

Strong lensing, plane gravitational waves and transient flashes

Abraham I. Harte

Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut
Am Mühlenberg 1, D-14476 Golm, Germany

E-mail: harte@aei.mpg.de

Abstract. Plane-symmetric gravitational waves are considered as gravitational lenses. Numbers of images, frequency shifts, mutual angles, and image distortion parameters are computed exactly in essentially all non-singular plane wave spacetimes. For a fixed observation event in a particular plane wave spacetime, the number of images is found to be the same for almost every source. This number can be any positive integer, including infinity. Wavepackets of finite width are discussed in detail as well as waves which maintain a constant amplitude for all time. Short wavepackets are found to generically produce up to two images of each source which appear (separately) only some time after the wave has passed. They are initially infinitely bright, infinitely blueshifted images of the infinitely distant past. Later, these images become dim and acquire a rapidly-increasing redshift. For sufficiently weak wavepackets, one such “flash” almost always exists. The appearance of a second flash requires that the Ricci tensor inside the wave exceed a certain threshold. This might occur if a gravitational plane wave is sourced by, e.g., a sufficiently strong electromagnetic plane wave.

1. Introduction

The theory of gravitational lensing has by now reached a considerable degree of sophistication [1, 2, 3, 4]. Theorems have been found predicting (or bounding) the number of images in very general systems [5, 6, 7]. Shapes of stable caustics have been exhaustively classified [2, 3, 8, 9], a non-perturbative notion of the lens map has been obtained [10], and various universal behaviors of images have been found for sources lying near caustics [11, 12, 13]. These general results have been complemented by a number of detailed calculations for specific types of lenses. The majority of such calculations have been performed within the quasi-Newtonian viewpoint of gravitational lensing commonly used in astrophysics [3]. While various flavors of this formalism exist, most require that bending angles be small and that all lenses be nearly-Newtonian mass distributions.

This is to be contrasted with the more fundamental picture of gravitational lensing where light rays are modelled as null geodesics in a Lorentzian spacetime. Within this context, specific lensing calculations have been performed in Kerr, Reissner-Nordström, and a handful of other geometries [1]. While curvatures and bending angles may be large in these examples, they all involve (at least conformally) stationary

spacetimes. It is of interest to understand if qualitatively new effects appear in dynamical cases.

This paper considers the bending of light by (decidedly non-stationary) plane-symmetric gravitational waves. Gravitational lensing by gravitational waves has previously been considered by a number of authors, although almost all of this work has been carried out within the weak-field regime [14, 15, 16, 17, 18]. One exception is [19], where redshifts were computed in an exact solution to the vacuum Einstein equation representing a plane gravitational wave. This work considered only very specific waveforms, and was confined to a coordinate patch too small to include caustics and other effects associated with the formation of multiple images. Separately, extensive work has been devoted to non-perturbatively understanding the geodesic structure of generic plane wave spacetimes [20, 21, 22, 23, 24, 25, 26]. This is clearly a subject closely related to gravitational lensing, although few explicit relations between the two subjects appear to have been made (see, however, remarks in [1]).

It is the purpose of this work to provide a comprehensive and non-perturbative discussion of lensing in plane wave spacetimes. These geometries are a well-known subclass of *pp*-waves; plane-fronted waves with parallel rays. Many plane wave spacetimes are exact solutions to the vacuum Einstein equation. Others may be interpreted as, e.g., exact solutions in Einstein-Maxwell theory. While rather idealized from the astrophysical perspective, plane wave spacetimes admit a wide variety of interesting phenomena. Depending on the waveform, any number of astigmatic and anastigmatic caustics may exist. Examples admitting any specified number of images – even an infinite number – are easily constructed. In particular, even numbers of images can exist [1]. The image count can also change in time even when a source does not cross an observer’s caustic. Despite all of these properties, plane wave spacetimes are geometrically very mild. They are topologically equivalent to \mathbb{R}^4 and admit coordinate systems which cover the entire manifold.

Aside from their value as models of gravitational radiation, the plane wave spacetimes considered here have also found numerous applications via the Penrose limit. This limit provides a sense in which the metric near any null geodesic in any spacetime is equivalent to the metric of an appropriate plane wave spacetime [27, 28]. It allows problems in relatively complicated spacetimes to be reduced to equivalent problems in plane wave spacetimes (which are often much simpler). This has been particularly valuable within string theory and related fields [29, 30]. Penrose limits have also been applied to ordinary quantum field theory in order to investigate causality and effective indices of refraction for photons and gravitons propagating in curved spacetimes [24, 25, 31]. More recently, Penrose limits were used to deduce the effect of caustics on Green functions associated with the propagation of classical fields in curved spacetimes [23]. Given the content of the Penrose limit, lensing in plane wave spacetimes might be related to lensing in generic spacetimes as seen by ultrarelativistic observers. We make no attempt, however, to justify this conjecture.

This paper starts by providing a self-contained review of plane wave spacetimes in Sect. 2. Although most of this material is not new [20, 21, 22, 23, 24, 25, 26], it is not widely known. After presenting this foundational material, Sect. 3 derives the number of images of a point source that may be viewed in plane wave spacetimes. Under generic conditions, this is found to depend only on the waveform and a certain time parameter associated with the observation event. The number of images does not depend on any properties of the source. Once this is established, Sect. 4 computes image positions, frequency shifts, angles, and image distortion parameters in general

plane wave spacetimes. Sect. 5 applies these results to symmetric plane waves, which have constant waveforms. These geometries produce an infinite number of images of almost every source. Their lensing properties are found to change significantly if the Ricci tensor is increased beyond a certain threshold. Lastly, Sect. 6 discusses “sandwich waves;” wavepackets with finite width. These spacetimes generically admit images which appear at discrete times and then persist indefinitely. Such images initially provide infinitely blueshifted, infinitely bright pictures into the infinitely distant past. Very quickly, however, such images become highly redshifted and effectively fade away. One of these “transient flashes” is produced by almost every sufficiently weak vacuum (Ricci-flat) wave. For weak waves, a second flash appears only if the Ricci tensor of the wavepacket exceeds a certain threshold. Throughout this work, the spacetime is assumed to be everywhere transparent. The language used also assumes that the geometric optics approximation [3] holds even in situations where it would be severely strained (such as when light is emitted near an observer’s caustic).

Notation

This paper restricts attention only to plane wave spacetimes in four spacetime dimensions. Our sign conventions follow those of Wald [32]. The signature is $-+++$. Latin letters a, b, \dots (and occasionally A, B, \dots) from the beginning of the alphabet are used to denote abstract indices. Greek letters μ, ν, \dots are used to denote four-dimensional coordinate indices. The Latin letters i, j, \dots are instead coordinate indices associated with the 2 directions transverse to the direction of wave propagation. Objects involving the latter type of index are often written in boldface with all indices suppressed. They are then manipulated using the standard notation of linear algebra [e.g., $A_{ki}B_{kj} = (\mathbf{A}^\top \mathbf{B})_{ij}$ and $|\mathbf{x}| = \sqrt{x_i x_i}$]. Overall, notation related to plane wave spacetimes closely follows the conventions of [23].

2. Geometry of plane wave spacetimes

Plane wave geometries may be interpreted as simple models for gravitational waves emitted from distant sources. Alternatively, they arise as universal limits for the geometry near null geodesics in any spacetime [27, 28]. The typical definition of a plane wave spacetime (M, g_{ab}) requires that $M = \mathbb{R}^4$ and that there exist global coordinates $(u, v, \mathbf{x}) = (u, v, x^1, x^2) : M \rightarrow \mathbb{R}^4$ such that the line element has the form

$$g_{\mu\nu} dx^\mu dx^\nu = -2dudv + H_{ij}(u)x^i x^j du^2 + (dx^1)^2 + (dx^2)^2. \quad (1)$$

$H_{ij} = (\mathbf{H})_{ij}$ is any symmetric 2×2 matrix. Its components describe the waveforms associated with a wave’s three polarization states (one of these polarizations is excluded in the vacuum case $R_{ab} = 0$). The u coordinate is interpreted as a phase parameter for the wave, while v affinely parametrizes its rays. The remaining two coordinates x^i span spacelike wavefronts transverse to the wave’s direction of propagation.

Note that if $\mathbf{H} = 0$ in some region, the spacetime is locally flat there. In terms of a Minkowski coordinate system (t, x^1, x^2, x^3) , u and v satisfy $u = (t + x^3)/\sqrt{2}$ and $v = (t - x^3)/\sqrt{2}$ in such a region. We consider only nontrivial plane waves, so \mathbf{H} cannot vanish everywhere.

The physical interpretation of the so-called Brinkmann metric (1) as a plane-symmetric gravitational wave follows from considering the integral curves of the

vector field $\ell^a = (\partial/\partial v)^a$. These curves form a null geodesic congruence which may be interpreted as the rays of the gravitational wave. $\nabla_a \ell^b = 0$, so these rays have vanishing expansion, shear, and twist. There is therefore a sense in which they are everywhere parallel to one another. All rays are also orthogonal to the family of spacelike 2-surfaces generated by the two commuting spacelike vector fields $X_{(i)}^a = (\partial/\partial x^i)^a$. The induced metric on each such surface is flat: The wavefronts are 2-planes. The curvature is constant on these planes in the sense that the $X_{(i)}^a$ are curvature collineations:

$$\mathcal{L}_{X_{(i)}} R_{abc}{}^d = 0. \quad (2)$$

Despite this, the $X_{(i)}^a$ are not everywhere Killing. There do, however, exist linear combinations of $X_{(i)}^a$ and ℓ^a which *are* Killing. ℓ^a itself is also Killing, which may be interpreted as a statement that plane waves do not deform along their characteristics. Plane wave spacetimes admit a minimum of five linearly independent Killing fields. These symmetries are discussed more fully in Sect. 2.3.

All non-vanishing coordinate components of the Riemann tensor may be determined from

$$R_{uiuj} = -H_{ij}. \quad (3)$$

It follows from this that the Ricci tensor is

$$R_{ab} = -\text{Tr } \mathbf{H} \ell_a \ell_b, \quad (4)$$

where Tr denotes the ordinary (Euclidean) trace of the 2×2 matrix \mathbf{H} . The Ricci scalar always vanishes in plane wave spacetimes. More generally, there are no nonzero scalars formed by local contractions of the metric, the curvature, and its derivatives: $R^{ab} R_{ab} = R_{abcd} R^{abcd} = R_{abcd} \epsilon^{abfh} R_{fh}{}^{cd} = \dots = 0$. This is analogous to the fact that plane electromagnetic waves satisfy, e.g., $F_{ab} F^{ab} = \epsilon^{abcd} F_{ab} F_{cd} = 0$. Note, however, that electromagnetic plane waves are not the only electromagnetic fields with vanishing field scalars. Similarly, plane wave spacetimes are not the only curved geometries with vanishing curvature scalars [33].

It follows from (4) that plane wave spacetimes satisfying the vacuum Einstein equation (and the vacuum equations of many alternative theories of gravity [34]) are characterized by the simple algebraic constraint $\text{Tr } \mathbf{H} = 0$. For vacuum waves, there exist two scalar functions h_+ and h_\times such that

$$\mathbf{H} = \begin{pmatrix} -h_+ & h_\times \\ h_\times & h_+ \end{pmatrix}. \quad (5)$$

h_+ and h_\times describe the waveforms for the two polarization states of a gravitational plane wave propagating in vacuum. A plane wave is said to be linearly polarized if h_+ and h_\times are linearly dependent (in which case one of these functions can be eliminated by a suitable rotation of the transverse coordinates \mathbf{x}).

If h_+ and h_\times have compact support, the geometry is said to be a sandwich wave. This name evinces the image of a curved region of spacetime “sandwiched” between null hyperplanes in a geometry that is otherwise Minkowski. Physically, it corresponds to a wavepacket of finite length. Note that the planar symmetry considered here is very special in the sense that passing waves do not necessarily leave any “tail” behind them. After interacting with a sandwich wave, all observers enter a region of spacetime which is perfectly flat. There is a sense in which test fields propagating on plane wave spacetimes also have no tails [23, 35].

A general (not necessarily vacuum) wave profile \mathbf{H} may be built by adding to (5) a term proportional to the identity matrix δ . There then exists a third polarization function h_{\parallel} such that

$$\mathbf{H} = \begin{pmatrix} -h_+ - h_{\parallel} & h_{\times} \\ h_{\times} & h_+ - h_{\parallel} \end{pmatrix}. \quad (6)$$

If the Ricci tensor of such a wave is associated with a stress-energy tensor via Einstein's equation, that stress-energy tensor obeys the weak energy condition if and only if $h_{\parallel} \geq 0$. Assuming this, the stress-energy tensors associated with (6) are very simple. They could be generated by, e.g., electromagnetic plane waves with the form

$$F_{ab} = 2h_{\parallel}^{\frac{1}{2}} \nabla_{[a} u \nabla_{b]} x^1. \quad (7)$$

Alternatively, (6) could be associated with the stress-energy tensor of the massless Klein-Gordon plane wave

$$\phi = \int_u h_{\parallel}^{\frac{1}{2}}(w) dw. \quad (8)$$

Besides the vacuum case $h_{\parallel} = 0$, another interesting class of wave profiles are those that are conformally flat. These satisfy $h_+ = h_{\times} = 0$, so $\mathbf{H} \propto \delta$. As gravitational lenses, all caustics of conformally-flat plane waves are associated with “perfect” anastigmatic focusing. For more general plane waves, caustics are typically (but not necessarily) associated with astigmatic focusing.

2.1. The matrices \mathbf{A} and \mathbf{B}

The geometry of plane wave spacetimes has been analyzed in detail by a number of authors [20, 21, 22, 23, 24, 25, 26]. One essential conclusion of this work is that nearly all interesting properties of plane wave spacetimes may be deduced from the properties of 2×2 matrices $\mathbf{E} = \mathbf{E}(u)$ satisfying the differential equation

$$\ddot{\mathbf{E}} = \mathbf{H}\mathbf{E}. \quad (9)$$

This is a “generalized oscillator equation” with $-\mathbf{H}$ acting like a matrix of squared frequencies. Eq. (9) arises when solving for geodesics or Jacobi fields in plane wave spacetimes. Bitensors such as Synge's function and the parallel propagator may be written explicitly in terms of its solutions. The same is also true for a plane wave's Killing vectors.

