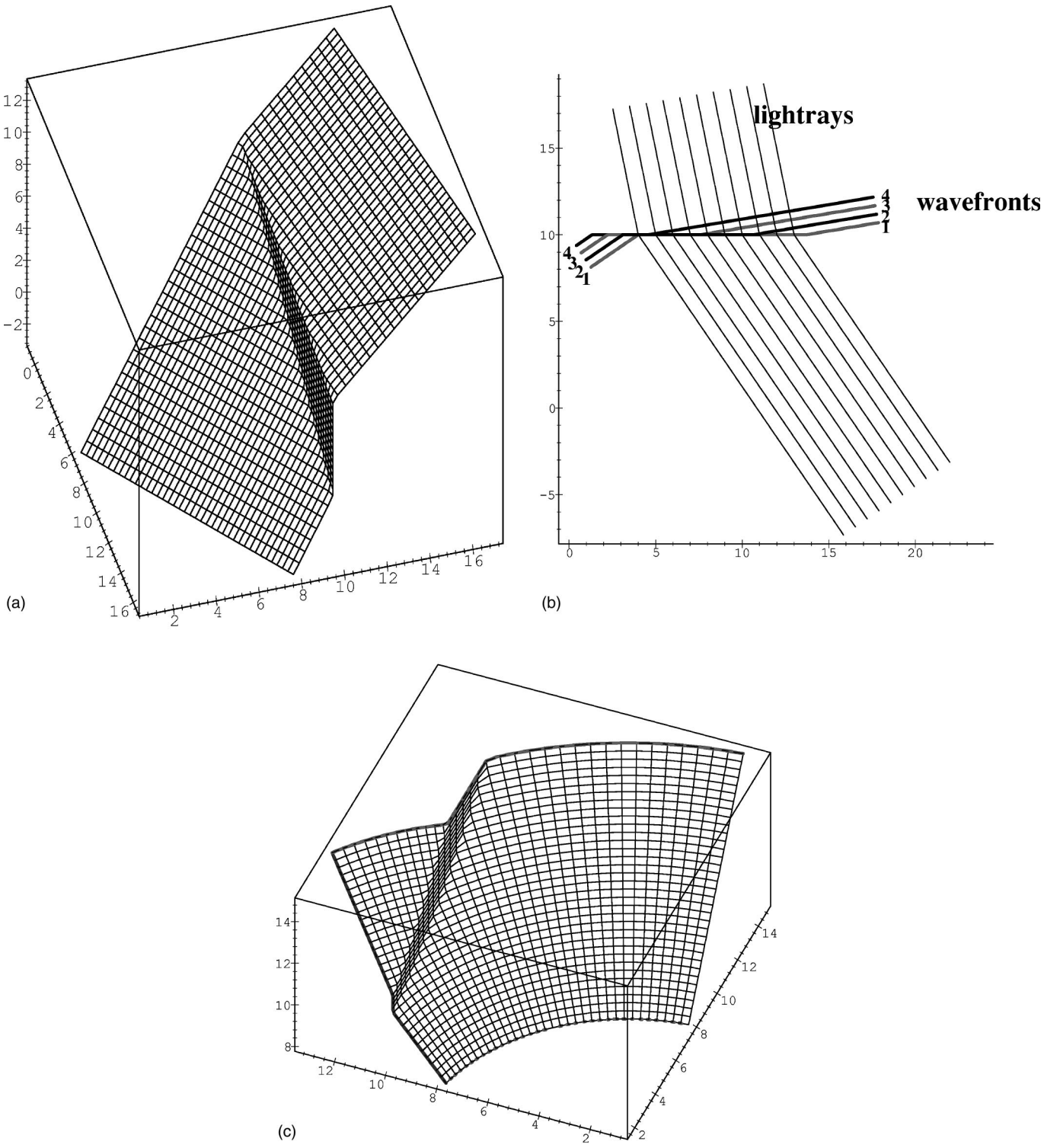

FIG. 2. Envelope construction for the case of a bending angle independent of the location on the lens line x_l. Panel (a) shows a spacetime picture of one member of the 1-parameter family of null surfaces that is used to construct the envelope. Panel (b) is a space picture of the null surface, where both the light rays and the associated wave fronts are shown. Panel (c) shows a spacetime picture of the envelope of the 1-parameter family. The family is labeled by the direction of the plane wave that hits the lens line.

the lens line. The envelope of the φ-dependent family of such null surfaces yields the light cone of the observer on the observer's side of the lens, which is matched to another light cone with a shifted origin on the other side of the lens.

A more physical example is provided by the use of a time delay of the form $\mathcal{T}(x_l)=\mathcal{T}_0(1-\ln[1+(x_l/x_0)^2])$. This reproduces qualitatively the time delay of a compact lens in the vicinity of the optical axis, since the delay is greater for smaller impact parameters. The spatial projection of the null surface obtained in this case is shown in the top panel of Fig.

3. The center panel of the figure shows the projection of the envelope. In this case, a caustic develops on the source's side of the observer's light cone. The caustic is traced by cusps in the wave fronts, as shown in the bottom panel of the figure.

The equation for the time delay $\Delta\tau$ (the difference in arrival time between a lensed light ray and an unlensed one) can be obtained from our approach by solving for τ in Eq. (46b) and subtracting from it the travel time in the absence of the lens:

$$\Delta\tau = \frac{y - y_l}{c\cos(\varphi - \alpha)} + \tau_l + \mathcal{T} - \frac{1}{c}\sqrt{x^2 + y^2}. \tag{63}$$

The equation for the time delay is usually given up to second order in the angles. We need expansions of x_l and τ_l, from Eqs. (43a), (43b) with $r=0$, up to second-order terms:

$$\tau_l = \frac{y_l}{c}\left[1 + \frac{\varphi^2}{2}\right] + O(\varphi^3), \tag{64a}$$

$$x_l = y_l\varphi + O(\varphi^3). \tag{64b}$$

These, as well as Eq. (62) for x, are used in Eq. (63). We obtain

$$\Delta\tau = \mathcal{T} + \frac{D_l D_{ls}}{2cD_s}\alpha^2, \tag{65}$$

which agrees with the usual expression for the time delay.

B. 3+1 static lens planes

The generalization of our scheme for the thin-lens spacetimes to 3+1 dimensions is straightforward. We consider a spacetime with coordinates (t,x,y,z) consisting of two Minkowski spacetimes matched together at the lens plane, which in this case is the timelike 3-surface $z=z_l$ and is parametrized by (t,x_l,y_l). A generic null direction is l^a given in terms of stereographic coordinates as in Eq. (33). For our present purposes, we relabel the components of l^a as

$$l^a = \frac{1}{\sqrt{2}}(1, p_x, p_y, p_z) \tag{66}$$

where (p_x,p_y) label the sphere of null directions and $p_z \equiv \sqrt{1-p_x^2-p_y^2}$. As in the (2+1) case, we choose our family of null surfaces as null planes on the observer's side of the spacetime, and write it parametrically as

$$\vec{r} = c\tau\vec{p} + \vec{s} \tag{67}$$

with $\vec{r}\equiv(x,y,z)$, $\vec{p}\equiv(p_x,p_y,p_z)$ and $\vec{s}=(s_x,s_y,s_z)$, with $\vec{s}\cdot\vec{p}=0$, representing the null surface at $\tau=0$. The values of $\vec{s}$ label the null geodesics of the $\vec{p}$ null plane.