It is convenient to write all possible matrices \mathbf{E} in terms of two particular solutions. Fix any $\ddagger u_o \in \mathbb{R}$ and define $\mathbf{A}(\cdot, u_o)$ and $\mathbf{B}(\cdot, u_o)$ to be solutions to (9) (where derivatives are applied to the first arguments of \mathbf{A} and \mathbf{B}) with the initial conditions

$$\lim_{u_s \rightarrow u_o} \mathbf{A}(u_s, u_o) = \lim_{u_s \rightarrow u_o} \partial_{(1)} \mathbf{B}(u_s, u_o) = \delta, \quad (10)$$

$$\lim_{u_s \rightarrow u_o} \partial_{(1)} \mathbf{A}(u_s, u_o) = \lim_{u_s \rightarrow u_o} \mathbf{B}(u_s, u_o) = 0. \quad (11)$$

Here, the notation $\partial_{(1)} \mathbf{A}$ indicates a partial derivative with respect to the first argument of \mathbf{A} . We assume for simplicity that \mathbf{H} is a matrix of piecewise-continuous

\ddagger The notation u_o is used here because this will later be interpreted as a u coordinate associated with some observer. Similarly, u_s is often interpreted below as a u coordinate associated with a source (not necessarily when that source is observable at any observation point with $u = u_o$). Later, the notation u_e is used to denote a u coordinate associated with the emission of light from a source.

functions and that \mathbf{A} and \mathbf{B} are at least C^1 (and piecewise- C^2) in both of their arguments.

Given any two solutions \mathbf{E}_1 and \mathbf{E}_2 to (9), it is easily verified that their Wronskian is conserved:

$$\mathbf{E}_1^\top \dot{\mathbf{E}}_2 - \dot{\mathbf{E}}_1^\top \mathbf{E}_2 = \text{constant}. \quad (12)$$

Here, \top denotes a matrix transpose. Applying this formula with $\mathbf{E}_1 \rightarrow \mathbf{A}$ and $\mathbf{E}_2 \rightarrow \mathbf{B}$ shows that

$$\mathbf{A}^\top \partial_{(1)} \mathbf{B} - \partial_{(1)} \mathbf{A}^\top \mathbf{B} = \delta. \quad (13)$$

Neither \mathbf{A} nor \mathbf{B} are necessarily symmetric matrices. Nevertheless, (12) and (13) may be used to show that the products

$$\mathbf{A}^\top \partial_{(1)} \mathbf{A}, \quad \partial_{(1)} \mathbf{A} \mathbf{A}^{-1}, \quad \mathbf{B}^\top \partial_{(1)} \mathbf{B}, \quad \partial_{(1)} \mathbf{B} \mathbf{B}^{-1}, \quad \mathbf{B} \mathbf{A}^\top, \quad \mathbf{B}^{-1} \mathbf{A} \quad (14)$$

are symmetric wherever they exist [23]. Letting $\mathbf{E}_1 \rightarrow \mathbf{B}(\cdot, u_o)$ and $\mathbf{E}_2 \rightarrow \mathbf{B}(\cdot, u_s)$ in (12) shows that

$$\mathbf{B}(u_s, u_o) = -\mathbf{B}^\top(u_o, u_s) \quad (15)$$

for all $u_s, u_o \in \mathbb{R}$. This is essentially Etherington's reciprocity law [1, 3, 36, 37]. A similar calculation may be used to show that

$$\partial_{(1)} \mathbf{A}(u_s, u_o) = -\partial_{(1)} \mathbf{A}^\top(u_o, u_s) \quad (16)$$

as well.

It is sometimes useful to consider partial derivatives $\partial_{(2)}$ with respect to the second arguments of \mathbf{A} and \mathbf{B} . The resulting matrices remain solutions to (9). Comparing initial conditions shows that

$$\partial_{(2)} \mathbf{A}(u_s, u_o) = -\mathbf{B}(u_s, u_o) \mathbf{H}(u_o), \quad \partial_{(2)} \mathbf{B}(u_s, u_o) = -\mathbf{A}(u_s, u_o). \quad (17)$$

\mathbf{A} may therefore be derived from \mathbf{B} . The opposite is also true wherever $\det \mathbf{H} \neq 0$.

Geometrically, \mathbf{B} corresponds to the transverse coordinate components of a Jacobi propagator describing the evolution of deviation vectors along geodesics passing between different pairs of points [23]. \mathbf{B} is also related to image distortion. Up to an overall time dilation factor, it translates small differences in image position on an observer's sky to spatial deviations from a fiducial source point. In the language of [11], \mathbf{B} is proportional to the Jacobi map. The symmetric matrix $\partial_{(1)} \mathbf{B} \mathbf{B}^{-1}$ plays a similar role, but translates source separations to *emission* (rather than observation) angles. It is proportional to an object typically referred to as the optical deformation matrix. These statements are explained more fully in Sect. 4.3.

2.2. Conjugate pairs

It is often useful when working with plane wave spacetimes to consider hypersurfaces of "constant phase." Recalling the interpretation of the u coordinate as a phase, let

$$S_{u_o} := \{p \in M : u(p) = u_o\} \quad (18)$$

denote such a hypersurface. Given two points $p_s \in S_{u_s}$ and $p_o \in S_{u_o}$ (with $u_s \neq u_o$), define the multiplicity or "index" of these points to be [20]

$$I(p_s, p_o) := 2 - \text{rank } \mathbf{B}(u_s, u_o). \quad (19)$$

We also set $I(p_o, p_s) = 0$ whenever $u_o = u_s$. It follows from (15) that $I(p_s, p_o) = I(p_o, p_s)$. For most pairs of points, $I = 0$. Such pairs are said to be "disconjugate."

Pairs p_s, p_o satisfying $I(p_s, p_o) \neq 0$ are instead said to be conjugate with multiplicity $I(p_s, p_o)$. Similarly, we call the pairs S_{u_s}, S_{u_o} “conjugate hyperplanes” and the pairs of real numbers u_s, u_o “conjugate phases” when $\text{rank } \mathbf{B}(u_s, u_o) < 2$. Despite the appearance of (19), the index map $I : M \times M \rightarrow \{0, 1, 2\}$ describes phenomena which do not depend on any choice of coordinate system.

Conjugate pairs as described here are closely related to the conjugate points commonly considered in differential geometry and optics. In general, distinct points p_s and p_o on a given geodesic are said to be conjugate if and only if there exist nontrivial deviation vectors along that geodesic which vanish at both p_o and p_s . In plane wave spacetimes, this condition reduces to $I(p_s, p_o) > 0$. Defining the multiplicity of a pair of conjugate points to be the number of linearly independent deviation vectors which vanish at those points, that multiplicity is equal to $I(p_s, p_o)$. The concept of conjugacy associated with I does not, however, require the specification of any particular geodesic. It is uniquely defined even for pairs of points connected by multiple geodesics or by none. Indeed, these are the only cases where $I \neq 0$.

All strong lensing effects associated with plane wave spacetimes are related to the existence of conjugate hyperplanes. It follows from (19) that every pair of conjugate phases u_s, u_o satisfies

$$\det \mathbf{B}(u_s, u_o) = \det \mathbf{B}(u_o, u_s) = 0. \quad (20)$$

Finding conjugate pairs and their multiplicities may be viewed as a matter of direct computation once \mathbf{H} is specified. Alternatively, various Sturm-type comparison theorems can be used to make general statements regarding the existence and separations of conjugate pairs for various classes of plane wave. See, e.g., Chapt. XI of [38] for results relating to mathematical problems of this type and [20] for a specific application to “tame” plane wave spacetimes.

The qualitative structure of geodesics in plane wave spacetimes is closely related to the index I . First note that every disconjugate pair of points p_s, p_o is connected by exactly one geodesic. If two points are conjugate, the number of connecting geodesics is either zero or infinity. Sizes of geodesically connected regions may be summarized by

$$\dim[(\text{all points geodesically connected to } p_o) \cap S_{u_s}] + I(p_s, p_o) = 3. \quad (21)$$

A similar relation exists for null cones when $u_o \neq u_s$:

$$\dim[(\text{all points connected to } p_o \text{ via null geodesics}) \cap S_{u_s}] + I(p_s, p_o) = 2. \quad (22)$$

The latter result has been referred to as an “index theorem” in [20]. Eq. (21) implies that geodesics emanating from p_o and intersecting a hyperplane S_{u_s} that is disconjugate to S_{u_o} form a 3-dimensional region. Indeed, these geodesics fill the entire hyperplane. More interestingly, geodesics intersecting a conjugate hyperplane with multiplicity 1 fill only a 2-dimensional region on that 3-dimensional surface. Geodesics intersecting a hyperplane with multiplicity 2 form a line. Similarly, the null cone of a point reduces to a parabola on every hyperplane with multiplicity 1. A null cone intersecting a hyperplane with multiplicity 2 is focused to a single point. These two cases correspond to astigmatic and anastigmatic focusing, respectively. Anastigmatic focusing tends to be unstable in the sense that perturbations typically split a single multiplicity 2 focus into two closely-spaced foci each with multiplicity 1.

In many applications, there exists a preferred point p_o , or perhaps a preferred hyperplane S_{u_o} . p_o may, for example, represent the position of an observer at a particular time. Fixing this point, the set of hyperplanes conjugate to S_{u_o} divides a

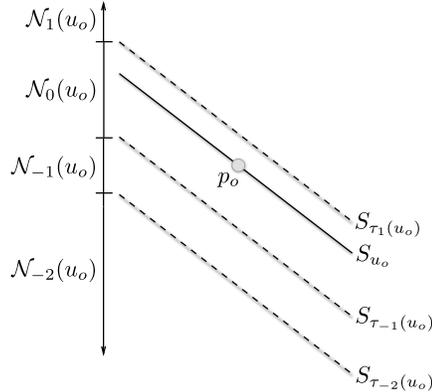


Figure 1: Given a preferred point p_o or a preferred $u = \text{constant}$ hyperplane S_{u_o} , plane wave spacetimes naturally divide into a number of open four-dimensional regions $\mathcal{N}_n(u_o)$. These regions are separated from each other by the hyperplanes $S_{\tau_n(u_o)}$ conjugate to S_{u_o} .

plane wave spacetime into a (possibly infinite) number of open regions $\mathcal{N}_n(u_o)$. Let $\mathcal{N}_0(u_o)$ denote the largest connected region containing S_{u_o} and excluding any portion of a hyperplane conjugate to S_{u_o} . If there exists a smallest $\tau_1(u_o) > u_o$ conjugate to u_o , the surface $S_{\tau_1(u_o)}$ is clearly contained in the boundary of $\mathcal{N}_0(u_o)$. $\mathcal{N}_1(u_o)$ may then be defined as the largest connected region which includes $S_{\tau_1(u_o)}$ as a boundary and contains points p_s satisfying $u_s > \tau_1(u_o)$ and $I(p_s, p_o) = 0$. This continues the spacetime “above” $S_{\tau_1(u_o)}$. Similar constructions may be used to define $\tau_n(u_o)$ and $\mathcal{N}_n(u_o)$ for values of n other than 1. See Fig. 1. Fixing a particular nonzero integer n and real number u_o , it is not necessary that $\tau_n(u_o)$ exist at all. In general, the domain of τ_n is an open subset of \mathbb{R} . This domain can be empty for some n .

Collecting all regions $\mathcal{N}_n(u_o)$ together, it is useful to define the open (and usually disconnected) set

$$\mathcal{N}(u_o) := \bigcup_n \mathcal{N}_n(u_o). \quad (23)$$

This includes the whole spacetime except for a set of measure zero formed from all of the hyperplanes conjugate to S_{u_o} :

$$\mathcal{N}(u_o) = M \setminus \left(\bigcup_n S_{\tau_n(u_o)} \right). \quad (24)$$

The geodesic uniqueness result described above can now be reduced to the statement that a point p_o is connected to another point $p_s \neq p_o$ by exactly one geodesic if and only if $p_s \in \mathcal{N}(u_o)$. In this sense, $\mathcal{N}(u_o)$ is a kind of “generalized normal neighborhood” for any p_o . Bitensors like Synge’s world function, the parallel propagator, and the van Vleck determinant may be defined unambiguously throughout $M \times \mathcal{N}$, which is “almost” $M \times M$. Unlike a true normal neighborhood, however, geodesics on M which connect points in $\mathcal{N}(u_o)$ needn’t lie entirely in this region.

2.3. Geodesics

Beyond the qualitative geodesic structure of plane wave spacetimes discussed above, it is not difficult to obtain explicit coordinate expressions for all geodesics. Let $\Gamma \subset M$ denote some geodesic and $\gamma : \mathbb{R} \rightarrow M$ an affine parametrization of it. The vector field $\ell^a = (\partial/\partial v)^a$ generating the characteristics of the gravitational wave is Killing, so $\dot{\gamma}^a \ell_a$ must be constant on Γ . If this constant vanishes, Γ is confined to a hypersurface of constant phase. Such a geodesic has the form of a (Euclidean) straight line in the coordinates (v, \mathbf{x}) .

Geodesics satisfying $\dot{\gamma}^a \ell_a \neq 0$ are more interesting. In these cases, the affine parameter can always be rescaled such that $\dot{\gamma}^a \ell_a = -1$. It is then possible to identify that parameter with the phase coordinate u . Doing, so $u(\gamma(u_s)) = u_s$ for all $u_s \in \mathbb{R}$. The spatial components $\boldsymbol{\gamma} := \mathbf{x}(\gamma)$ of any geodesic are fixed everywhere once $\boldsymbol{\gamma}_o = \boldsymbol{\gamma}(u_o)$ and $\dot{\boldsymbol{\gamma}}_o = \dot{\boldsymbol{\gamma}}(u_o)$ have been specified at some fiducial phase u_o . In terms of the matrices \mathbf{A} and \mathbf{B} defined in Sect. 2.1,

$$\boldsymbol{\gamma}(u_s) = \mathbf{A}(u_s, u_o)\boldsymbol{\gamma}_o + \mathbf{B}(u_s, u_o)\dot{\boldsymbol{\gamma}}_o. \quad (25)$$

The associated v coordinate of Γ may be efficiently derived using that fact that the vector field $(2v\partial_v + x^i\partial_i)^a$ is a homothety [23]. This implies that

$$v(\gamma(u_s)) = v(\gamma(u_o)) + \frac{1}{2}[\kappa_s(u_s - u_o) + \boldsymbol{\gamma}(u_s) \cdot \dot{\boldsymbol{\gamma}}(u_s) - \boldsymbol{\gamma}_o \cdot \dot{\boldsymbol{\gamma}}_o], \quad (26)$$

where

$$\kappa_s := -\dot{\gamma}_a \dot{\gamma}^a \quad (27)$$

is a constant.

κ_s is closely related to the conserved quantity on Γ associated with the Killing field ℓ^a . The (unit) 4-velocity U_s^a tangent to Γ is related to $\dot{\gamma}^a$ via

$$U_s^a = \frac{\dot{\gamma}^a}{\sqrt{\kappa_s}}. \quad (28)$$

It follows that

$$-\ell_a U_s^a = \frac{1}{\sqrt{\kappa_s}}. \quad (29)$$

If the spacetime is nearly flat (so $\mathbf{H} \approx 0$), κ_s reduces to a particle's specific energy minus its specific momentum in the direction of the gravitational wave. The limit $-\ell_a U^a \rightarrow 0$ (or $\kappa_s \rightarrow \infty$) may therefore be interpreted as an ultrarelativistic boost in the direction of the gravitational wave. By contrast, the limit $-\ell_a U^a \rightarrow \infty$ (or $\kappa_s \rightarrow 0$) corresponds to ultrarelativistic motion *against* the background wave.

Other Killing fields present in essentially all plane wave spacetimes may be written as

$$(x^i \dot{\Xi}_i) \ell^a + \Xi^i X_{(i)}^a, \quad (30)$$

where $\Xi_i = \Xi^i$ is any 2-vector with the form

$$\boldsymbol{\Xi} = \mathbf{A}(\cdot, u_o)\boldsymbol{\Xi}(u_o) + \mathbf{B}(\cdot, u_o)\dot{\boldsymbol{\Xi}}(u_o). \quad (31)$$

For each choice of u_o , there exists a four-parameter family of such vector fields. Each of these is associated with a conservation law. Two such conserved quantities may be summarized by

$$\mathcal{P}(u_o) := \frac{1}{\sqrt{\kappa_s}} [\mathbf{A}^\top(u_s, u_o)\dot{\boldsymbol{\gamma}}(u_s) - \partial_{(1)}\mathbf{A}^\top(u_s, u_o)\boldsymbol{\gamma}(u_s)]. \quad (32)$$

Fixing any u_o , this 2-vector is conserved in the sense that it is independent of u_s . Another conserved 2-vector may be defined by

$$\mathcal{C}(u_o) := \frac{1}{\sqrt{\kappa_s}} [\mathbf{B}^\Gamma(u_s, u_o) \dot{\gamma}(u_s) - \partial_{(1)} \mathbf{B}^\Gamma(u_s, u_o) \gamma(u_s)]. \quad (33)$$

In the weak field limit, \mathcal{P} corresponds to the specific momentum transverse to the gravitational wave. In this same context, \mathcal{C} may be interpreted as the conserved quantity associated with boosts transverse to the gravitational wave. It constrains transverse displacements. Note that both \mathcal{P} and \mathcal{C} depend on a choice of u_o . This is analogous to the choice of origin necessary to define angular momentum in elementary mechanics.

In stationary spacetimes, it is common to discuss various quantities related to gravitational lensing in terms of stationary observers (and often stationary sources). While plane wave spacetimes are not stationary, there does exist sufficient symmetry to define similarly preferred sources and observers. Two geodesics Γ and Γ' can be said to be “instantaneously comoving” at $u = u_o$ when

$$\kappa_s = \kappa'_s, \quad \mathcal{P}(u_o) = \mathcal{P}'(u_o). \quad (34)$$

These conditions imply that the 4-velocities of both geodesics are parallel-transported versions of each other on the constant-phase hyperplane S_{u_o} .

As implied by (21), bundles of geodesics are strongly focused on conjugate hyperplanes. Consider such a hyperplane associated with a phase $\tau_n(u_o)$ conjugate to u_o with multiplicity 1. It is clear from (19) that $\hat{\mathbf{B}}_n(u_o) := \mathbf{B}(\tau_n(u_o), u_o)$ is a matrix with rank 1. There therefore exists a unit 2-vector $\hat{\mathbf{q}}_n(u_o)$ such that

$$\hat{\mathbf{q}}_n^\Gamma(u_o) \hat{\mathbf{B}}_n(u_o) = 0. \quad (35)$$

$\hat{\mathbf{q}}_n(u_o)$ is unique up to sign. Choosing any $\hat{\mathbf{p}}_n(u_o)$ orthogonal to $\hat{\mathbf{q}}_n(u_o)$, the transverse spatial coordinates of all geodesics starting at a given point p_o focus to the line

$$\gamma(\tau_n) = \hat{\mathbf{A}}_n \gamma_o + w \hat{\mathbf{p}}_n \quad (36)$$

as they pass through $S_{\tau_n(u_o)}$. Here, w is any real number. All v coordinates may be reached on $S_{\tau_n(u_o)}$ by appropriate geodesics.

Geodesics starting at p_o and intersecting a hyperplane $S_{\tau_n(u_o)}$ with multiplicity 2 all focus to the single transverse position

$$\gamma(\tau_n) = \hat{\mathbf{A}}_n \gamma_o \quad (37)$$

as they pass through $S_{\tau_n(u_o)}$. As in the multiplicity 1 case, all values of v may be reached by appropriate geodesics. Eqs. (36) and (37) illustrate explicitly how conjugate hyperplanes with multiplicities 1 and 2 are associated with astigmatic and anastigmatic focusing, respectively. The former case involves focusing in only one transverse direction, while the latter case involves simultaneous focusing in both transverse directions.

2.4. Distances

As noted above, all pairs of points not lying on conjugate hyperplanes are connected by exactly one geodesic. There is therefore no ambiguity in ascribing geodesic distances to these pairs. In particular, Synge’s world function

$$\sigma(p_s, p_o) := \frac{1}{2} (\text{squared geodesic distance between } p_s \text{ and } p_o) \quad (38)$$

is well-defined whenever its arguments do not lie on conjugate hyperplanes [i.e., as long as $p_o \in \mathcal{N}(u_s)$ or $p_s \in \mathcal{N}(u_o)$]. Plane wave spacetimes constitute one of the few examples where σ is known essentially in closed form:

$$\begin{aligned} \sigma(p_s, p_o) = & \frac{1}{2}(u_s - u_o) \left[-2(v_s - v_o) + \mathbf{x}_s^\top \partial_{(1)} \mathbf{B}(u_s, u_o) \mathbf{B}^{-1}(u_s, u_o) \mathbf{x}_s \right. \\ & \left. + \mathbf{x}_o^\top \mathbf{B}^{-1}(u_s, u_o) \mathbf{A}(u_s, u_o) \mathbf{x}_o - 2\mathbf{x}_o^\top \mathbf{B}^{-1}(u_s, u_o) \mathbf{x}_s \right]. \end{aligned} \quad (39)$$

This is symmetric in its arguments: $\sigma(p_s, p_o) = \sigma(p_o, p_s)$.

The appearance of \mathbf{B}^{-1} in (39) indicates that σ tends to diverge when its arguments approach conjugate hyperplanes. More specifically, suppose that $u_s \approx \tau_n(u_o)$ and $I(\tau_n(u_o), u_o) = 1$. Then,

$$\sigma(p_s, p_o) \approx -\frac{1}{2} \left(\frac{u_o - u_s}{u_s - \tau_n(u_o)} \right) \left[\hat{\mathbf{q}}_n(u_o) \cdot (\mathbf{x}_s - \hat{\mathbf{A}}_n(u_o) \mathbf{x}_o) \right]^2 \quad (40)$$

is an asymptotic approximation for σ if the bracketed term on the right-hand side of this equation is nonzero [23]. It follows from (36) that this expression is valid only when there does not exist any geodesic passing from p_o to a point on $S_{\tau_n(u_o)}$ with the same transverse coordinates as p_s . The equivalent result if p_s is near a hyperplane $S_{\tau_n(u_o)}$ with multiplicity 2 is

$$\sigma(p_s, p_o) \approx -\frac{1}{2} \left(\frac{u_o - u_s}{u_s - \tau_n(u_o)} \right) |\mathbf{x}_s - \hat{\mathbf{A}}_n(u_o) \mathbf{x}_o|^2. \quad (41)$$

Again, this is valid only when there does not exist any geodesic passing from p_o to a point on $S_{\tau_n(u_o)}$ with the same transverse coordinates as p_s .

3. Image counting

One of the most basic questions that can be asked regarding a gravitational lens is the number of images that it produces of a particular source. Stated somewhat differently, how many future-directed null geodesics connect a given timelike curve (the source) to a particular spacetime event (the observer)? This may be answered using the geodesic structure of plane wave spacetimes summarized above.

To fix the notation, let p_o denote a fixed observation event and Γ the timelike worldline of a point source. Assume that Γ may be parametrized by an everywhere- C^1 function $\gamma : \mathbb{R} \rightarrow M$. The phase coordinate u serves as a useful “quasi-time” [21] for plane wave spacetimes[§], so let γ satisfy $u(\gamma(u_s)) = u_s$ for all $u_s \in \mathbb{R}$. The source’s worldline is not required to be geodesic.

This section establishes that under generic conditions, an ideal observer at p_o may see exactly one image of Γ from each of the “epochs” $\mathcal{N}_n(u_o)$ described in Sect. 2.2. Somewhat more precisely, there typically exists exactly one future-directed null geodesic from

$$\Gamma_n := \Gamma \cap \mathcal{N}_n(u_o) \quad (42)$$

to p_o for each $n \leq 0$ such that $\mathcal{N}_n(u_o)$ exists. See Fig. 2. Recall that the boundaries of $\mathcal{N}_n(u_o)$ depend only on the spacetime under consideration, and not at all on the

[§] u is not a “true” time function because $\nabla_a u$ is null. Generically, there do exist well-behaved functions $t : M \rightarrow \mathbb{R}$ where $\nabla_a t$ is everywhere timelike [39]. These are time functions in the usual sense, although they do not appear to simplify any computations performed here. Incidentally, the existence of such time functions implies that plane wave spacetimes are stably causal. Plane waves are not, however, globally hyperbolic [40]. Hypersurfaces of constant t are not Cauchy surfaces.

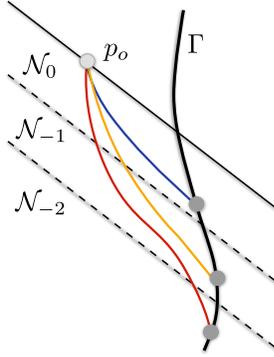


Figure 2: Schematic illustration of the gravitational lensing problem. Images correspond to future-directed null geodesics from a timelike source Γ to an observation event p_o . The arguments of Sect. 3 show that under generic conditions, exactly one image is emitted from each region $\mathcal{N}_n(u_o)$ lying before the observation event. The dashed lines correspond to constant- u hyperplanes conjugate to S_{u_o} .

behavior of any particular source. This allows generic bounds to be placed on time delays associated with the various images that may be observed.

Recall from Sect. 2.2 that each point in Γ_n is connected to p_o via exactly one (not necessarily null) geodesic. Synge's function $\sigma(\gamma(u_s), p_o)$ is therefore well-defined and explicitly given by (39) for all u_s such that $\gamma(u_s) \in \Gamma_n$. Consider instead the rescaled function

$$\Sigma_n(u_s) := \frac{\sigma(\gamma(u_s), p_o)}{u_o - u_s}. \quad (43)$$

If $n \leq 0$, the domain of $\Sigma_n(u_s)$ is equal to all u_s such that $\gamma(u_s) \in \Gamma_n$. If $n = 0$, we additionally suppose that $u_s < u_o$ for reasons of causality. Images of Γ_n produced by a plane gravitational wave correspond to the zeros of Σ_n .

Recalling the form (39) for σ , it is clear that Σ_n depends on Γ as well as the matrices \mathbf{A} , \mathbf{B} , \mathbf{B}^{-1} , and $\partial_{(1)}\mathbf{B}$. We have assumed in Sect. 2.1 that \mathbf{A} and \mathbf{B} are at least C^1 . So is γ . The definition of $\mathcal{N}_n(u_o)$ ensures that $\det \mathbf{B}(u_s, u_o) \neq 0$ everywhere $\Sigma_n(u_s)$ is defined. \mathbf{B}^{-1} is therefore C^1 and Σ_n is continuous.

Σ_n is also monotonic. To see this, note that (9), (13), (14), (39), and (43) may be used to show that

$$\dot{\Sigma}_n = \frac{1}{2} [-\dot{\gamma}_a \dot{\gamma}^a + |\dot{\gamma} + \mathbf{B}^{-\top}(\mathbf{x}_o - \partial_{(1)}\mathbf{B}^\top \gamma)|^2] > 0. \quad (44)$$

Σ_n is both continuous and monotonic, so at most one zero can exist for each n . This means that at most one image of a source may reach an observer from each epoch $\mathcal{N}_n(u_o)$. *Exactly* one such image exists if some u', u'' are known to satisfy

$$\Sigma_n(u') < 0, \quad \Sigma_n(u'') > 0. \quad (45)$$

Such bounds are easily established.

3.1. Lensing between conjugate hyperplanes

The simplest case to consider is one where $\mathcal{N}_n(u_o)$ lies “in between” hyperplanes conjugate to S_{u_o} . Suppose that both $\tau_n(u_o)$ and $\tau_{n-1}(u_o)$ exist for some $n < 0$. This

is true in Fig. 2 for $n = -1$. More generally, (40) and (41) imply that if there are no geodesics connecting p_o to either $\gamma(\tau_n(u_o))$ or $\gamma(\tau_{n-1}(u_o))$,

$$\lim_{u_s \rightarrow \tau_{n-1}^+(u_o)} \Sigma_n(u_s) = -\infty, \quad \lim_{u_s \rightarrow \tau_n^-(u_o)} \Sigma_n(u_s) = \infty. \quad (46)$$

It follows from these limits together with continuity that Σ_n is surjective on \mathbb{R} . Since this function is also monotonic, there must exist exactly one $u_e \in (\tau_{n-1}(u_o), \tau_n(u_o))$ such that p_o and $\gamma(u_e) \in \Gamma_n$ are connected by a null geodesic.

Under the same assumptions, projectiles moving on timelike geodesics may be thrown from $\gamma(u_s) \in \Gamma_n$ to p_o only if $u_s < u_e$. Choosing $u_s - \tau_{n-1}(u_o)$ to be sufficiently small (but positive), these projectiles can require an arbitrarily large amount of proper time to intersect p_o . It is somewhat curious that points $\gamma(u_s) \in \Gamma_n$ satisfying $u_s > u_e$ cannot be connected to p_o by any causal geodesic. Such points may, however, be reached by suitably accelerated curves which are everywhere causal.