Then, if τ_l represents the time that it takes for a light ray to reach the point $\vec{r}_l\equiv(x_l,y_l,z_l)$ measured from the time that the wave front passes the observer, we must have

$$\vec{r}_l = c\tau_l\vec{p} + \vec{s} \tag{68}$$

and thus

$$\tau_l = \frac{\vec{r}_l\cdot\vec{p}}{c}. \tag{69}$$

It is very useful to change the geodesic labels from $\vec{s}$ to $\vec{r}_l$ by means of

$$\vec{s} = \vec{r}_l - (\vec{r}_l\cdot\vec{p})\vec{p} \tag{70}$$

for use on both the observer and source side of the spacetime. Equation (67) becomes

$$\vec{r} = (c\tau - (\vec{r}_l\cdot\vec{p}))\vec{p} + \vec{r}_l. \tag{71}$$

We thus have in this picture that the wave front, in direction $\vec{p}$, consists of null geodesics parallel to each other, parametrized by the values of (x_l,y_l) where they hit the lens plane at z_l. Once a ray reaches the lens plane, it is delayed for a time $\mathcal{T}(x_l,y_l)$, during which it just "sits" at the lens plane, before leaving in a different direction $\vec{p}\,'$. Thus, in our model, we have

$$\vec{r} = \vec{r}_l \tag{72}$$

and then

$$\vec{r} = c(T_0 - t - \tau_l - \mathcal{T})\vec{p}\,' + \vec{r}_l \tag{73}$$

where

$$\vec{p}\,' = \vec{p} - \vec{\alpha}(x_l,y_l). \tag{74}$$

Here we have defined $\vec{\alpha}(x_l,y_l)$ in the usual manner, representing the deviation of the null geodesic, which depends on the point at which the geodesic hits the lens plane. Equations (71), (72), and (73) constitute a parametric expression of $G(t,x,y,z,p_x,p_y)=0$. In this approach, the time delay $\mathcal{T}(x_l,y_l)$ is assumed to be prescribed, whereas $\vec{\alpha}(x_l,y_l)$ is determined by the envelope condition. The envelope condition in this case takes the form

$$\frac{\partial G}{\partial p_x} = \frac{\partial G}{\partial p_y} = 0 \tag{75}$$

which is carried out implicitly as in the (2+1) case, and which yields

$$c\frac{\partial\mathcal{T}}{\partial x_l} = -\alpha_x(x_l,y_l), \tag{76a}$$

$$c\frac{\partial\mathcal{T}}{\partial y_l} = -\alpha_y(x_l,y_l), \tag{76b}$$

thus reproducing the relationship between the time delay and the bending angle in the more standard astrophysical approach to lensing. The detailed calculations, which can be omitted, are reproduced below.

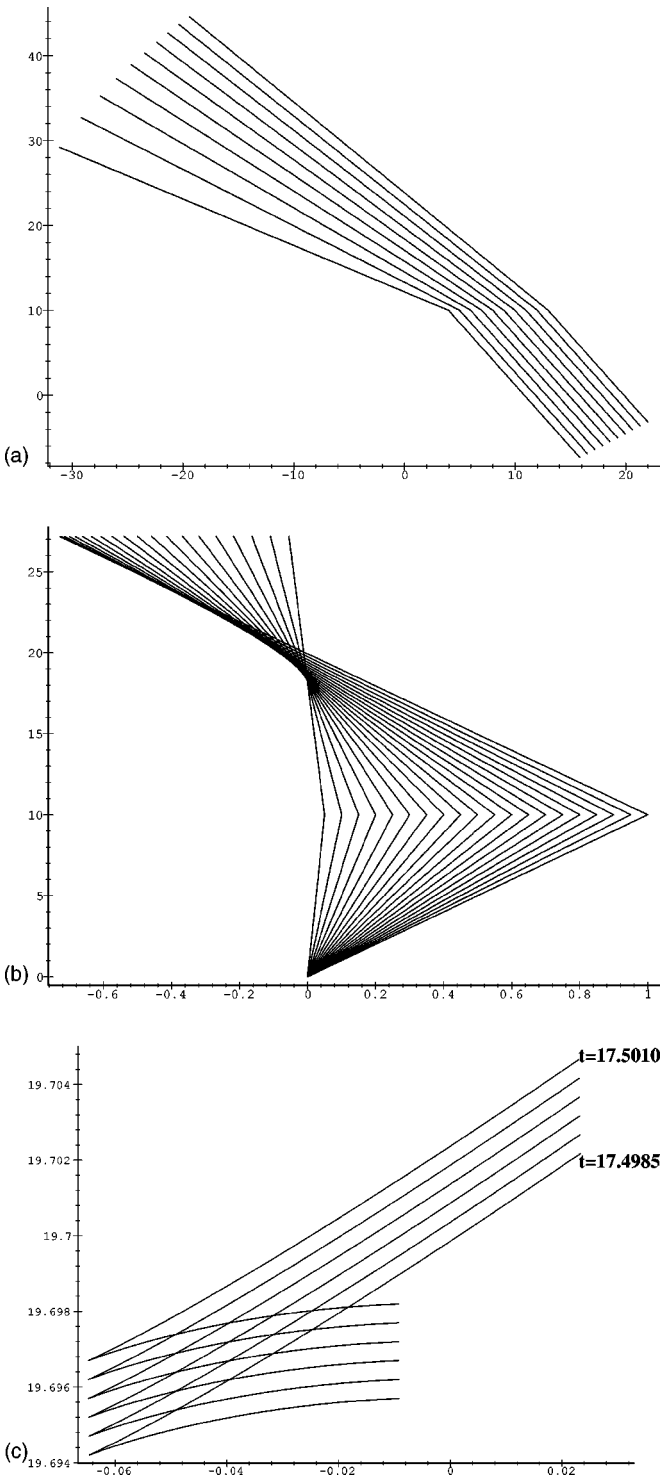

FIG. 3. Envelope construction for the case of a time delay $T(x_l)=T_0-\ln(1+(x_l/x_0)^2)$. The top panel shows a space projection of one member of the family of null surfaces. The null surfaces are labeled by the direction of the plane that hits the lens line. The center panel contains a space picture of the envelope showing the light rays. Notice the caustic line that develops after the light rays pass through the lens line. The bottom panel shows a greatly magnified view of a few of the wave fronts associated with the light rays of the center panel. Notice the cusp in each wave front. These wave fronts occur around the value 20 in the vertical axis of the center panel.