3.2. The youngest image

Next, consider the case $n = 0$ when there exists at least one conjugate hyperplane in the observer's past (as occurs in the example illustrated by Fig. 2). The arguments given above imply that if $\gamma(\tau_{-1}(u_o))$ is geodesically disconnected from p_o ,

$$\lim_{u_s \rightarrow \tau_{-1}^+(u_o)} \Sigma_0(u_s) = -\infty. \quad (47)$$

The other boundary of the domain of Σ_0 occurs at $u_s = u_o$. Here,

$$\lim_{u_s \rightarrow u_o^-} \Sigma_0(u_s) = \frac{1}{2} \lim_{u_s \rightarrow u_o^-} \frac{|\mathbf{x}_o - \gamma(u_s)|^2}{u_o - u_s}. \quad (48)$$

This limit clearly tends to $+\infty$ if $\gamma(u_o) \neq \mathbf{x}_o$. Physically, $\gamma(u_o) \neq \mathbf{x}_o$ implies that p_o cannot be connected to $\gamma(u_o)$ by any null geodesic. Assuming that this is true, there must exist exactly one $u_e \in (\tau_{-1}(u_o), u_o)$ such that p_o and $\gamma(u_e) \in \Gamma_0$ are connected by a null geodesic.

3.3. The oldest image

The results of Sects. 3.1 and 3.2 typically suffice to describe the images formed in plane wave spacetimes containing an infinite number of conjugate hyperplanes in an observer's past. It is, however, important to consider cases where only a finite number of conjugate points exist. This occurs, e.g., for finite wavepackets where \mathbf{H} has compact support. It is also true of linearly polarized vacuum waves that are tame in the sense described in [20].

If there is at least one conjugate hyperplane in the observer's past, let N denote the smallest negative integer such that $\tau_N(u_o)$ exists. If there are no conjugate hyperplanes in the observer's past, set $N = 0$. The case illustrated in Fig. 2 corresponds to $N = -2$ if $\mathcal{N}_{-2}(u_o)$ extends into the infinite past. Regardless, we ask whether there exist any future-directed light rays from Γ_N to p_o . Unlike in the cases considered above, sources spend an infinite amount of proper time in $\mathcal{N}_N(u_o)$.

Even in flat spacetime, a source that accelerates for an infinitely long time may be causally-disconnected from certain observers via a Rindler horizon. Such phenomena can be ruled out here by supposing that there exists some finite constant $\kappa_{\min} > 0$ such that

$$-\dot{\gamma}_a \dot{\gamma}^a = 2\dot{\gamma}^a(u_s) \nabla_a v - [|\dot{\gamma}(u_s)|^2 + \gamma^\top(u_s) \mathbf{H}(u_s) \gamma(u_s)] > \kappa_{\min} \quad (49)$$

for all u_s less than some cutoff. As is clear from (27) and (28), the left-hand side of this inequality acts like the square root of a time dilation factor between the coordinate u and the source's proper time. Eq. (49) implies that the source's 4-velocity $U_s^a \propto \dot{\gamma}^a$ satisfies

$$0 < -\ell_a U_s^a < \frac{1}{\sqrt{\kappa_{\min}}} \quad (50)$$

sufficiently far in the past. It therefore excludes sources which experience arbitrarily large boosts against the background gravitational wave in the distant past. It is satisfied by, e.g., sources whose motion is geodesic sufficiently far into the past. Assuming that κ_{\min} exists, it is clear from (44) that u' may be chosen sufficiently small that

$$\Sigma_N(u') < 0. \quad (51)$$

Further assuming that $\gamma(\tau_N(u_o))$ is geodesically disconnected from p_o (if $N \neq 0$) or that $\gamma(u_o)$ is not null-separated from p_o (if $N = 0$), it follows that there exists exactly one u_e smaller than $\tau_N(u_o)$ (if $N \neq 0$) or u_o (if $N = 0$) such that $\gamma(u_e) \in \Gamma_N$ is connected to p_o via a future-directed null geodesic.

3.4. Total image count for generic sources

The results just described may be summarized as follows: Suppose that there does not exist any past-directed geodesic segment from p_o to Γ whose endpoints are conjugate (in the usual sense). Also assume that the source and observer are not instantaneously aligned with the background wave: $\mathbf{x}_o \neq \gamma(u_o)$. If there exists an ‘‘oldest’’ phase conjugate to u_o , further require that there be some $\kappa_{\min} > 0$ such that the source's motion is bounded by (50) sufficiently far in the past.

For each $n \leq 0$, these assumptions imply that an observer at p_o sees exactly one image of Γ as it appeared in $\mathcal{N}_n(u_o)$. This provides a strong bound on the possible emission times of different images. If $\mathcal{N}_N(u_o)$ exists for every negative integer N , an infinite number of images are formed. If, however, there is some smallest $N \leq 0$ such that $\mathcal{N}_N(u_o)$ exists, $|N| + 1$ images appear at p_o . Note that these results depend only on the waveform \mathbf{H} and the phase coordinate u_o associated with the observer. The total number of images is the same for all sources satisfying the hypotheses outlined above.

These hypotheses are generic. Recalling (21), the requirement that Γ exclude any points conjugate to p_o along a connecting geodesic is equivalent to demanding that the source's worldline avoid certain one- or two-dimensional subsets of the three-dimensional hyperplanes conjugate to S_{u_o} . Similarly, the assumption that $\mathbf{x}_o \neq \gamma(u_o)$ demands only that Γ avoid a certain line on S_{u_o} . Also note that the bound (50) on a source's asymptotic 4-velocity can fail to hold only for sources which accelerate for an infinitely long time. Violating any of these conditions for a fixed observer would require that a source's worldline be quite exceptional. Moreover, we now show that $|N| + 1$ images appear under even broader (but more difficult to state) conditions than those just discussed.

3.5. Non-generic imaging

Despite the comments made above, our assumptions on the behavior of a source's worldline can be violated in certain cases. Suppose, contrary to these assumptions,

that there exists at least one point $\gamma_e \in \Gamma$ which is conjugate to p_o along a geodesic connecting these two points. An infinite number of geodesics then pass between γ_e and p_o . If the connecting geodesics are null, γ_e lies on a caustic of the observer's past light cone. Continuous images of point sources – “Einstein rings” – are then formed at p_o (ignoring the associated breakdown of geometric optics). Such cases are not considered any further here. If conjugate points between the source and observer are associated with non-null geodesics, discrete images of Γ appear at p_o . In these cases, the methods used above are easily adapted to find how many images of Γ_n arrive at p_o . As already mentioned, there can be no more than one root for each Σ_n . Depending on the details of the system, the associated image from $\mathcal{N}_n(u_o)$ may or may not exist.

The case considered in Sect. 3.1 where Γ_n lies in between successive conjugate hyperplanes is the simplest to analyze. Assuming that Γ does not intersect the caustic of p_o , no images are formed of Γ_n if either $\gamma(\tau_n(u_o))$ is connected to p_o via a timelike geodesic or $\gamma(\tau_{n-1}(u_o))$ is connected to p_o via a spacelike geodesic. Otherwise, exactly one image exists from this region.

If there exists an oldest conjugate hyperplane $S_{\tau_N(u_o)}$ as described in Sect. 3.3, suppose that the source satisfies (49) for some $\kappa_{\min} > 0$. There are then zero images of Γ_N if $\gamma(\tau_N(u_o))$ is connected to p_o via a timelike geodesic. There is exactly one image if these points are either geodesically disconnected or are connected by a spacelike geodesic.

The last cases to consider concern images of Γ_0 . First suppose that $\mathbf{x}_o \neq \gamma(u_o)$. If there exists at least one conjugate hyperplane in the observer's past, one image is formed of Γ_0 if either $\gamma(\tau_{-1}(u_o))$ is geodesically disconnected from p_o or it is connected by a timelike geodesic. No images are formed if $\gamma(\tau_{-1}(u_o))$ and p_o are connected by a spacelike geodesic. If $\mathbf{x}_o \neq \gamma(u_o)$ and there are no conjugate points in the observer's past, condition (49) implies that there exists exactly one image of Γ_0 .

Cases where the source and observer are instantaneously aligned are more interesting. Suppose that $\mathbf{x}_o = \gamma(u_o)$. There then exists one image of Γ_0 with $u_e = u_o$. Recalling that u_o is not in the domain of Σ_0 , it is possible for a second image to be emitted from Γ_0 if $\Sigma_0 = 0$ somewhere. This may be seen by noting that

$$\lim_{u_s \rightarrow u_o^-} \Sigma_0 = v(\gamma(u_o)) - v_o. \quad (52)$$

Two images of Γ_0 can therefore exist when $v(\gamma(u_o)) > v_o$ and $\mathbf{x}_o = \gamma(u_o)$.

If a source includes points which are conjugate to the observer (in the ordinary sense), there is no simple result for the total number of images formed. Nevertheless, it is always possible to say that the total number of images is less than or equal to $|N| + 2$ if the source does not intersect a caustic of the observer's light cone.

4. Properties of lensed images

Plane wave spacetimes typically produce multiple images of each source. Even for sources whose intrinsic properties remain constant, these images can appear with different spectra, brightnesses, etc. We now compute these properties for generic configurations satisfying the hypotheses summarized in Sect. 3.4.

For each image of a timelike worldline Γ seen at p_o , there is an associated null geodesic segment connecting p_o to an appropriate emission point $\gamma_e = \gamma(u_e) \in \Gamma$. In terms of Synge's function (38), these points satisfy

$$\sigma(\gamma_e, p_o) = 0. \quad (53)$$

First derivatives of σ are always tangent to the connecting light ray. In particular, the vector

$$r_o^a := -\frac{\nabla^a \sigma(\gamma_e, p_o)}{u_o - u_e} \quad (54)$$

at p_o points along the geodesic which eventually intersects γ_e (and is therefore past-directed). The derivative operator here is understood to act on the second argument of σ . Also note that r_o^a is normalized such that $\ell_a r_o^a = 1$. Parallel-transporting r_o^a to the observation point yields

$$r_e^a = \frac{\nabla^a \sigma(\gamma_e, p_o)}{u_o - u_e}. \quad (55)$$

The derivative operator in this equation is understood to act on the first argument of σ . Both r_o^a and r_e^a may be viewed as (dimensionless) separation vectors between p_o and γ_e .

Eq. (39) and the various identities of Sect. 2.1 may be used to compute the explicit coordinate components of r_e^a and r_o^a . Components transverse to the direction of wave propagation are

$$\mathbf{r}_e = \mathbf{B}^{-\top}(u_e, u_o) [\mathbf{x}_o - \partial_{(1)} \mathbf{B}^\top(u_e, u_o) \boldsymbol{\gamma}_e], \quad (56)$$

where $\mathbf{x}_o = \mathbf{x}(p_o)$ and $\boldsymbol{\gamma}_e = \mathbf{x}(\gamma(u_e))$ denote the transverse coordinates of the observer and source. A similar calculation shows that

$$\mathbf{r}_o = \mathbf{B}^{-1}(u_e, u_o) [\mathbf{A}(u_e, u_o) \mathbf{x}_o - \boldsymbol{\gamma}_e]. \quad (57)$$

\mathbf{r}_o and \mathbf{r}_e are related via

$$\mathbf{r}_e = \partial_{(1)} \mathbf{B}(u_e, u_o) \mathbf{r}_o - \partial_{(1)} \mathbf{A}(u_e, u_o) \mathbf{x}_o. \quad (58)$$

Much of the discussion below considers sources moving on geodesics. In these cases, use of (25) shows that

$$\mathbf{r}_o = \mathbf{B}^{-1}(u_e, u_o) \mathbf{A}(u_e, u_o) \delta \mathbf{x}_o - \dot{\boldsymbol{\gamma}}_o. \quad (59)$$

Here, $\delta \mathbf{x}_o := \mathbf{x}_o - \boldsymbol{\gamma}(u_o) = \mathbf{x}_o - \boldsymbol{\gamma}_o$. The various identities involving \mathbf{A} and \mathbf{B} discussed in Sect. 2.1 may also be used to reduce the imaging condition (53) to

$$\kappa_s(u_o - u_e) = 2(\dot{\boldsymbol{\gamma}}_o \cdot \delta \mathbf{x}_o - \delta v_o) - \delta \mathbf{x}_o^\top \mathbf{B}^{-1}(u_e, u_o) \mathbf{A}(u_e, u_o) \delta \mathbf{x}_o. \quad (60)$$

Here, $\delta v_o := v(p_o) - v(\gamma_o)$. Eq. (60) is a nonlinear relation for the emission ‘‘time’’ u_e in terms of the observer’s position p_o and the parameters $\boldsymbol{\gamma}_o$, $\dot{\boldsymbol{\gamma}}_o$, $v(\boldsymbol{\gamma}_o)$, κ_s describing the source’s worldline. As discussed in Sect. 3, there can be many solutions to (60). These correspond to different images.

Neither u_e nor r_o^a depends on the observer’s motion. Nevertheless, redshifts and angles on the observer’s sky do depend on that motion (as is true even in flat spacetime). It is often useful to minimize this dependency by supposing that the observer is instantaneously comoving with the source. Following (34), this is taken to mean that the unit 4-velocities U_s^a, U_o^a of the source and observer on S_{u_o} satisfy

$$\ell_a U_s^a(u_o) = \ell_a U_o^a = -\frac{1}{\sqrt{\kappa_s}}, \quad (61)$$

$$\mathcal{P}_s(u_o) = \mathbf{U}_s(u_o) = \mathbf{U}_o = \frac{\dot{\boldsymbol{\gamma}}(u_o)}{\sqrt{\kappa_s}} = \frac{\dot{\mathbf{x}}_o}{\sqrt{\kappa_s}}. \quad (62)$$

Recall from Sect. 2.3 that the ‘‘transverse momentum’’ $\mathcal{P}_s(u_o)$ is generated by contracting $U_s^a(u_o)$ with the two Killing fields equal to $X_{(\hat{i})}^a$ at p_o and having vanishing first derivative at that point. Also note that (62) implicitly defines an instantaneous observer velocity $\dot{x}_o^a = \sqrt{\kappa_s} U_o^a$ normalized (like $\dot{\boldsymbol{\gamma}}^a$) such that $\ell_a \dot{x}_o^a = -1$.