The implicit G function on the observer's side is given by the z component of Eq. (71)

$$G(\tau,x,y,z,p_x,p_y)=z-z_l-(\tau-\tau_l)p_z=0, \tag{77}$$

while x_l and y_l are defined implicitly by the (x,y) components of Eq. (71). Solving for x_l and y_l from these equations and substituting back into G yields

$$G=\frac{1}{p_z}(-t+xp_x+yp_y+zp_z)=0. \tag{78}$$

On the source side we implicitly define the thin-lens G function by

$$G=z-z_l-(\tau-\tau_l-\mathcal{T})p_z'=0, \tag{79}$$

where x_l and y_l are implicitly defined by the (x,y) components of Eq. (73). The envelope of G is obtained by setting to zero the partial derivatives of G with respect to p_x and p_y, holding (t,x,y,z) fixed. Applying this prescription to the G function before the lens plane, we have

$$\frac{\partial G}{\partial p_x}=0\Rightarrow x-\frac{p_x}{p_z}z=0, \tag{80a}$$

$$\frac{\partial G}{\partial p_y}=0\Rightarrow y-\frac{p_y}{p_z}z=0. \tag{80b}$$

We note that Eqs. (80a), (80b) imply that

$$\frac{x_l}{p_x}=\frac{y_l}{p_y}=\frac{z_l}{p_z}. \tag{81}$$

Solving for p_x and p_y from Eqs. (80a), (80b) and substituting into the G function, we obtain our final expression for the thin-lens G function on the observer's side

$$p_zG=-t+xp_x+yp_y+zp_z=-t+\frac{p_z}{z}(x^2+y^2+z^2)$$

$$=-t\pm\sqrt{x^2+y^2+z^2}=0, \tag{82}$$

the observer's light cone.

On the source's side, the envelope construction requires computing the following partial derivatives: $\partial x_l/\partial p_x, \partial y_l/\partial p_x, \partial x_l/\partial p_y$ and $\partial y_l/\partial p_y$ from the (x,y) components of Eq. (73). Since $\vec{p}\,'=\vec{p}-\vec{\alpha}(x_l,y_l)$, we need the derivatives of $\vec{p}\,'$.

We define

$$\frac{\delta p_x'}{\delta p_x}\equiv\frac{\partial p_x'}{\partial p_x}+\frac{\partial p_x'}{\partial x_l}\frac{\partial x_l}{\partial p_x}+\frac{\partial p_x'}{\partial y_l}\frac{\partial y_l}{\partial p_x}, \tag{83}$$

and similar quantities for all combinations of (p_x',p_y',p_x,p_y). After a lengthy calculation, we obtain

$$\frac{\partial x_l}{\partial p_x}=\frac{(\tau-\tau_l-\mathcal{T})\left\{-\frac{\delta p_x'}{\delta p_x}\left(1-\frac{\partial \mathcal{T}}{\partial y_l}p_y'-p_y'p_y\right)-\frac{\delta p_y'}{\delta p_x}\left(\frac{\partial \mathcal{T}}{\partial y_l}p_x'+p_x'p_y\right)\right\}}{1-p_x'\frac{\partial \mathcal{T}}{\partial x_l}-p_y'\frac{\partial \mathcal{T}}{\partial y_l}-p_xp_x'-p_yp_y'}, \tag{84a}$$

$$\frac{\partial y_l}{\partial p_x}=\frac{(\tau-\tau_l-\mathcal{T})\left\{-\frac{\delta p_y'}{\delta p_x}\left(1-\frac{\partial \mathcal{T}}{\partial x_l}p_x'-p_x'p_x\right)-\frac{\delta p_x'}{\delta p_x}\left(\frac{\partial \mathcal{T}}{\partial x_l}p_y'+p_y'p_x\right)\right\}}{1-p_x'\frac{\partial \mathcal{T}}{\partial x_l}-p_y'\frac{\partial \mathcal{T}}{\partial y_l}-p_xp_x'-p_yp_y'}, \tag{84b}$$

and similar expressions for $\partial x_l/\partial p_y$ and $\partial y_l/\partial p_y$.

The envelope construction consists in setting the p_x and p_y partial derivatives [holding (t,x,y,z) fixed] of Eq. (79) to zero. Explicitly, we have

$$0=-(\tau-\tau_l-\mathcal{T})\left(\frac{\partial p_z'}{\partial p_x'}\frac{\delta p_x'}{\delta p_x}+\frac{\partial p_z'}{\partial p_y'}\frac{\delta p_y'}{\delta p_x}\right)+p_z'\left(p_x\frac{\partial x_l}{\partial p_x}+p_y\frac{\partial y_l}{\partial p_x}+\frac{\partial \mathcal{T}}{\partial x_l}\frac{\partial x_l}{\partial p_x}+\frac{\partial \mathcal{T}}{\partial y_l}\frac{\partial y_l}{\partial p_x}\right), \tag{85a}$$

$$0=-(\tau-\tau_l-\mathcal{T})\left(\frac{\partial p_z'}{\partial p_x'}\frac{\delta p_x'}{\delta p_y}+\frac{\partial p_z'}{\partial p_y'}\frac{\delta p_y'}{\delta p_y}\right)+p_z'\left(p_x\frac{\partial x_l}{\partial p_y}+p_y\frac{\partial y_l}{\partial p_y}+\frac{\partial \mathcal{T}}{\partial x_l}\frac{\partial x_l}{\partial p_y}+\frac{\partial \mathcal{T}}{\partial y_l}\frac{\partial y_l}{\partial p_y}\right), \tag{85b}$$

where we have again used Eq. (81). In these two equations, we substitute the partial derivatives given above and solve for $\partial \mathcal{T}/\partial x_l$ and $\partial \mathcal{T}/\partial y_l$, obtaining

$$\frac{\partial \mathcal{T}}{\partial x_l}=p_x'-p_x=\alpha_x(x_l,y_l), \tag{86a}$$

$$\frac{\partial \mathcal{T}}{\partial y_l}=p_y'-p_y=\alpha_y(x_l,y_l). \tag{86b}$$

Equations (86a), (86b) determine the new direction of the light ray on the source's side with respect to the initial direction of the light ray on the observer's side (i.e., the bending angle) as the gradient of the time delay function. Note that we have not made use of the small angle approximation in these calculations.

V. MOVING LENS PLANES IN 2+1 DIMENSIONS

As an application of our approach, we consider a generalization of our (2+1) results to the case of a lens plane that is moving in the observer's frame of reference, but is otherwise unchanging. In other words, the deflector has a given configuration, which does not change with time, but moves rigidly along the line of sight. The idea is to obtain the bending angle and compare with the bending angle by the same deflector at rest.

Again our basic picture is a family of parallel null geodesics at an angle φ with the normal to the lens plane forming a null plane. The rays each arrive at the lens plane at a time τ_l, and then remain in the lens plane for a length of time $\mathcal{T}$ that depends only on the point at which they hit the now moving lens plane, before exiting the lens plane at an angle $\varphi'=\varphi-\alpha_m$.

We point out that the "detention time" $\mathcal{T}$ of a thin rigid lens in (slow) motion is the same as that of the same deflector at rest. The reason for this becomes clear if we consider the situation of a glass sheet of thickness Δ that is moving. In the rest frame of the glass, light has a speed u_0 and takes a time $t_0=\Delta/u_0$ between entering and exiting the glass. The time in the laboratory frame in which the glass is moving with speed v is $t'=t_0-v\Delta/c^2$ (neglecting higher powers in v/c) following from the fact that the length for light to travel changes due to the motion of the glass, and with the fact that the speed of light also changes if the glass is moving (by the Fizeau effect). Thus t' differs from t_0 only by a term proportional to the thickness of the glass, but completely unrelated to the glass properties. In the limit in which the thickness of the glass tends to zero while the rest-frame travel time, t_0, is kept constant, the difference vanishes.

As before, we assume that the observer lies at the origin of coordinates $x_0^a=(T_0,0,0)$. The lens plane, however, is moving, being described by $x_l^a=[t,x_l,y_l(t)]$. The (affine) parameter τ describing the evolution of the null geodesics of the null surface is such that the wave front $\tau=0$ passes by the observer at $t=T_0$. The parameter τ runs backwards in time, so that the time coordinate of the light rays is given by $t=T_0-\tau$ as before. We will slightly abuse our notation by using $y_l(\tau)$ to denote $y_l(t=T_0-\tau)$. If the individual light rays in the beam are labeled by r, with $r=0$ being the ray that passes through the origin, then we can give them parametrically as

$$x=r\cos\varphi+c\tau\sin\varphi, \tag{87a}$$

$$y=-r\sin\varphi+c\tau\cos\varphi. \tag{87b}$$

Individual light rays reach the lens plane at time τ_l, at which

$$x_l = r\cos\varphi + c\tau_l \sin\varphi \tag{88a}$$

$$y_l(\tau_l) = -r\sin\varphi + c\tau_l\cos\varphi. \tag{88b}$$

The rays remain at the lens plane for a time $\mathcal{T}(x_l)$ and then leave in a direction $\varphi - \alpha_m(x_l)$, where the bending angle for a moving lens, $\alpha_m(x_l)$, is to be determined. At the time the rays leave the lens plane, however, the lens plane has moved to a point $y_l(\tau_l+\mathcal{T})$, carrying the rays with it. On the other side of the lens, the source side, we have

$$x = c(\tau-\tau_l-\mathcal{T})\sin(\varphi-\alpha_m)+x_l, \tag{89a}$$

$$y = c(\tau-\tau_l-\mathcal{T})\cos(\varphi-\alpha_m)+y_l(\tau_l+\mathcal{T}). \tag{89b}$$

Equations (87a), (87b) hold when $\tau<\tau_l$, whereas Eqs. (89a), (89b) hold for $\tau>\tau_l+\mathcal{T}$. On the observer side, Eq. (87b) represents $G(x,y,t,\varphi)=0$ if r is thought of as a function of (t,x,φ) given implicitly by Eq. (87a). The envelope condition $\partial G/\partial\varphi=0$ on the observer side, implemented by taking $\partial/\partial\varphi$ of Eq. (87a), yields, as we had earlier, $r=0$ or

$$x_l = c\tau_l\sin\varphi, \tag{90a}$$

$$y_l(\tau_l) = c\tau_l\cos\varphi. \tag{90b}$$

As we had earlier, we consider Eq. (89b) as the equation $G(x,y,t,\varphi)=0$ on the source's side by interpreting x_l as a function of (τ,x,φ):

$$x_l = x_l(\tau,x,\varphi), \tag{91}$$

given implicitly by Eq. (89a). The implementation of the envelope condition, $\partial G/\partial\varphi=0$, involves several steps that must be explained.

First, we have to evaluate $dy_l/d\tau$, the velocity of the lens when the light first enters it, at $\tau_l+\mathcal{T}$, and when it exits the lens, at τ_l, on the observer side. Since the lens is thin and moving slowly, we consider the two values to be the same (or equal to an average) and call

$$\frac{dy_l}{d\tau} = -\hat{v}_l. \tag{92}$$

The second point is that, though Eq. (89b) contains x_l via α and $\mathcal{T}$, it also contains τ_l which is a function of both x_l and φ and is thus needed in the envelope condition. It is obtained by eliminating r from Eqs. (87a), (87b):

$$c\tau_l - x_l\sin\varphi - y_l(\tau_l) = 0. \tag{93}$$

Taking the φ derivative of Eq. (93) and using Eq. (92), with $r=0$, yields

$$\frac{\partial\tau_l}{\partial\varphi} = \frac{\partial x_l}{\partial\varphi}\frac{\sin\varphi}{(c+\cos\varphi\hat{v}_l)}. \tag{94}$$

Next we need the φ derivative of x_l, calculated from Eq. (89a), which, using Eq. (94), yields

$$\frac{\partial x_l}{\partial\varphi} = \frac{-c(\tau-\tau_l-\mathcal{T})\cos(\varphi-\alpha_m)}{1-c(\tau-\tau_l-\mathcal{T})\cos(\varphi-\alpha_m)}\frac{\partial\alpha_m}{\partial x_l} - c\left(\frac{\partial\mathcal{T}}{\partial x_l}+\frac{\sin\varphi}{(c+\hat{v}_l\cos\varphi)}\right)\sin(\varphi-\alpha_m). \tag{95}$$

Finally, setting $\partial G/\partial\varphi=0$ from Eq. (89b) yields

$$0 = \left(\frac{\partial\tau_l}{\partial\varphi}+\frac{\partial\mathcal{T}}{\partial x_l}\frac{\partial x_l}{\partial\varphi}\right)(c\cos(\varphi-\alpha_m)+\hat{v}_l)+c(\tau-\tau_l-\mathcal{T}) \times\sin(\varphi-\alpha_m)\left(1-\frac{\partial\alpha_m}{\partial x_l}\frac{\partial x_l}{\partial\varphi}\right) \tag{96}$$

which, using Eqs. (94), (95) and the fact that $x_l\cos\varphi - \sin\varphi y_l(\tau_l)=0$ [from Eqs. (90a), (90b)] becomes

$$\sin(\varphi-\alpha_m)-(c+\hat{v}_l\cos(\varphi-\alpha_m)) \times\left(\frac{\partial\mathcal{T}}{\partial x_l}+\frac{\sin\varphi}{(c+\hat{v}_l\cos\varphi)}\right)=0. \tag{97}$$