4.1. Frequency shifts

Gravitational lenses typically discussed in astrophysics involve nearly-Newtonian mass distributions which may be regarded as approximately stationary (at least on sub-cosmological timescales). If both a source and an observer are sufficiently far from such a lens, there can be no significant redshift or blueshift except from relative motion. Roughly speaking, a light ray falling into any stationary gravitational potential must climb out of that same potential. This result breaks down if light passes through non-stationary regions of spacetime. Indeed, plane wave spacetimes may produce images with significant frequency shifts [19].

Consider an approximately monochromatic beam of light emitted from γ_e and received at p_o . A future-directed tangent vector $k_e^a \propto -r_e^a$ to the emitted light ray may always be chosen such that

$$\omega_e = -k_e \cdot U_e \quad (63)$$

is the angular frequency of the light as seen by its source. The frequency ω_o of this same light ray as measured by an observer at p_o is $-k_o \cdot U_o$, where k_o^a is equal to k_e^a parallel transported from the source to the observer. The observed and emitted frequencies are therefore related by

$$\frac{\omega_o}{\omega_e} = \frac{U_o \cdot k_o}{U_e \cdot k_e} = \frac{U_o \cdot r_o}{U_e \cdot r_e} = \sqrt{\frac{\kappa_e}{\kappa_o}} \left(\frac{\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}_o|^2}{\kappa_e + |\dot{\gamma}_e + \mathbf{r}_e|^2} \right). \quad (64)$$

Here, $\kappa_e := 1/(\ell \cdot U_s(u_e))^2$ and $\kappa_o := 1/(\ell \cdot U_o)^2$. The 2-vectors \mathbf{r}_e and \mathbf{r}_o appearing here are determined by the source and observer positions via (39), (53), (56), and (57). The resulting frequency shift is valid for all emission points not contained in a caustic of p_o .

Now suppose that a source moves on a geodesic and that the observer is instantaneously comoving with this geodesic in the sense of (61) and (62). Then $\kappa_e = \kappa_s = 1/(\ell \cdot U_s)^2$ doesn't depend on which image is chosen. Eqs. (58), (59), and the symmetry of $\mathbf{B}^{-1}\mathbf{A}$ may be used to rewrite (64) as

$$\frac{\omega_o}{\omega_e} = 1 + (\mathbf{B}^{-\top} \delta \mathbf{x}_o)^\top \left(\frac{\mathbf{A}\mathbf{A}^\top - \boldsymbol{\delta}}{\kappa_s + |\mathbf{B}^{-\top} \delta \mathbf{x}_o|^2} \right) (\mathbf{B}^{-\top} \delta \mathbf{x}_o). \quad (65)$$

The matrix in parentheses on the right hand side of this equation acts like a metric for the ‘‘separation’’ 2-vector $\mathbf{B}^{-\top}(u_e, u_o)[\mathbf{x}_o - \gamma(u_o)]$. If both eigenvalues of $\mathbf{A}(u_e, u_s)\mathbf{A}^\top(u_e, u_s) - \boldsymbol{\delta}$ are negative, the source is necessarily redshifted. Conversely, sources are always blueshifted when this matrix is positive definite. If $\mathbf{A}\mathbf{A}^\top - \boldsymbol{\delta}$ has both positive and negative eigenvalues, the sign of the frequency difference depends on the direction of $\mathbf{B}^{-\top} \delta \mathbf{x}_o$. For special configurations, there is no frequency shift at all.

4.2. Angles

Various images formed from a single source appear at different points on an observer's sky. Like redshifts, the relative angles between images change depending on an observer's 4-velocity U_o^a . The angle θ between two images arriving at p_o with tangents r_o^a and $r_o'^a$ is

$$\cos \theta = \frac{(g_{ab} + U_{o,a}U_{o,b})r_o^a r_o'^b}{(U_o \cdot r_o)(U_o \cdot r_o')} = 1 + \frac{r_o \cdot r_o'}{(U_o \cdot r_o)(U_o \cdot r_o')}. \quad (66)$$

Simplifying,

$$\cos \theta = 1 - \frac{2\kappa_o |\mathbf{r}_o - \mathbf{r}'_o|^2}{(\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}_o|^2)(\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}'_o|^2)}. \quad (67)$$

This expression is valid for arbitrary source and observer configurations. Specializing to geodesic sources and comoving observers,

$$\cos \theta = 1 - \frac{2\kappa_s |(\mathbf{B}^{-1}\mathbf{A} - \mathbf{B}'^{-1}\mathbf{A}')\delta\mathbf{x}_o|^2}{(\kappa_s + |\mathbf{B}^{-1}\mathbf{A}\delta\mathbf{x}_o|^2)(\kappa_s + |\mathbf{B}'^{-1}\mathbf{A}'\delta\mathbf{x}_o|^2)}. \quad (68)$$

Here, $\mathbf{A} = \mathbf{A}(u_e, u_o)$ and $\mathbf{A}' = \mathbf{A}(u'_e, u_o)$. It is evident that angles are largely controlled by the difference between $\mathbf{B}^{-1}\mathbf{A}$ at the two emission times.

Another interesting angle to consider is the observed separation ψ between a single image (emitted at γ_e) and a generator ℓ^a of the background gravitational wave. For arbitrarily moving source and observer configurations,

$$\cos \psi = 1 - \frac{2\kappa_o}{\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}_o|^2}. \quad (69)$$

For observers comoving with geodesic sources,

$$\cos \psi = 1 - \frac{2\kappa_s}{\kappa_s + |\mathbf{B}^{-1}\mathbf{A}\delta\mathbf{x}_o|^2}. \quad (70)$$

This may be used to rewrite the angle θ between two different images partially in terms of the angles ψ and ψ' those images make with ℓ^a . Using (68),

$$\cos \theta = \cos \psi \cos \psi' + \frac{(\mathbf{B}^{-1}\mathbf{A}\delta\mathbf{x}_o) \cdot (\mathbf{B}'^{-1}\mathbf{A}'\delta\mathbf{x}_o)}{|\mathbf{B}^{-1}\mathbf{A}\delta\mathbf{x}_o||\mathbf{B}'^{-1}\mathbf{A}'\delta\mathbf{x}_o|} \sin \psi \sin \psi'. \quad (71)$$

Similarly, the frequency shift (65) of an individual image may be rewritten as

$$\frac{\omega_o}{\omega_e} = \frac{\kappa_s \csc^2(\psi/2)}{\kappa_s + |\mathbf{B}^{-1}\delta\mathbf{x}_o|^2}. \quad (72)$$

It is evident from this equation that images which appear highly blueshifted to comoving observers must satisfy $\psi \approx 0$.

4.3. Image distortion and magnification

Thus far, all sources here have been modelled as though they were confined to timelike worldlines. Real objects are not pointlike, however. They form extended worldtubes in spacetime. Images of such worldtubes form null geodesic congruences which converge on p_o . These images can be significantly distorted by the curvature of spacetime. It is simplest to quantify such distortions by first fixing a particular null geodesic \mathcal{Z} passing between some part of the source and p_o . Precisely which geodesic is chosen is not important. \mathcal{Z} serves only as an origin from which to discuss nearby light rays connecting p_o to other points in the source. Once this origin has been fixed, the image of an extended source may be described entirely using deviation vectors on \mathcal{Z} (at least for sufficiently small sources). See Fig. 3.

Deviation vectors (or Jacobi fields) satisfy the geodesic deviation (or Jacobi) equation along \mathcal{Z} . Letting r^a denote the past-directed null vector tangent to \mathcal{Z} and obtained by parallel-transporting r^a_o from p_o , every deviation vector ξ^a is a solution to

$$r^b \nabla_b (r^c \nabla_c \xi^a) = R^a{}_{bcd} r^b r^c \xi^d. \quad (73)$$

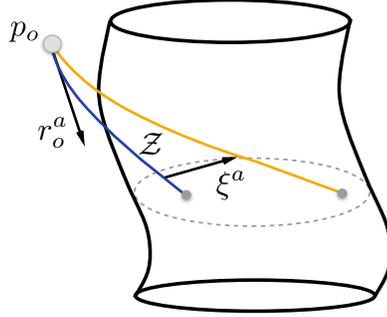


Figure 3: Imaging for an extended source. The fiducial light ray \mathcal{Z} is indicated in blue together with another light ray separated from it by a deviation vector ξ^a . The vector r_o^a is also drawn. This is tangent to \mathcal{Z} at the observation point p_o .

This equation is linear, so ξ^a must depend linearly on initial data. In particular, all deviation vectors can be written as linear combinations of appropriate bitensors contracted into the initial data $\parallel \xi^A(u_o)$ and $\dot{\xi}^A(u_o)$. All light rays observed at p_o must necessarily intersect that point, so it suffices to set $\xi^A(u_o) = 0$. The first derivative of a deviation vector at p_o describes an angular deviation between one point of an image and the center associated with \mathcal{Z} . We therefore consider deviation vectors ξ^a with the form

$$\xi^a = B^a_{\ A}(\cdot, u_o) \dot{\xi}^A(u_o). \quad (74)$$

$B^a_{\ A}$ is known as a Jacobi propagator. It satisfies the Jacobi equation

$$r^b \nabla_b (r^c \nabla_c B^a_{\ A}) = R^a_{\ bcd} r^b r^c B^d_{\ A} \quad (75)$$

along \mathcal{Z} together with the initial conditions

$$\lim_{u_s \rightarrow u_o} B^a_{\ A}(u_s, u_o) = 0, \quad \lim_{u_s \rightarrow u_o} r^b \nabla_b B^a_{\ A}(u_s, u_o) = \delta^a_A. \quad (76)$$

Note that $B^a_{\ A}$ is a bitensor. It maps vectors at p_o into vectors at others points on \mathcal{Z} . The transverse components of $B^a_{\ A}$ are

$$B_{aA} X^a_{(i)} X^A_{(j)} = (\mathbf{B})_{ij}, \quad (77)$$

where \mathbf{B} is the matrix defined in Sect. 2.1. Other components of $B^a_{\ A}$ may be deduced from the eigenvector relations [23]

$$B^a_{\ A}(u_s, u_o) r_o^A = (u_s - u_o) r^a, \quad r^a B_a^{\ A}(u_s, u_o) = (u_s - u_o) r_o^A, \quad (78)$$

$$B^a_{\ A}(u_s, u_o) \ell^A = (u_s - u_o) \ell^a, \quad \ell_a B^a_{\ A}(u_s, u_o) = (u_s - u_o) \ell_A. \quad (79)$$

All parts of an image must arrive at an observer along null geodesics. Additionally, an observer with 4-velocity U_o^a can only measure angles of vectors orthogonal to U_o^a . It therefore suffices to restrict attention to deviation vectors satisfying

$$r_o \cdot \dot{\xi}(u_o) = U_o \cdot \dot{\xi}(u_o) = 0 \quad (80)$$

at p_o . These constraints restrict all interesting initial data to a two dimensional space. The orthonormal vectors

$$e^A_{(i)} = X^A_{(i)} - r_o^i \ell^A - 2 \left(\frac{r_o^i + \dot{x}_o^i}{\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}_o|^2} \right) r_o^A \quad (81)$$

\parallel Capital letters are used in this subsection to denote abstract indices associated with the observation point. This is done to avoid confusion when writing down two-point tensors such as $B^a_{\ A}$ [see (74)].

form a basis for this space at p_o . They satisfy

$$e_{(i)} \cdot e_{(j)} = \delta_{ij}, \quad r_o \cdot e_{(i)} = U_o \cdot e_{(i)} = 0. \quad (82)$$

Parallel-transporting $e_{(i)}^A$ to another point on \mathcal{Z} yields

$$e_{(i)}^a = X_{(i)}^a - r^i \ell^a - 2 \left(\frac{r_o^i + \dot{x}_o^i}{\kappa_o + |\dot{\mathbf{x}}_o + \mathbf{r}_o|^2} \right) r^a. \quad (83)$$

This results forms a Sachs basis [1] on \mathcal{Z} . Initial data appearing in (74) must be of the form

$$\dot{\xi}^A(u_o) = (U_o \cdot r_o) \alpha^{(i)} e_{(i)}^A, \quad (84)$$

where α is an unconstrained 2-vector. The factor $U_o \cdot r_o > 0$ is included here so that the α are directly related to angles on an observer's sky. A sufficiently small image may be described by a suitable set of 2-vectors α representing the angular locations of each portion of the image with respect to the center defined by \mathcal{Z} .

Each α may be translated into a physical displacement using (74) and (84). First note that for every particular α , (78) implies that $r \cdot \xi = 0$ throughout \mathcal{Z} . Indeed, ξ^a is always a linear combination of the $e_{(i)}^a$ together with r^a . Components of ξ^a proportional to r^a are physically irrelevant, so we consider only the Sachs components

$$\xi_{(i)} := e_{(i)} \cdot \xi = [(U_o \cdot r_o) B_{aA} e_{(i)}^a e_{(j)}^A] \alpha^{(j)}. \quad (85)$$

Defining the 2×2 matrix

$$D_{(i)(j)}(u_s, u_o) := (U_o \cdot r_o) B_{aA}(u_s, u_o) e_{(i)}^a(u_s) e_{(j)}^A(u_o), \quad (86)$$

it is then clear that $\xi = \mathbf{D}\alpha$ for any α . \mathbf{D} is referred to as the Jacobi matrix or Jacobi map [1, 11]. Using (77)-(79), (81), (83), and (86),

$$\mathbf{D}(u_s, u_o) = (U_o \cdot r_o) \mathbf{B}(u_s, u_o). \quad (87)$$

If a source moves on a geodesic which is instantaneously comoving with the observer,

$$U_o \cdot r_o = \frac{\kappa_s + |\mathbf{B}^{-1} \mathbf{A} \delta \mathbf{x}_o|^2}{2\sqrt{\kappa_s}}. \quad (88)$$

This discussion implies that a portion of an image with angular separation $\epsilon\alpha$ from the fiducial direction associated with r_o^a is spatially separated from the fiducial emission point $\gamma_e \in \mathcal{Z}$ by

$$\epsilon \xi(u_e) = (U_o \cdot r_o) \mathbf{B}(u_e, u_o) (\epsilon \alpha). \quad (89)$$

The factors of $\epsilon \ll 1$ have been introduced here to emphasize that this description is valid only for infinitesimal deviations. Regardless, (89) shows that up to the time dilation factor $(U_o \cdot r_o)$, the matrix \mathbf{B} central to all aspects of plane wave geometry may be physically interpreted as a transformation converting infinitesimal angles on the vertex of a light cone into infinitesimal separations elsewhere on that light cone. \mathbf{B} depends only on the u coordinates of the source and emission points, and not on any other aspects of the physical configuration. It may be computed for all possible observer-source pairs directly from the wave profile \mathbf{H} .