Now, using the results of the envelope condition on the observer side where φ is a function of x_l given implicitly by eliminating τ_l from Eqs. (90a), (90b), we can interpret Eq. (97) as an equation defining the bending angle, α_m, as a function of x_l if $\mathcal{T}(x_l)$ is known. In the regime of small angles, keeping only linear terms in φ and α_m, it reduces to

$$\varphi-\alpha_m-(c+\hat{v}_l)\left(\frac{\partial\mathcal{T}}{\partial x_l}+\frac{\varphi}{(c+\hat{v}_l)}\right)=0 \tag{98}$$

or, the equation for α_m in terms of the gradient of $\mathcal{T}$:

$$\alpha_m = -\left(1+\frac{\hat{v}_l}{c}\right)c\frac{\partial\mathcal{T}}{\partial x_l}\equiv\left(1+\frac{\hat{v}_l}{c}\right)\alpha. \tag{99}$$

This is the relationship between the bending angle and the gravitational time delay at the lens plane when the lens is moving. It represents a correction of a factor of $(1+v/c)$ with respect to the bending angle for the same deflector at rest. This is in agreement with aberration [8], according to which the angles made by light rays as observed in a moving frame are corrected by a factor of $(1+v/c)$ with respect to the angles made by the same light rays as observed in the rest frame. In our case both the directions of incoming and outgoing light rays are corrected by the same factor, and consequently their deviation α is corrected by the same factor.

The lens equation

$$x = x_l+(y-y_l(\tau_l+\mathcal{T}))\tan(\varphi-\alpha_m), \tag{100}$$

is obtained from by eliminating $\tau-\tau_l$ from Eqs. (89a), (89b). We are interested in the small angle regime so, neglecting higher order terms, we have

$$x = x_l+(y-\hat{y}_l)(\varphi-\alpha_m), \tag{101}$$

where we have used the assumption that $\mathcal{T}$ is small (of the same order as φ and α_m) and introduced the notation $\hat{y}_l \equiv y_l(\tau_l)$. From the ratio of Eq. (90a) over Eq. (90b) and for small φ, we have

$$\varphi = \frac{x_l}{\hat{y}_l} \tag{102}$$

which leads, via Eq. (101), to

$$x = \frac{y}{\hat{y}_l} x_l - (y - \hat{y}_l)\alpha_m \tag{103}$$

or

$$\frac{x}{y} = \frac{x_l}{\hat{y}_l} - \frac{y - \hat{y}_l}{y}\alpha_m . \tag{104}$$

Since in flat space the coordinate distance is equal to the angular diameter distance, we can use the substitutions $y = D_s$, $\hat{y}_l = \hat{D}_l$ and $y - \hat{y}_l = \hat{D}_{ls}$ to denote the distance to the source, the distance to the lens and the distance between the lens plane and the source plane at the time the wave front passes by the lens plane. With these substitutions the lens equation becomes

$$\frac{x}{D_s} = \frac{x_l}{\hat{D}_l} - \frac{\hat{D}_{ls}}{D_s}\alpha_m . \tag{105}$$

Thus the changes in the lens equation from lens motion amount to a factor of $(1 + \hat{v}_l/c)$ in the bending angle [from Eq. (99)] and the evaluation of the instantaneous position of the lens plane at the average time that the wave front reaches it.

The time delay can be obtained by solving for τ in Eq. (89b) and subtracting the travel time of a ray in the absence of the lens. We have thus

$$\Delta\tau = \frac{1}{c}\frac{y - y_l(\tau_l + \mathcal{T})}{\cos(\varphi - \alpha)} + \tau_l + \mathcal{T} - \frac{1}{c}\sqrt{x^2 + y^2}. \tag{106}$$

From Eq. (105) and Eq. (90b), for small φ,

$$\tau_l = c^{-1}\hat{y}_l\left(1 + \frac{1}{2}\varphi^2\right) \tag{107}$$

but *keeping up to second-order terms*, we find the time delay equation corrected for a moving lens:

$$\Delta\tau = (1 + \hat{v}_l/c)\mathcal{T} + \frac{\hat{D}_l \hat{D}_{ls}}{2cD_s}\alpha^2. \tag{108}$$

VI. COSMOLOGICAL THIN LENSES

In Secs. IV and V, we derived the implicit Fermat potentials for lensing by static and moving thin lenses where the lens plane separated two regions of flat spacetimes. In this section, we expand these results to consider lensing in spacetimes that (apart from the lens) are described by Friedmann-Robertson-Walker (FRW) metrics. Our approach is to utilize the fact that all the FRW spacetimes are conformally flat to rescale physical quantities in the flat lensing case into physical quantities for observers living in a FRW universe. We begin by noting that there is, for each FRW spacetime ($k = \pm 1,0$), a coordinate transformation between the "natural" FRW coordinates $(\tau, \chi, \theta, \phi) = x_F^a$ to "Minkowski" coordinates $(t, x, y, z) = x_M^a$,

$$x_M^a = x_M^a(x_F^a), \tag{109}$$

such that the metric can be written as

$$\begin{aligned} d\tilde{s}^2 &= d\tau^2 - a^2(\tau)dl_k^2(\chi, \theta, \phi) \\ &= \Omega^2(dt^2 - dx^2 - dy^2 - dz^2). \end{aligned} \tag{110}$$

Here, dl_k^2 is the metric of a constant curvature three-space (sphere for $k = +1$, hyperboloid for $k = -1$, or flat space for $k = 0$). In general, the conformal transformation is such that

$$\Omega = f_k(x^a)a(\tau), \tag{111}$$

for some function f_k. Because conformal transformations preserve light rays, the implicit Fermat potentials for lensing in our previous sections are implicit Fermat potentials for the cosmological spacetime and can be expressed in terms of either coordinate system, x_F^a or x_M^a, via Eq. (109).

Often, when a symbol is used for a quantity described in the Minkowski space and we want to distinguish it from the same quantity in the cosmological space we will indicate that by a twiddle, i.e., A versus $\tilde{A}$. Also, often when we are refering to a quantity associated with a moving lens we will indicate that by a "hat," e.g., the y_l position of the moving lens becomes $y_l(\tau) = \hat{y}_l$.

Several subtle issues arise from the fact that, in general, a fixed space point $(\chi, \theta, \phi) = \text{const}$ moving with the cosmological flow in the $k = \pm 1$ cases, has in the associated Minkowski space a coordinate velocity; thus comoving lenses in the FRW coordinates are modeled by lenses that move, i.e., have a coordinate velocity, in the Minkowski space. This, in turn, leads to the consideration of aberration affects in the bending angle and source angles which then influence angular-diameter distances. A second subtle issue is the relationship between the gravitational time delay as computed in the FRW spaces compared to those computed in the associated Minkowski space. We point out that the time delays are not simply conformally related; the cosmological time delay must be calculated independently from the Minkowski delays. This issue will be further explored later in this section.

We consider the universe to be two sections of a FRW spacetime appropriately matched at a lens plane. The matter distribution in the lens plane determines two related gravitational time delay functions: the (Minkowski) *coordinate time* delay, $\mathcal{T}$, and the related cosmological proper time delay, $\mathcal{T}_C$, with

$$\mathcal{T}_C = \Omega_l \mathcal{T} \tag{112}$$

where Ω_l is the conformal factor evaluated at the lens at the time of arrival of the ray. The relevant functions depend on the Minkowski coordinates in the lens plane, x_l. We interpret $\mathcal{T}_C(x_l)$ as the cosmological proper time measured by a person just past the lens (on the observer's side) between a light ray emitted at the source at time τ_s that passed through the lens and one emitted at the same τ_s that was not influenced by the lens.

As we saw for the moving lens in Sec.V, our implicit Fermat potential, in 2+1 dimensions, is given by Eqs. (89a), (89b) with the envelope condition yielding

$$\alpha_m = -(c+\hat{v}_l)\frac{\partial \mathcal{T}}{\partial x_l} \equiv (1+\hat{v}_l/c)\,\alpha, \tag{113}$$

where $\hat{v}_l$ is the average speed of the lens during the time the light ray is delayed and α_m is the bending angle observed by a stationary observer with a moving lens and α the bending angle for an observer comoving with the lens. Our lens equation (in the small angle approximation) is simply

$$x = \varphi y - (y-\hat{y}_l)\,\alpha_m, \tag{114}$$

where (x,y) are the source position, $\hat{y}_l$ is the position of the lens plane when the light ray leaves the lens, and φ is the observation angle.

We consider how to introduce some (modified) angular-diameter distance **s** for lenses that are moving in a Minkowski spacetime. Here, we interpret the source coordinate, *x*, as the *metric distance*, Δ_s, between the source (in the source plane) and the optical axis (i.e. $x \equiv \Delta_s$). We then convert the coordinate distances, *y*, and $y-\hat{y}_l$ into angular-diameter distances, D_s and $\hat{D}_{ls}$, as follows. Let β be the unlensed angle that the source at *x* subtends at the observer and let γ be the angle that *x* subtends at the lens plane (when the light ray leaves the lens plane). Then

$$y \equiv D_s = \frac{x}{\beta} = \frac{\Delta_s}{\beta}, \tag{115a}$$

$$y-\hat{y}_l \equiv \hat{D}_{ls} = \frac{x}{\gamma} = \frac{\Delta_s}{\gamma}. \tag{115b}$$

Eliminating y and $y-\hat{y}_l$ from Eq. (114), the lens equation becomes

$$\beta = \varphi - \frac{\hat{D}_{ls}}{D_s}\alpha_m. \tag{116}$$

This expression for the lens equation with a moving lens is in terms of distances and angles relative to the frame of a stationary observer in the Minkowski space. As we saw earlier, by the aberration (of light) equation, the bending angle seen by the stationary observer, α_m, is related to the bending angle α seen in the frame comoving with the lens, by Eq. (113),

$$\alpha_m = \alpha\left(1+\frac{\hat{v}_l}{c}\right), \tag{117}$$

so that Eq. (116) becomes

$$\beta = \varphi - \frac{\hat{D}_{ls}}{D_s}\alpha_m = \varphi - \frac{\hat{D}_{ls}}{D_s}\alpha\left(1+\frac{\hat{v}_l}{c}\right) \tag{118}$$

with $\hat{D}_{ls}$ the time-dependent angular-diameter distance of the lens to the source *as seen by a stationary observer*. On the other hand, if we wish to use (as we will in a moment) the angular-diameter distance measured from the lens to the source *in the comoving lens frame*, $D_{ls}^{(co)}$, we have that

$$D_{ls}^{(co)} = \frac{x}{\gamma^{(co)}} = \frac{\Delta_s}{\gamma^{(co)}}, \tag{119}$$

where $\gamma^{(co)}$ is the angle that *x* subtends according to an observer comoving with the lens plane. Using the aberration equation, Eq. (117), for the angle γ, we have

$$\hat{D}_{ls} = D_{ls}^{(co)}\frac{1}{\left(1+\frac{v}{c}\right)}. \tag{120}$$

The (moving) lens equation becomes

$$\beta = \varphi - \frac{D_{ls}^{(co)}}{D_s}\alpha. \tag{121}$$

Note that neither $D_{ls}^{(co)}$ or $\hat{D}_{ls}$ are directly observable quantities. The issue of which to use depends on the physical situation that is being addressed. If we utilize the coordinate transformation, Eq. (109), that relates the Minkowski coordinates to the ''natural'' FRW coordinates, the moving lens in the Minkowski background becomes a stationary lens in the conformally related cosmological spacetime. Hence, $D_{ls}^{(co)}$ should be used in the lens equation in the case where our interest lies in the FRW space.

The just completed discussion pertained to the Minkowski background; we now transform the lens equation, Eq. (121), to the cosmological background. We assume that the observer's world line is given by $(\tau,\chi=0,\theta=0,\phi=0)$ in the standard FRW coordinates. In any of the three FRW models, such an observer is stationary at $(t,r=0,\theta=0,\phi=0)$. However, the lens and source, located along the world lines of constant (χ,θ,ϕ) will be ''moving'' in the Minkowski coordinates. (The motion of the source in the Minkowski coordinates plays no important role.) In the cosmological space, the metric distance between the source location and the optical axis (in the source plane) is given by $\tilde{\Delta}_s = \Omega(x_s^a)\Delta_s$ where the conformal factor is evaluated at the source location when the light ray leaves the source. Using

$$\tilde{D}_s = \frac{\tilde{\Delta}_s}{\beta} = \Omega(x_s^a)D_s \tag{122a}$$

$$\tilde{D}_{ls}=\frac{\tilde{\Delta}_s}{\gamma^{(co)}}=\Omega(x_s^a)D_{ls}^{(co)} \tag{122b}$$

for the cosmological angular-diameter distances, our lens equation immediately becomes

$$\beta=\varphi-\frac{\tilde{D}_{ls}}{\tilde{D}_s}\alpha, \tag{123}$$

which is identical (in form) to the flat-space lens equation with a stationary lens and to the conventional cosmological lens equation [1,3].