Angles of emission (as opposed to observation) of the various light rays travelling from the source to the observer may be found by differentiating (89) and applying the appropriate time dilation factor:

$$\frac{\dot{\xi}(u_e)}{U_e \cdot r_e} = \left(\frac{U_o \cdot r_o}{U_e \cdot r_e} \right) \partial_{(1)} \mathbf{B}(u_e, u_o) \alpha = \left(\frac{\omega_o}{\omega_e} \right) \partial_{(1)} \mathbf{B}(u_e, u_o) \alpha. \quad (90)$$

The last equality here makes use of (64). Applying (89) shows that

$$\dot{\xi} = \partial_{(1)} \mathbf{B} \mathbf{B}^{-1} \xi. \quad (91)$$

The symmetric matrix $\partial_{(1)} \mathbf{B} \mathbf{B}^{-1} / (U_e \cdot r_e)$ therefore converts spatial locations to emission angles within the source (with the constraint that all light rays intersect p_o). It is referred to as the optical deformation matrix [1, 11].

Eq. (89) implies that there is a sense in which circles on the observer's sky correspond to ellipses near γ_e . This deformation may be parametrized by performing a polar decomposition on $\mathbf{D}(u_e, u_o)$:

$$\mathbf{D} = \mathbf{R}_\beta^\top \begin{pmatrix} D_+ & 0 \\ 0 & D_- \end{pmatrix} \mathbf{R}_\chi. \quad (92)$$

Here, \mathbf{R}_β and \mathbf{R}_χ represent rotation matrices through some angles β and χ . The ratio D_+/D_- is related to the ellipticity of the aforementioned ellipse. χ represents the angle between the principal axes of that ellipse and the Sachs basis. D_\pm and χ are referred to as shape parameters [1, 41].

Recalling that polarization vectors are parallel-transported in the geometric optics approximation [3], any polarization vector must have Sachs components which are constant along \mathcal{Z} . In principle, the angle χ might therefore be measured by comparing the relative ‘‘rotation’’ between an object's observed shape and an appropriate polarization angle [1, 42]. For linearly polarized waves where \mathbf{H} can be made diagonal by an appropriate coordinate choice, $\chi = 0$ with respect to this coordinate system and the basis (83). It is shown in Sect. 6 that χ also vanishes in a natural way for all sufficiently weak wavepackets which are nonzero only for short times.

Eq. (89) implies that \mathbf{D} converts angles at the observer to separations within the source. The determinant of \mathbf{D} must therefore relate solid angles at p_o to physical areas near γ_e :

$$\frac{d\mathcal{A}}{d\Omega} = |\det \mathbf{D}| = |D_+ D_-| = (U_o \cdot r_o)^2 |\det \mathbf{B}|. \quad (93)$$

It follows that

$$d_{\text{ang}} := \sqrt{|\det \mathbf{D}|} = (U_o \cdot r_o) \sqrt{|\det \mathbf{B}|} \quad (94)$$

may be interpreted as an ‘‘angular diameter distance.’’ Absolute value signs are necessary here because $\det \mathbf{B}$ changes sign after each pass through a conjugate hyperplane with multiplicity 1. Physically, such sign changes represent parity inversions of the resulting image. Note that d_{area} does not necessarily increase monotonically with the age of an image (as computed using the source's proper time). Closely related to the angular diameter distance is the luminosity distance

$$d_{\text{lum}} := (\omega_o/\omega_e)^{-2} d_{\text{ang}}. \quad (95)$$

One factor of ω_o/ω_e arises here from considering light cones emanating from the source instead of the observer. The other factor of ω/ω_e is related to the energy change associated with frequency shifts.

5. Symmetric plane waves

Now that various optical quantities have been computed for general plane wave spacetimes, we consider their application to special classes of plane waves. The simplest nontrivial class is composed of the symmetric waves. These are locally

symmetric in the sense that $\nabla_a R_{bcd}{}^f = 0$. It follows from (3) that symmetric plane waves must have constant waveforms. Also note that $(\partial/\partial u)^a$ is Killing in these examples [as well as $\ell^a = (\partial/\partial v)^a$, which is Killing in all plane wave spacetimes]. Particular symmetric plane waves may be specified entirely by the (constant) eigenvalues of \mathbf{H} .

Recalling the decomposition (6) of \mathbf{H} into h_+ , h_\times and h_\parallel , a coordinate rotation may always be used to set $h_\times = 0$ for symmetric waves. It is then evident that the two eigenvalues of \mathbf{H} are given by $\pm h_+ - h_\parallel$. It is always possible to set

$$\mathbf{H} = \begin{pmatrix} -h_1 & 0 \\ 0 & -h_2 \end{pmatrix}, \quad (96)$$

where

$$h_1 := h_\parallel + h_+, \quad h_2 := h_\parallel - h_+. \quad (97)$$

The weak energy condition implies that $h_\parallel \geq 0$, so at least one eigenvalue of \mathbf{H} must be negative. Up to a possible parity inversion, it is always possible to assume that $h_+ \geq 0$. Then,

$$h_1 > 0, \quad h_1 \geq |h_2|. \quad (98)$$

If the vacuum Einstein equation is imposed, $h_\parallel = 0$ and $h_2 = -h_1$. For conformally-flat geometries representing spacetimes associated with, e.g., pure electromagnetic plane waves, $h_+ = 0$ and $h_2 = h_1$. Other cases may be viewed as superpositions of gravitational and (“gravito”-)electromagnetic waves.

All symmetric waves produce an infinite number of images of almost every source. It is clear from (96) that these waves are also linearly polarized. The angles χ and β appearing in (92) therefore vanish when considering image deformations with respect to the Sachs basis (83). Other lensing properties depend on the sign of h_2 . We call the case $h_2 < 0$ “gravity-dominated” and the case $h_2 > 0$ “matter-dominated.”

5.1. Gravity-dominated symmetric waves

Consider symmetric plane wave spacetimes where $h_2 = h_\parallel - h_+ < 0$. Gravity-dominated waves such as these generalize the vacuum waves satisfying $h_1 = -h_2$. Symmetric vacuum waves arise from, e.g., the Penrose limit of a null geodesic orbiting a Schwarzschild black hole on the light ring.

For any gravity-dominated symmetric wave, the matrices \mathbf{A} and \mathbf{B} defined in Sect. 2.1 are

$$\mathbf{A}(u_s, u_o) = \begin{pmatrix} \cos h_1^{\frac{1}{2}}(u_s - u_o) & 0 \\ 0 & \cosh |h_2|^{\frac{1}{2}}(u_s - u_o) \end{pmatrix}, \quad (99)$$

$$\mathbf{B}(u_s, u_o) = \begin{pmatrix} h_1^{-\frac{1}{2}} \sin h_1^{\frac{1}{2}}(u_s - u_o) & 0 \\ 0 & |h_2|^{-\frac{1}{2}} \sinh |h_2|^{\frac{1}{2}}(u_s - u_o) \end{pmatrix}. \quad (100)$$

It is clear that $\det \mathbf{B}(\cdot, u_o)$ has an infinite number of zeros for any choice of u_o . Each of these zeros represents a phase conjugate to u_o . There are an infinite number of such phases in both the past and future of every observer. Explicitly,

$$\tau_n(u_o) = u_o + n\pi h_1^{-\frac{1}{2}} \quad (101)$$

for every nonzero integer n . It is evident from (19) that all of these phases have multiplicity 1. It follows from the discussion in Sect. 3 that under generic conditions,

an infinite number of images appear for almost every source. Given any integer $n < 0$, there is exactly one image of the source as it appeared in $\mathcal{N}_n(u_o)$. This corresponds to the region between $u = u_o + n\pi h_1^{-\frac{1}{2}}$ and $u = u_o + (n-1)\pi h_1^{-\frac{1}{2}}$. Note that $\det \mathbf{B}(\cdot, u_o)$ switches sign on each pass through a conjugate phase. The parity of an image emitted from $\mathcal{N}_n(u_o)$ is therefore opposite to the parity of an image emitted from $\mathcal{N}_{n-1}(u_o)$.

Specializing to cases where the source is a geodesic and the observer is instantaneously comoving with that source on S_{u_o} , some configurations lead to redshifts and others to blueshifts. Using (65),

$$\frac{\omega_o}{\omega_e} = 1 + \frac{\delta \mathbf{x}_o^\top \mathbf{H} \delta \mathbf{x}_o}{\kappa_s + |\mathbf{B}^{-\top} \delta \mathbf{x}_o|^2}. \quad (102)$$

An image is therefore redshifted if and only if

$$\delta \mathbf{x}_o^\top \mathbf{H} \delta \mathbf{x}_o = -h_1(\delta x_o^1)^2 + |h_2|(\delta x_o^2)^2 < 0. \quad (103)$$

It is blueshifted when $\delta \mathbf{x}_o^\top \mathbf{H} \delta \mathbf{x}_o > 0$. There is no frequency shift at all if

$$(\delta x_o^1 / \delta x_o^2)^2 = |h_2 / h_1|. \quad (104)$$

The direction of the frequency shift clearly depends only on the instantaneous orientation $\delta \mathbf{x}_o / |\delta \mathbf{x}_o|$ of the source and the observer on S_{u_o} . It does not, in particular, depend on which image is considered. All images of a particular source experience the same type of frequency shift.

Emission times u_e for an observer comoving with a geodesic source may be found by solving (60). For gravity-dominated symmetric waves, this equation reduces to

$$\begin{aligned} \kappa_s(u_o - u_e) &= 2(\dot{\gamma}_o \cdot \delta \mathbf{x}_o - \delta v_o) + h_1^{\frac{1}{2}}(\delta x_o^1)^2 \cot h_1^{\frac{1}{2}}(u_o - u_e) \\ &\quad + |h_2|^{\frac{1}{2}}(\delta x_o^2)^2 \coth |h_2|^{\frac{1}{2}}(u_o - u_e). \end{aligned} \quad (105)$$

If $-n \gg 1$, this implies that the image from $\mathcal{N}_n(u_o)$ must satisfy $\cot h_1^{\frac{1}{2}}(u_o - u_e) \gg 1$. Images emitted in the distant past are therefore emitted at phases u_e very nearly conjugate to u_o :

$$u_e \approx \tau_n(u_o) - \frac{h_1^{\frac{1}{2}}(\delta x_o^1)^2}{|n|\pi\kappa_s}. \quad (106)$$

Substituting this into (70) and (102) shows that very old images cluster near ℓ^a on the observer's sky and experience negligible frequency shifts:

$$\psi \propto |n|^{-1}, \quad |\omega_o/\omega_e - 1| \propto |n|^{-2}. \quad (107)$$

Old images of slightly extended sources are also highly distorted and demagnified. Their angular diameter and luminosity distances both scale like

$$d_{\text{ang}} \sim d_{\text{lum}} \propto |n|^{\frac{3}{2}} \exp\left(\frac{1}{2}\sqrt{|h_2/h_1|}|n|\pi\right). \quad (108)$$

Gravity-dominated symmetric waves therefore produce an infinite number of exponentially dimming images for almost every source.

5.2. Matter-dominated symmetric waves

Matter-dominated symmetric waves satisfying $h_2 > 0$ act somewhat differently than gravity-dominated waves. In these cases,

$$\mathbf{A}(u_s, u_o) = \begin{pmatrix} \cos h_1^{\frac{1}{2}}(u_s - u_o) & 0 \\ 0 & \cos h_2^{\frac{1}{2}}(u_s - u_o) \end{pmatrix}, \quad (109)$$

$$\mathbf{B}(u_s, u_o) = \begin{pmatrix} h_1^{-\frac{1}{2}} \sin h_1^{\frac{1}{2}}(u_s - u_o) & 0 \\ 0 & h_2^{-\frac{1}{2}} \sin h_2^{\frac{1}{2}}(u_s - u_o) \end{pmatrix}. \quad (110)$$

Phases conjugate to u_o occur at $u_o + n\pi h_1^{-\frac{1}{2}}$ and at $u_o + n'\pi h_2^{-\frac{1}{2}}$, where n, n' are any nonzero integers. If $\sqrt{h_1/h_2}$ is an irrational number, these two families of phases are distinct. Each conjugate pair then has multiplicity 1. If $\sqrt{h_1/h_2}$ is rational, some conjugate pairs have multiplicity 2. In the conformally-flat case where $h_1 = h_2$, all conjugate phases have multiplicity 2. In all other cases where $\sqrt{h_1/h_2}$ is rational, an infinite number of conjugate phases occur with each multiplicity.

Now consider a luminous source moving on a geodesic. If the source and observer are instantaneously comoving in the sense of (61) and (62), frequency shifts associated with each image are given by (65). Since

$$\mathbf{B}^{-1}(\mathbf{A}\mathbf{A}^\top - \delta)\mathbf{B}^{-\top} = \mathbf{H} \quad (111)$$

is negative-definite in this case, all images are redshifted.

If $-n \gg 1$, an image originating from $\mathcal{N}_n(u_o)$ must be emitted just before the source intersects $S_{\tau_n(u_o)}$. All such images cluster towards ℓ^a in the observer's sky and have negligible frequency shifts. Images emitted near conjugate hyperplanes with multiplicity 1 are highly distorted. Images emitted near conjugate hyperplanes with multiplicity 2 are not significantly distorted at all. In both cases, however, older images are dimmer (although the rate at which this occurs is much slower than for gravity-dominated symmetric waves).

Regardless of the sign of h_2 , the oldest images formed by symmetric plane wave spacetimes depend on the spacetime structure at arbitrarily large transverse distances. If the metric is modified so that the wave decays at large distances, only a finite number of images would be expected to appear.

6. Sandwich waves

Symmetric plane waves are mathematically simple, but are not reasonable models for gravitational radiation emitted from compact sources. More interesting are waves where \mathbf{H} is nonzero only for a finite time: Sandwich waves. Suppose, in particular, that there exists some $u_+ > 0$ such that $\mathbf{H}(u) = 0$ for all $u \notin [0, u_+]$. It follows from (3) that such spacetimes are locally flat whenever $u < 0$ or $u > u_+$. The curved region containing the gravitational wave is effectively sandwiched between the two null hyperplanes S_0 and S_{u_+} . Every timelike curve eventually passes entirely through such a wave.