We now turn from the cosmological lens equation to the cosmological time of arrival equation. In Sec. V, we derived an expression for the ''time delay'' at the observer between the true ''lensed'' path and the ''unlensed'' path in a Minkowski model with a moving lens, namely,

$$\Delta\tau=(1+\hat{v}_l/c)\mathcal{T}+\frac{\hat{D}_l}{2c}\frac{\hat{D}_{ls}}{D_s}\alpha_m^2. \tag{124}$$

Using the transformation properties of α_m and $\hat{D}_{ls}$ just described, Eq. (124) has the form

$$\Delta\tau=(1+\hat{v}_l/c)\left(\mathcal{T}+\frac{D_l}{2c}\frac{D_{ls}^{(co)}}{D_s}\alpha^2\right), \tag{125}$$

where α is the bending angle seen in the comoving frame and $\mathcal{T}$ is the *Minkowski coordinate time delay* measured along a world line just to the observer's side of the lens between the arrival time for a lensed and unlensed ray. To transform Eq. (125) to cosmological variables, we first note that

$$\frac{D_{ls}^{(co)}}{D_s}=\frac{\tilde{D}_{ls}}{\tilde{D}_s} \tag{126}$$

where the ratio on the right side refers to cosmological angular-diameter distances. Next, we see that if Δ_l is the Minkowski metric distance from the optical axis to the ''image location'' in the lens plane and φ is the angle this distance subtends at the observer we have

$$\tilde{D}_l=\frac{\Omega(x_l^a)\Delta_l}{\varphi}=\Omega(x_l^a)\hat{D}_l, \tag{127}$$

which relates the angular-diameter distance from observer to the lens in the FRW cosmology to the conformally related ''Minkowski'' angular-diameter distance. (Here, the conformal factor is evaluated at the location of the lens at the time the light ray leaves the lens.) If we multiply both sides of Eq. (125) by $\Omega_0\Omega_l$ where Ω_0 and Ω_l are respectively the conformal factor evaluated at the observer at the observation time and at the lens at the time the light ray leaves the lens, we have

$$\Delta\tilde{\tau}_0=(1+\hat{v}_l/c)\frac{\Omega_0}{\Omega_l}\left(\mathcal{T}_C+\frac{\tilde{D}_l}{2c}\frac{\tilde{D}_{ls}}{\tilde{D}_s}\alpha^2\right), \tag{128}$$

where we have used

$$\mathcal{T}_C=\Omega_l\mathcal{T} \tag{129a}$$

$$\Delta\tilde{\tau}_0=\Omega_0\Delta\tau \tag{129b}$$

the latter being the cosmological proper time between reception of the two rays at the observer.

From the relation

$$(1+\hat{v}_l/c)\frac{\Omega_0}{\Omega_l}\approx\frac{a_0}{a_l}\equiv 1+z, \tag{130}$$

which is derived in the Appendix for small coordinate velocities, $\hat{v}_l/c$, in the associated Minkowski space, from Eq. (128), our final cosmological time delay equation is

$$\Delta\tilde{\tau}_0=(1+z)\left(\mathcal{T}_C+\frac{\tilde{D}_l}{2c}\frac{\tilde{D}_{ls}}{\tilde{D}_s}\alpha^2\right). \tag{131}$$

Our last issue is to describe how the cosmological proper time delay $\mathcal{T}_C$ is to be computed. To compute $\mathcal{T}_C$, and, by association, the Minkowksi coordinate time delay, $\mathcal{T}=\Omega_l^{-1}\mathcal{T}_C$, we first consider the Newtonian potential

$$U_C=-\int\frac{dm'}{|\vec{x}-\vec{x}'|_C}, \tag{132}$$

where $dm'=\rho_C(x')d^3x'$ is the ''observable'' mass element in the cosmological space, and *the norm in the denominator is taken in the cosmological metric*. To find the cosmological gravitational time delay, one integrates the Newtonian potential over the path from the source to the observer,

$$\mathcal{T}_C=\frac{-2}{c^3}\int U_C dl_C, \tag{133}$$

where dl_C is *the distance element along the path in the cosmology*. In the standard thin-lens approximation, it is assumed that $\rho_C(x')$ is nonzero only in a thin region perpendicular to the optical axis. Thus, when computing the time delay from the integral in Eq. (133), one need only consider the region along the trajectory that is very close to the lens plane. First, we point out that under conformal rescaling, since dm' is conformally invariant, that

$$U_C=-\int\frac{dm'}{|\vec{x}-\vec{x}'|_C}\approx-\int\frac{dm'}{\Omega_l|\vec{x}-\vec{x}'|_M}=\Omega_l^{-1}U_M \tag{134}$$

where $|\vec{x}-\vec{x}'|_M=$ is the metric distance between two points in the lens plane taken in the conformally related Minkowski metric. Because U_C only has support in the vicinity of the lens and $dl_C=\Omega dl_M$, where dl_M is the distance element in a Minkowski spacetime, we have

$$\mathcal{T}_C=\frac{-2}{c^3}\int U_C dl_C=\frac{-2}{c^3}\int U_M dl_M=\mathcal{T}_M. \tag{135}$$

There is an apparent conundrum here. $\mathcal{T}_C$ is the physical proper-time delay which *appears to be the same* as the Minkowski space proper-time delay for the same physical situation, i.e., arising from potentials

$$U_C=-\int \frac{dm'}{|\vec{x}-\vec{x}'|_C} \tag{136a}$$

$$U_M=-\int \frac{dm'}{|\vec{x}-\vec{x}'|_M} \tag{136b}$$

with the same dm' in both cases. The conundrum is resolved by noting that though the dm' are the same the density functions are different.

We compute the mass densities by introducing *physical coordinates* at the lens (local Lorentzian coordinates), so that the cosmological metric *at the lens* takes the form

$$\begin{aligned} ds^2&=(dt_C^2-dx_C^2-dy_C^2-dz_C^2)\\ &=\Omega_l^2(dt^2-dx^2-dy^2-dz^2) \end{aligned} \tag{137}$$

with

$$x_C^a=\Omega_l x^a. \tag{138}$$

Using

$$\begin{aligned} dm'&=\rho_C(x_C^a)d^3x_C=\rho_C(\Omega_l x^a)\Omega_l^3 d^3x\\ &=\rho_M(x^a)d^3x \end{aligned} \tag{139}$$

we have the relationship between the physical density, ρ_C, and the *equivalent fictitious* Minkowski space density $\rho_M(x^a)$,

$$\rho_M(x^a)=\rho_C(\Omega_l x^a)\Omega_l^3. \tag{140}$$

Returning to the issue of the cosmological lens equation, Eq. (123), and associated bending angle, Eq. (113), we see that the present results concerning the cosmological time delay agree with the earlier results, since from Eqs. (129a), (138), and (113), we have

$$\alpha=-c\frac{\partial\mathcal{T}}{\partial x_l}=-c\frac{\Omega_l\partial\mathcal{T}}{\Omega_l\partial x_l}=-c\frac{\partial\mathcal{T}_C}{\partial x_{Cl}}, \tag{141}$$

the cosmological bending angle.

All our results obtained from the envelope construction now agree with the standard results appearing in the published literature. In addition, it is clear that the method can be used to extend the results to the case of a lens plane that has a peculiar motion above the Hubble flow by letting the velocity of the lens plane in the underlaying flat space be arbitrary.

As a final comment here, we point out that this method of obtaining the FRW lens equation and time delay equation by conformal rescaling of an associated Minkowski space can be extended, largely unchanged, to any spacetime that is conformally related to Minkowski space. It is not clear whether this observation has any physical application.

VII. CONCLUDING REMARKS AND OUTLOOK

We have introduced a novel approach to gravitational lensing built on a variational principle that is analogous to, but distinctly different from, the conventional version of Fermat's principle.

Our approach is to find the envelope of an appropriate family of null surfaces containing the observer's location. We have shown how this approach reproduces the astrophysical scenario of static thin lenses as an illustration. More interestingly, though, we have been able to use this approach to obtain the correction to first order in v/c that the motion of the deflector imposes on the bending angle in the approximation of thin lenses. This calculation can alternatively be done by direct integration of the null geodesics in time-dependent spacetime perturbations off flat space. In this respect, our result agrees with such a direct calculation of the bending angle by moving thin lenses carried out by Pyne and Birkinshaw [9], whereas it differs from others [10] by an overall sign and a factor of 2. Work is in progress to clarify this difference and will be reported elsewhere. Notice that a redshift of 0.001 might conceivably bring this effect into the observable regime in the future.

Because in the thin-lens case in cosmology our approach is based entirely on the underlying conformally flat space, we also have an alternative derivation of the lens equation in cosmology, which does not seem to have been exploited in the literature so far. We feel our derivation clarifies some points that remain obscure in the presentations of the lens mapping in cosmology that appear in [1] and [3]. On the other hand, the reader interested in a very complete derivation of the lens mapping in cosmology by integration of the null geodesics will definitely benefit from the excellent article by Pyne and Birkinshaw [11].

Throughout this work, we have assumed that a metric is given that represents the structure of the deflector or lens. In this sense, we have kept ourselves within the kinematics of the implicit Fermat potential. We have not addressed the issue of the dynamics of the Fermat potential; namely, how the implicit Fermat potential is directly affected by the structure of the deflectors. The resolution of this important question in principle involves two steps: to solve the Einstein equations for the metric, and then to use the metric as a given source into the eikonal equation for the implicit Fermat potential. By contrast, in the standard thin-lens scenario, the Fermat potential is obtained directly from the surface mass distribution by means of a 2-dimensional equation of the Poisson type. It is reasonable to ask whether an analogous scheme to obtain field equations for the implicit Fermat potential directly in terms of the structure of the deflector exists in cases other than the thin-lens scenario. This question will be discussed elsewhere.

ACKNOWLEDGMENTS

We are indebted to Juergen Ehlers for a very enlightening discussion. This work was supported by the NSF under grants No. PHY-0088951 and No. PHY-0070624.

APPENDIX

In the main text we quoted the approximate equation,

$$\frac{a_0}{\Omega_0} \approx \frac{a_l}{\Omega_l}(1+v_l/c), \tag{A1}$$

which relates the ratio of conformal factors Ω to the cosmological scale factor a, at two different points, the observer and the lens, that are connected by a null geodesic. The quantity v_l is the velocity of the lens in the associated Minkowski coordinates evaluated as the ray leaves the lens.

For the case of $k=0$ this result is trivially true since then $a=\Omega$ and $v_l=0$. For ease of presentation we will only give the proof for the $k=-1$ spacetime, but the calculations are exactly paralled in the case of $k=+1$ and the result is exactly the same. We start by recalling that we have the metric in the two forms

$$d\tilde{s}^2 = a^2(d\eta^2 - d\chi^2 - \sinh^2\chi(d\theta^2 + \sin^2\theta d\varphi^2)) \tag{A2a}$$

$$= \Omega^2(c^2dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2)) \tag{A2b}$$

with Eq. (109) explicitly given by

$$ct = \frac{\sinh\eta}{\cosh\eta + \cosh\chi} \tag{A3a}$$

$$r = \frac{\sinh\chi}{\cosh\eta + \cosh\chi} \tag{A3b}$$

and

$$\Omega^2 = \frac{4a^2}{(1-(ct+r)^2)(1-(ct-r)^2)}, \tag{A4}$$

or, from Eq. (A3),

$$\frac{\Omega}{a} = \cosh\chi + \cosh\eta. \tag{A5}$$

We are interested in a lens at some fixed value of θ,φ and $\chi=\chi_l$, sending a light signal to the observer located at $\chi=0$. The lens and arrival times are η_l and η_0, respectively, and we have

$$\eta_0 - \eta_l = \chi_l. \tag{A6}$$

Using Eqs. (A5) and (A6) we can construct

$$\frac{\Omega_l}{a_l} = \frac{\Omega_0}{a_0}\frac{\cosh\chi_l + \cosh(\chi_l - \eta_0)}{1+\cosh\eta_0} \tag{A7}$$

or, for small χ_l,

$$\frac{\Omega_l}{a_l} \approx \frac{\Omega_0}{a_0}\left(1 - \frac{\chi_l \sinh\eta_0}{1+\cosh\eta_0}\right). \tag{A8}$$

By calculating the velocity from

$$\frac{dr}{dt} = \frac{\left.\frac{\partial r}{\partial\eta}\right|_\chi}{\left.\frac{\partial t}{\partial\eta}\right|_\chi} \tag{A9}$$

we have

$$\frac{v}{c} \equiv -\frac{\sinh\chi\sinh\eta}{1+\cosh\chi\cosh\eta}. \tag{A10}$$

We thus see that a FWR lens plane approaches the observer in the conformally related Minkowski space. Evaluating at the lens time we have

$$\frac{v_l}{c} = \frac{\sinh\chi_l\sinh(\chi_l-\eta_0)}{1+\cosh\chi_l\cosh(\chi_l-\eta_0)} \tag{A11}$$

and for small values of χ_l

$$\frac{v_l}{c} \approx -\frac{\chi_l\sinh\eta_0}{1+\cosh\eta_0}. \tag{A12}$$

Using Eq. (A12) with Eq. (A8), we have our result, Eq. (A1).

[1] P. Schneider, J. Ehlers, and E.E. Falco, *Gravitational Lenses* (Springer-Verlag, New York, 1992).

[2] V. Perlick, *Ray Optics, Fermat's Principle and Applications to General Relativity* (Springer-Verlag, Berlin, 2000).

[3] A.O. Petters, H. Levine, and J. Wambsganss, *Singularity Theory and Gravitational Lensing* (Birkhauser, Boston, 2001).

[4] S. Frittelli, G. Silva-Ortigoza, and E.T. Newman, J. Math. Phys. **40**, 383 (1999).

[5] J. Ehlers and E.T. Newman, J. Math. Phys. **41**, 3344 (2000).

[6] S. Frittelli and E.T. Newman, Phys. Rev. D **59**, 124001 (1999).

[7] J. Ehlers, S. Frittelli, and E.T. Newman, *Graviational Lensing from a Space-Time Perspective*, Boston Studies in the Philosophy of Science: Festschrift for John Stachel, edited by J. Renn (Kluwer, Dordrecht, 2001).

[8] See, for example, L. Landau and E. Lifschitz, *The Classical Theory of Fields* (Addison-Wesley, Cambridge, MA, 1951).

[9] T. Pyne and M. Birkinshaw, Astrophys. J. **415**, 459 (1993).

[10] S. Capozziello, G. Lambiase, G. Papini, and G. Scarpetta, Phys. Lett. A **254**, 11 (1999).

[11] T. Pyne and M. Birkinshaw, Astrophys. J. **458**, 46 (1996).