Before an observer interacts with the wave, spacetime is flat and optics is trivial. The case $u_o > u_+$ where an observer has already passed through the wave is more interesting. In this case, \mathbf{A} and \mathbf{B} reduce to their flat space forms

$$\mathbf{A}(u_s, u_o) = \delta, \quad \mathbf{B}(u_s, u_o) = (u_s - u_o)\delta \quad (112)$$

when $u_s > u_+$. The forms of \mathbf{A} and \mathbf{B} inside the wave depend on the details of \mathbf{H} , and will not be discussed here. If $u_s < 0$, however, there always exist four constant 2×2 matrices $\boldsymbol{\alpha}$, $\dot{\boldsymbol{\alpha}}$, $\boldsymbol{\beta}$ and $\dot{\boldsymbol{\beta}}$ such that

$$\mathbf{B}(u_s, u_o) = (\boldsymbol{\alpha} + \dot{\boldsymbol{\alpha}}u_o) + (\boldsymbol{\beta} + \dot{\boldsymbol{\beta}}u_o)u_s. \quad (113)$$

Note that the dots on $\dot{\boldsymbol{\alpha}}$ and $\dot{\boldsymbol{\beta}}$ do not refer to derivatives in this case. It follows from (17) that $\mathbf{A}(u_s, u_o)$ is independent of u_o . Moreover,

$$\mathbf{A}(u_s, u_o) = -\dot{\boldsymbol{\alpha}} - \dot{\boldsymbol{\beta}}u_s. \quad (114)$$

If there were no wave at all, $\boldsymbol{\alpha} = \dot{\boldsymbol{\beta}} = 0$ and $-\dot{\boldsymbol{\alpha}} = \boldsymbol{\beta} = \boldsymbol{\delta}$.

In general, $\dot{\boldsymbol{\alpha}}$ and $\dot{\boldsymbol{\beta}}$ have a simple physical interpretation. If two geodesics are comoving and have a transverse separation $\delta\mathbf{x}_o$ when $u > u_+$, it follows from (25) that the transverse separation between these geodesics is $-\dot{\boldsymbol{\alpha}}\delta\mathbf{x}_o$ immediately before they interact with the wave at $u = 0$. Similarly, the relative transverse velocity of these geodesics is $-\dot{\boldsymbol{\beta}}\delta\mathbf{x}_o$ when $u < 0$.

Interpretations for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are somewhat less direct. Consider two geodesics which intersect at some time $u_o > u_+$ after the wave has passed, but which have a relative transverse velocity $\delta\dot{\mathbf{x}}_o$ at $u = u_o$. The difference in transverse velocities between these two geodesics is then $(\boldsymbol{\beta} + u_o\dot{\boldsymbol{\beta}})\delta\dot{\mathbf{x}}_o$ when $u < 0$. Similarly, the difference in the transverse positions of these geodesics is $(\boldsymbol{\alpha} + u_o\dot{\boldsymbol{\alpha}})\delta\dot{\mathbf{x}}_o$ at $u = 0$. It follows that $\boldsymbol{\alpha}$ controls shifts in position that are independent of the time u_o at which the two geodesics cross each other.

In principle, $\boldsymbol{\alpha}$, $\dot{\boldsymbol{\alpha}}$, $\boldsymbol{\beta}$, $\dot{\boldsymbol{\beta}}$ may all be found by solving (9) if \mathbf{H} is known. At first glance, this would appear to imply that $4 \cdot 4 = 16$ numbers are required to describe observations through a sandwich wave. The actual number of required parameters is somewhat less than this. First note that the Wronskian identity (13) implies that $\dot{\boldsymbol{\alpha}}^\top\dot{\boldsymbol{\beta}}$ is a symmetric matrix. It also implies that

$$\dot{\boldsymbol{\beta}}^\top\boldsymbol{\alpha} - \dot{\boldsymbol{\alpha}}^\top\boldsymbol{\beta} = \boldsymbol{\delta}. \quad (115)$$

Further simplifications arise by recalling from (14) that $\mathbf{B}\mathbf{A}^\top$ and $\mathbf{B}^\top\partial_{(1)}\mathbf{B}$ are symmetric. It follows that

$$\boldsymbol{\alpha}\dot{\boldsymbol{\alpha}}^\top, \quad \boldsymbol{\beta}\dot{\boldsymbol{\beta}}^\top, \quad \boldsymbol{\alpha}^\top\boldsymbol{\beta} \quad (116)$$

are symmetric as well. These expressions are completely general, and hold for any sandwich wave. They significantly constrain the number of independent parameters needed to specify \mathbf{A} and \mathbf{B} . Equivalently, they limit the number of parameters that must be extracted from \mathbf{H} .

It follows from the arguments of Sect. 3 that the number of images of a generic source observable in any plane wave spacetime is governed by the number of hypersurfaces conjugate to the $u = \text{constant}$ hypersurface S_{u_o} containing the observation event p_o . Continuing to assume that $u_o > u_+$, all phases conjugate to u_o must be smaller than u_+ . It follows from (20) and (113) that conjugate phases occurring before the wave may be found by solving

$$\det[(\boldsymbol{\alpha} + \dot{\boldsymbol{\alpha}}u_o) + (\boldsymbol{\beta} + \dot{\boldsymbol{\beta}}u_o)\tau] = 0 \quad (117)$$

for all $\tau < 0$. This equation is quadratic, so at most two solutions exist. It follows that an observer ahead of the wave may see at most three images of a source as it appeared behind the wave. There may also be at most one image of a source as it appeared ahead of the wave. In principle, any number of images may arise from inside the wave [where (113) is not valid] if \mathbf{H} is sufficiently large.

Conjugate phases found by solving (117) clearly depend on the observation time u_o . This means that $\tau = \tau(u_o)$. Less obviously, the *number* of conjugate phases can also depend on u_o . For an observer moving on a timelike worldline (where u_o increases monotonically), new conjugate phases – and therefore new images – sometimes appear at discrete times. These images correspond to observation times where (117) momentarily degenerates to a linear equation. New images can therefore arise when $u_o = \bar{u}_o$ and

$$\det(\boldsymbol{\beta} + \dot{\boldsymbol{\beta}}\bar{u}_o) = 0. \quad (118)$$

If $\det \dot{\boldsymbol{\beta}} \neq 0$, the two solutions to this equation are

$$\bar{u}_o = \frac{-[\text{Tr } \boldsymbol{\beta} \text{Tr } \dot{\boldsymbol{\beta}} - \text{Tr}(\boldsymbol{\beta}\dot{\boldsymbol{\beta}})] \pm \sqrt{[\text{Tr } \boldsymbol{\beta} \text{Tr } \dot{\boldsymbol{\beta}} - \text{Tr}(\boldsymbol{\beta}\dot{\boldsymbol{\beta}})]^2 - 4 \det \boldsymbol{\beta} \det \dot{\boldsymbol{\beta}}}}{2 \det \dot{\boldsymbol{\beta}}}. \quad (119)$$

Only solutions satisfying $\bar{u}_o > u_+ > 0$ are physically relevant. When a conjugate phase of this type first appears, it satisfies

$$\lim_{u_o \rightarrow \bar{u}_o^+} \tau(u_o) = -\infty. \quad (120)$$

The associated image therefore provides a picture of the infinitely distant past. Furthermore, an infinite amount of the source's history appears to the observer within a finite amount of proper time. This implies that new images are highly blueshifted.

No matter how long an observer waits, no conjugate phase can exceed u_+ . The emission time for an associated image might therefore be expected to tend towards a constant value as $u_o \rightarrow \infty$. This means that a very large amount of proper time at the observer corresponds to only a small amount of proper time at a source. Images which appear suddenly and are initially highly blueshifted become highly redshifted at late times.

Weak wavepackets

One important class of sandwich waves are those that are very weak and last only for a short time. In these cases, $\boldsymbol{\alpha}$, $\dot{\boldsymbol{\alpha}}$, $\boldsymbol{\beta}$, and $\dot{\boldsymbol{\beta}}$ may be expanded as integrals involving successively higher powers of \mathbf{H} . To lowest order in such a scheme, \mathbf{A} and \mathbf{B} are approximately unaffected by the wave while inside of it. Assuming that $-u_s, u_o \gg u_+$, the first corrections to this assumption are

$$\mathbf{A}(u_s, u_o) \approx \boldsymbol{\delta} - u_s \int_0^{u_+} dw \mathbf{H}(w), \quad (121)$$

$$\mathbf{B}(u_s, u_o) \approx (u_s - u_o)\boldsymbol{\delta} + u_s u_o \int_0^{u_+} dw \mathbf{H}(w). \quad (122)$$

This approximation is consistent with (13) [and therefore (115) as well]. In terms of the matrices appearing in (113), $\boldsymbol{\alpha} \approx 0$, $-\dot{\boldsymbol{\alpha}} \approx \boldsymbol{\beta} \approx \boldsymbol{\delta}$, and

$$\dot{\boldsymbol{\beta}} \approx \int_0^{u_+} dw \mathbf{H}(w). \quad (123)$$

Note that Eqs. (121) and (122) should be applied with care if \mathbf{H} involves many oscillations of an approximately periodic function. In these cases, the integral of \mathbf{H} can be very nearly zero. Terms nonlinear in \mathbf{H} might then be significant.

Assuming that (121) and (122) are indeed adequate approximations for \mathbf{A} and \mathbf{B} , a coordinate rotation may always be used to diagonalize $\dot{\beta}$. There then exist two constants \mathcal{H}_1 and \mathcal{H}_2 such that

$$\dot{\beta} = \begin{pmatrix} -\mathcal{H}_1 & 0 \\ 0 & -\mathcal{H}_2 \end{pmatrix}. \quad (124)$$

In this sense, all sufficiently short gravitational plane waves act as though they are linearly polarized [so $\beta = \chi = 0$ in (92)]. In terms of the individual wavefunctions appearing in (6),

$$\mathcal{H}_1 = \int_0^{u_+} dw [h_{\parallel}(w) + h_+(w)], \quad (125)$$

$$\mathcal{H}_2 = \int_0^{u_+} dw [h_{\parallel}(w) - h_+(w)]. \quad (126)$$

The transverse coordinates x^i have also been chosen such that

$$\int_0^{u_+} dw h_{\times}(w) = 0. \quad (127)$$

If a wave satisfies the vacuum Einstein equation, $h_{\parallel} = 0$ and $\mathcal{H}_1 = -\mathcal{H}_2$. More generally, it follows from the weak energy condition that

$$\mathcal{H}_1 + \mathcal{H}_2 \geq 0. \quad (128)$$

Now assume that the integral of h_+ is non-negative, which entails only a minimal loss of generality. Then,

$$\mathcal{H}_1 > 0, \quad \mathcal{H}_1 \geq |\mathcal{H}_2|. \quad (129)$$

We say that a wave is gravity-dominated if $\mathcal{H}_2 < 0$ and matter-dominated if $\mathcal{H}_2 > 0$. These definitions are closely analogous to those used to classify symmetric plane waves in Sect. 5. There, a wave was said to be gravity- or matter-dominated depending on the sign of the constant $h_2 = h_{\parallel} - h_+$ appearing in (96).

For weak gravity-dominated wavepackets, there can be at most one phase conjugate to an observer satisfying $u_o > u_+$. If this exists, it is evident from (122) that

$$\tau_{-1}(u_o) = -\frac{u_o}{\mathcal{H}_1 u_o - 1} \quad (130)$$

is conjugate to u_o with multiplicity 1. This equation is valid only if $\tau_{-1}(u_o) < 0$. A conjugate phase therefore exists only for observers satisfying

$$\mathcal{H}_1 u_o > 1. \quad (131)$$

Note that $\mathcal{H}_1 \bar{u}_o = 1$ is the unique physically-relevant solution to (118) in the gravity-dominated case.

Waves that are matter-dominated (so $\mathcal{H}_2 > 0$) also admit the conjugate phase (130) when u_o satisfies (131). In the conformally-flat case where $\mathcal{H}_1 = \mathcal{H}_2$, this is the only conjugate phase. Unlike in the gravity-dominated case, the multiplicity of τ_{-1} is equal to 2 for conformally-flat waves. In all other matter-dominated cases, τ_{-1} has multiplicity 1. A second conjugate phase

$$\tau_{-2}(u_o) = -\frac{u_o}{\mathcal{H}_2 u_o - 1} \quad (132)$$

is also admitted in these cases for observers satisfying

$$\mathcal{H}_2 u_o > 1. \quad (133)$$

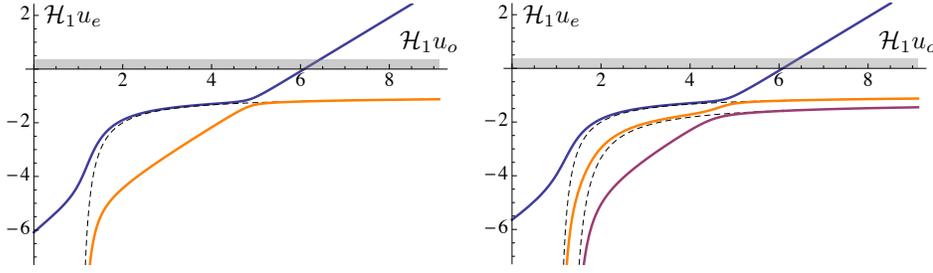


Figure 4: Example emission versus observation times in the presence of weak wavepackets. The left figure assumes a vacuum wave, while the right figure is a matter-dominated wave with $\mathcal{H}_2 = 4\mathcal{H}_1/5$. Dashed lines represent $\tau_{-1}(u_o)$ and $\tau_{-2}(u_o)$. The wave's location is indicated schematically by a thin grey rectangle. The source and observer are placed on geodesics assumed to be comoving after the wave has passed. In both cases, $\kappa_s = 1$, $\mathcal{H}_1 \delta x_o^1 = \mathcal{H}_1 \delta x_o^2 = 1/2$, and $\mathcal{H}_1(\dot{\gamma}_o \cdot \delta \mathbf{x}_o - \delta v_o) = 3$.

Note that this is a more stringent condition than (131). As implied by the notation, $\tau_{-2}(u_o) < \tau_{-1}(u_o)$. The second conjugate phase always has multiplicity 1.

Consider a point source moving on a timelike worldline Γ in a weak sandwich wave spacetime. Such a source appears differently when observed at different times. It is clear that when $u_o < 0$, exactly one image of Γ is viewable under generic conditions. As time passes, the wave eventually passes through the observer. A second image then appears when $u_o = \mathcal{H}_1^{-1} > u_+$. This image is always emitted before the first. If the wave involves a sufficient amount of Ricci curvature \P (from e.g., electromagnetic plane waves) and is not conformally-flat, a third image appears when $u_o = \mathcal{H}_2^{-1} > \mathcal{H}_1^{-1}$. This is emitted before the first two images. All images persist indefinitely once they appear. Sufficiently far in the future, one image is observed of the source as it appeared after

\P Distinguishing between different cases based on the Ricci tensor inside a wave requires that the approximations leading to (121) and (122) be valid. It is possible for, e.g., sufficiently strong vacuum waves to admit two conjugate phases in the region $u < 0$.

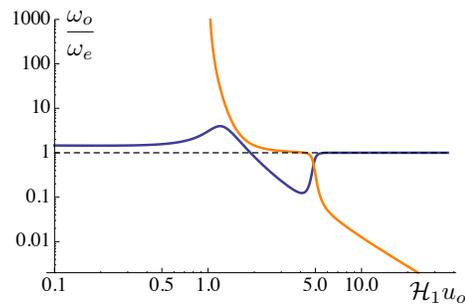


Figure 5: Frequency shifts for a vacuum wavepacket. The parameters and color coding used here are the same as those in the left panel of Fig. 4. The younger image (blue) experiences a temporary blueshift and then a redshift before settling down to $\omega_o = \omega_e$. The slight initial blueshift as $u_o \rightarrow 0^+$ is due to the source and observer moving towards each other before interacting with the wave. The older image (orange) initially appears with infinite blueshift. It then suffers an ever-increasing redshift.

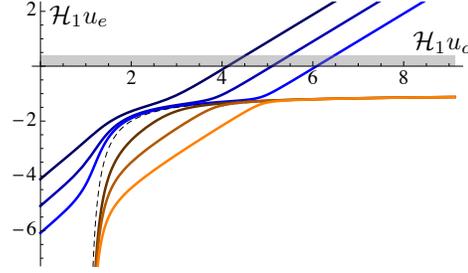


Figure 6: Bending of emission curves for a vacuum wavepacket. All parameters are the same as in the left panel of Fig. 4 except that three different choices are made for the value of δv_o . Curves for both images are shifted to the right as δv_o is decreased. The younger (blue) curves are almost unaffected by the wave if δv_o is sufficiently large. For smaller values of δv_o , both curves are strongly bent by the presence of τ_{-1} .

interacting with the gravitational wave. All other images predate this interaction. See Fig. 4.

When the second image first appears at $u_o = \mathcal{H}_1^{-1}$, the new conjugate phase $\tau_{-1}(u_o)$ is divergent. At all later times, it is finite. The same is also true for the emission times associated with the second image. Almost the entire past history of the source is therefore observable within a finite proper time. This implies an infinite blueshift. At late times, $\tau_{-1} \rightarrow -\mathcal{H}_1^{-1}$. The observed evolution of the source via the second image effectively freezes as u_e asymptotes to $-\mathcal{H}_1^{-1}$ (which predates the source’s interaction with the wave). Images such as these are highly redshifted, as indicated in Fig. 5. Note that a similar transition from infinite blueshift to infinite redshift also applies to the third image if it exists.

Another qualitative feature of the emission times plotted in Fig. 4 is that there is a sense in which pairs of images can “switch roles.” Consider, e.g., the left panel of that figure. At late times, the blue curve (corresponding to the younger image) is perfectly linear. Indeed, it remains very nearly linear until $\mathcal{H}_1 u_o \approx 5$. A rapid transition then occurs where the orange curve effectively takes over this linear behavior while the blue curve strongly deviates from it. In a sense, the two images reverse their roles. This phenomenon occurs one more time (somewhat less sharply) around $\mathcal{H}_1 u_o \approx 1$ when the second image first appears. It arises essentially because $u_e > \tau_{-1}$ for the

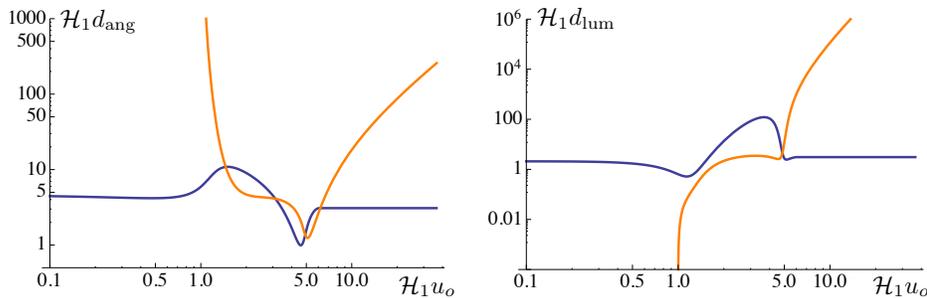


Figure 7: Angular diameter and luminosity distances for a vacuum wavepacket. All parameters are the same as those used in the left panel of Fig. 4.

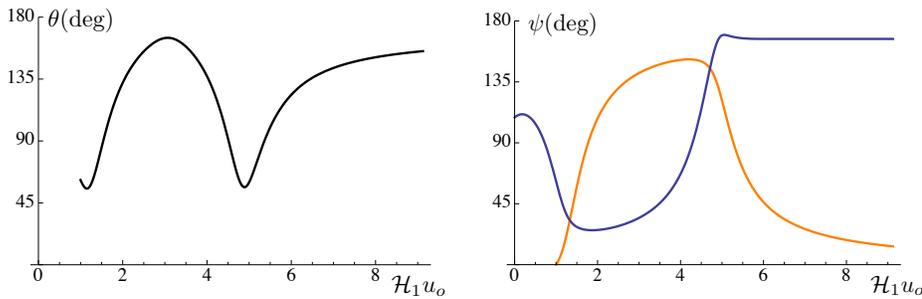


Figure 8: Angles (in degrees) for images produced by a vacuum wavepacket with the same parameters as those used in the left panel of Fig. 4. θ measures the angle between both images, while ψ measures the angle between each image and ℓ^a on an observer’s sky.

younger image and $u_e < \tau_{-1}$ for the older image. These constraints can cause emission curves to bend sharply – with large accompanying frequency shifts – in order to avoid intersecting τ_{-1} . Whether or not this occurs depends on whether the “average” linear increase of u_e ever comes near τ_{-1} . If it does, this role switching occurs. If not, the younger image is barely affected by the gravitational wave at all. This is illustrated in Fig. 6, where emission curves for several sources are plotted simultaneously.

To summarize, images which appear at discrete times briefly appear as bright blue “flashes.” Indeed, Fig. 7 shows that their luminosity distances go to zero. Simultaneously, the angular diameter distance of each new image tends to infinity. It is implied by (72) that all highly blueshifted images must make a very small angle ψ with ℓ^a on the observer’s sky. The second (and possibly third) images therefore appear aligned with the direction of propagation of the gravitational wave when they first appear. This direction could be quite different from the location of the other image(s). As time progresses, all images migrate across the observer’s sky as illustrated in Fig. 8. Different images may remain separated from each other by large angles at all times.

7. Discussion

Despite their simplicity, plane gravitational waves behave in qualitatively different ways from lenses associated with quasi-Newtonian mass distributions. As expected from their dynamic nature, plane waves generically shift the observed frequencies of various images. They may also admit images which appear to move, deform, change brightness, and shift color as time progresses.

More subtle differences relate to the number of images that are produced of a given source. For example, even numbers of images can appear generically (which has led plane wave spacetimes to be cited [1] as well-behaved examples where the odd number theorem [5, 7] does not apply). Some plane waves can even produce an infinite number of discrete images. Perhaps most striking of all are the bright flashes shown to be produced by sandwich waves in Sect. 6. These correspond to individual images which appear at discrete times. More typical gravitational lenses can produce new images if a source crosses an observer’s caustic. Individual images then split into two (or vice versa). The flashes produced by sandwich waves are quite different.

Their appearance does not require that a source pass through an observer's caustic. Such images appear individually from the infinitely distant past. Initially, they are infinitely bright and infinitely blueshifted points of light appearing in the direction of propagation associated with the gravitational wave.

Many of these effects depend at least partially on the idealization that a plane wave extends undiminished to infinitely-large transverse distances. Plane wave spacetimes are not asymptotically flat. Despite being topologically trivial and locally well-behaved, they are not even globally hyperbolic: Null geodesics passing between appropriately-chosen pairs of points can extend to arbitrarily large transverse distances in between those points. It is this property which permits the infinite number of images described in Sect. 5 to be produced by symmetric waves. The flashes described in Sect. 6 also depend on the spacetime structure at arbitrarily large distances. This structure likely affects the formation of even numbers of images as well. Indeed, the usual proofs of the odd number theorem require global hyperbolicity, among other assumptions [1, 5] (see, however, [7] for a more general formulation).

If a spacetime has the geometry of a plane wave only out to some finite transverse distance, all results derived here remain valid if the associated images involve light rays which never extend sufficiently far to interact with any large-distance modifications. The infinite sequence of images formed by a symmetric plane wave would then be expected to become finite for spacetimes which are only approximately plane waves. Calculations involving the oldest images could no longer be trusted. Similarly, the bright flashes associated with ideal sandwich waves are likely to be somewhat less extreme for waves which decay at infinity. Large brightnesses and large blueshifts can still exist, but these will be cut off at some finite maximum.

While modifications of the geometry at large distances can clearly remove some images, others may be added by such changes. Consider, for example, the pp -wave spacetimes obtained by substituting

$$H_{ij}(u)x^i x^j \rightarrow H(u, \mathbf{x}) \quad (134)$$

in the metric (1). These generalize the plane wave spacetimes. It has been shown that if H grows subquadratically as $|\mathbf{x}| \rightarrow 0$, the resulting geometries are globally hyperbolic [43]. Moreover, every pair of points is connected by at least one geodesic in these cases (unlike in pure plane wave spacetimes where the growth of H is precisely quadratic). This implies that modifications of the geometry at large transverse distances may introduce new null geodesics even between points at small transverse distances. It would be interesting to explore these effects in more depth to understand precisely how modifications of this sort (or more general ones) alter the lensing properties described here for ideal plane waves.

References

- [1] Perlick V 2004 *Living Rev. Relativity* **7** 9
- [2] Petters A O, Levine H and Wambsganss J 2001 *Singularity theory and gravitational lensing* (Boston, MA: Birkhäuser)
- [3] Schneider P, Ehlers J and Falco E E 1992 *Gravitational Lenses* (New York, NY: Springer-Verlag)
- [4] Petters A O and Werner M C 2010 *Gen. Rel. Grav.* **42** 2011
- [5] McKenzie R H 1985 *J. Math. Phys.* **26** 1592
- [6] Perlick V 1996 *Class. Quantum Grav.* **13** 529
- [7] Perlick V 2001 *Comm. Math. Phys.* **220** 403
- [8] Ehlers J and Newman E T 2000 *J. Math. Phys.* **41** 3344
- [9] Friedrich H and Stewart J M 1983 *Proc. R. Soc. London A* **385** 345

- [10] Frittelli S and Newman E T 1999 *Phys. Rev. D* **59** 124001
- [11] Seitz S, Schneider P and Ehlers J 1994 *Class. Quantum Grav.* **11** 2345
- [12] Werner M C 2009 *J. Math. Phys.* **50** 082504
- [13] Blandford R and Narayan R 1986 *Astrophys. J.* **310** 568
- [14] Damour T and Esposito-Farèse G 1998 *Phys. Rev. D* **58** 044003
- [15] Kovner I 1990 *Astrophys. J.* **351** 114
- [16] Frieman J A, Harari D D and Surpi G C 1994 *Phys. Rev. D* **50** 4895
- [17] Faraoni V 2008 *New Astron.* **13** 178
- [18] Faraoni V 1998 *Int. J. Mod. Phys. D* **7** 409
- [19] Faraoni V 1991 *Gen. Rel. Grav.* **23** 583
- [20] Ehrlich P E and Emch G G 1991 The conjugacy index and simple astigmatic focusing *Geometry and nonlinear partial differential equations* eds V Oliker and A Treibergs (Providence, RI: American Mathematical Society)
- [21] Ehrlich P E and Emch G G 1993 Geodesic and causal behavior of gravitational plane waves: astigmatic conjugacy *Differential Geometry: Geometry in mathematical physics and related topics* eds R E Greene R E and S-T Yau (Providence, RI: American Mathematical Society)
- [22] Beem J, Ehrlich P and Easley K 1996 *Global Lorentzian geometry* (New York, NY: Dekker)
- [23] Harte A I 2012 *Phys. Rev. D* **85** 124039
- [24] Hollowood T J and Shore G M 2008 *J. High Energy Phys.* **12** 091
- [25] Hollowood T J, Shore G M and Stanley R J 2009 *J. High Energy Phys.* **08** 089
- [26] Bondi H and Pirani F A E 1989 *Proc. R. Soc. London A* **421** 395
- [27] Penrose R 1976 Any spacetime has a plane wave as a limit *Differential geometry and relativity* eds. M Cahen and M Flato (Dordrecht: Reidel)
- [28] Blau M, Frank D and Weiss S 2006 *Class. Quantum Grav.* **23** 3993
- [29] Berenstein D, Maldacena J M and Nastase H 2002 *J. High Energy Phys.* **04** 013
- [30] Grignani G, Harmark T, Marini A and Orselli M 2010 *J. High Energy Phys.* **06** 034
- [31] Stanley R and Hollowood T J 2011 *J. High Energy Phys.* **10** 096
- [32] Wald R M 1984 *General relativity* (Chicago, IL: University of Chicago Press)
- [33] Pravda V, Pravdová A, Coley A and Milson R 2002 *Class. Quantum Grav.* **19** 6213
- [34] Torre C G 2006 *Gen. Rel. Grav.* **38** 653
- [35] Günther P 1988 *Huygens' principle and hyperbolic equations* (London: Academic Press)
- [36] Etherington I M H 1933 *Phil. Mag.* **15** 761
- [37] Ellis G F R 1971 Relativistic cosmology *General relativity and cosmology: Proceedings of the 47th international school of physics Enrico Fermi* ed R K Sachs (New York, NY: Academic Press)
- [38] Hartman P 1982 *Ordinary Differential Equations* (Boston, MA: Birkhäuser)
- [39] Hubeny V, Rangamani M and Ross S 2004 *Phys. Rev. D* **69** 024007
- [40] Penrose R 1965 *Rev. Mod. Phys.* **37** 215
- [41] Frittelli S, Kling T P and Newman E T 2000 *Phys. Rev. D* **63** 023006
- [42] Surpi G C and Harari D D 1999 *Astrophys. J.* **515** 455
- [43] Flores J L and Sánchez M 2006 *Lect. Notes Phys.* **692** 79