

Geometric optics in general relativity using bilocal operators

Michele Grasso,^{*} Mikołaj Korzyński,[†] and Julius Serbenta[‡]
*Center for Theoretical Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668 Warsaw, Poland.*

We consider the standard problem of observational astronomy, i.e. the observations of light emission from a distant region of spacetime in general relativity. The goal is to describe the changes between the measurements of the light performed by a sample of observers slightly displaced with respect to each other and moving with different 4-velocities and 4-accelerations. In our approach, all results of observations can be expressed as functions of the kinematic variables, describing the motions of the observers and the emitting bodies with respect to their local inertial frames, and four linear bilocal geodesic operators, describing the influence of the spacetime geometry on light propagation. The operators are functionals of the curvature tensor along the line of sight. The results are based on the assumption that the regions of emissions and observations are sufficiently small so that the spacetime curvature effects are negligible within each of them, although they are significant for the light propagation between them. The new formulation provides a uniform approach to optical phenomena in curved spacetimes and, as an application, we discuss the problem of a fully relativistic definition of the parallax and position drifts (or proper motions). We then use the results to construct combinations of observables which are completely insensitive to the motion of both the observer and the emitter. These combinations by construction probe the spacetime geometry between the observation and emission regions and in our formalism we may express them as functionals of the Riemann tensor along the line of sight. For short distances one of these combinations depends only on the matter content along the line of sight. This opens up the possibility to measure the matter content of a spacetime in a tomography-like manner irrespective of the motions of the emitter and the observer.

PACS numbers: 04.20.-q, 04.25.D; 98.80.-k, 95.10.Jk, 97.10.Vm, 97.10.Wn

I. INTRODUCTION

In the previous paper [1] the expressions for the position, redshift, Jacobi matrix, luminosity and angular distance drift, valid in any spacetime, have been derived assuming only that light propagation can be described using the geometric optics approximation. In general relativity, this simply means that light or gravitational radiation follows the null geodesics and does not influence the spacetime geometry [2, 3]. This paper extends the results of the previous one by including the GR corrections to the parallax and reformulates all derivations in a simple and convenient, geometric way. We approach here in a general way the typical situation in observational astronomy: in a region of spacetime, whose size is small with respect to the spacetime curvature scale, we have one or more bodies emitting electromagnetic radiation, called emitters or sources. The region will act like a “stage”, where various physical processes lead to the emission of radiation from certain points in the spacetime. Depending on the details of the situation the sources can be considered infinitesimally small points or extended luminous objects. They may move through the spacetime with arbitrary, not necessary geodesic motion and their emission may be continuous or pulsed. Far away, in another small region (the “auditorium”) this radiation is received by a number of observers, again moving in an arbitrary way. Both regions can be considered effectively flat due to their size, but the spacetime between them cannot. The observers register the moment of receiving a signal, measured by their proper time, and the direction from which they have seen it coming (the position on the sky). They are also able to compare these positions between each other and measure the rate of change of that position across the sky in their proper time (the position drifts, or the proper motions in astronomical terminology).

All observables in question obviously depend on the motions of both the sources and the observers, giving rise to the well-known effects of the Bradley (or stellar) aberration, parallax, relativistic time dilation etc. In a flat spacetime, they are easy to understand within the framework of special relativity. However, in a general, curved spacetime the results of observations will certainly depend also on the geometry of the spacetime in a non-trivial way. Recall that

^{*}Electronic address: grasso@cft.edu.pl

[†]Electronic address: korzynski@cft.edu.pl

[‡]Electronic address: julius@cft.edu.pl

gravity affects the light rays by the gravitational ray bending, which results in gravitational lensing, i.e. the distortion and (de)magnification of the images in the background, as well as the delays in the electromagnetic signal arrival times. We may, therefore, expect it to affect the position drift effects (by the position- and time-dependent light bending and delays) as well as the parallax. Thus all results of observations registered by the observers will inevitably contain contributions from the momentary motions of both the observer and the emitter as well as various effects of light propagation in a curved spacetime. The goal of this paper is to understand the dependence of these results on the spacetime geometry and the motions of the observers and sources in the most general situation.

The key observation from the theoretical point of view is that if the two regions in question are sufficiently small, then the influence of the spacetime geometry on light propagation can be understood using the first order geodesic deviation equation around a null geodesic. Geometrically the problem becomes the question of the behaviour of null geodesics contained in a narrow, 4-dimensional tube built around an arbitrary chosen, fiducial null geodesic connecting both regions. Irrespective of the global spacetime geometry, the geometry within the narrow connecting tube is close to the flat Minkowski geometry plus small corrections due to the curvature [4]. These corrections affect the null geodesics and produce the well-known optical effects of linear shape distortions, magnification etc. [5, 6]. As we will show in this paper, this way of conceptualizing the geometric optics in GR has many advantages: mathematically all resulting effects of light propagation can be encoded in four linear mappings, or bitensors, called the bilocal geodesic operators, that map the tangent space on one side of the connecting geodesic to the tangent space on the other side. Thanks to the geodesic deviation equation these bilocal mappings can be expressed as functionals of the Riemann curvature tensor along the null geodesic. They can be defined covariantly, without invoking any particular coordinate system or frame. They relate the behaviour of null geodesics near the observer’s and emitter’s end of the narrow tube. Once these mappings are known, including the effects of motions of the observer and the sources is then fairly straightforward, since both the “stage” and the “auditorium” are effectively flat regions: we simply need to apply the machinery of special relativity formulated in a geometric manner to calculate the observables. If we additionally assume that we may use the distant observer approximation, in which the perspective distortions are absent and the light cone structure of the spacetime is simplified, we find out that only a part of the four bilocal geodesic operator matters. The relevant information turns out to be included in two optical operators, constructed in a simple way from the bilocal geodesic operators. The first of them is the well-known Jacobi operator, relating a small deviation of the direction between two close null geodesics passing through the observation point to the transverse spatial distance between the same geodesics further away. The second one, the emitter-observer asymmetry operator, is related to the parallax effects and the position drift. It has a more complicated geometric interpretation which we shall elucidate in this paper. The natural domains of both operators are appropriate quotient spaces whose elements represent the null geodesics in a reparametrization-invariant way.

The advantage of this approach is that the whole problem can be formulated in a completely covariant and frame-independent way. Therefore in the resulting expressions for observables, we can clearly separate the dependence on the spacetime geometry (via the curvature along the line of sight) and the dependence on the momentary motions (i.e. the exact positions, the 4-velocities and the 4-accelerations of the emitter and the observer with respect to their local inertial frames). In the derivations we do not refer to any external structures like a 3+1 splitting of the spacetime, preferred frames such as the statistical isotropy frames, a background (conformal) metric, or (conformal) Killing vectors. In general relativity motions are not absolute, so it is possible to derive expressions for observables without invoking any fixed, external reference frames provided this way. Moreover, since we consider geodesics displaced in all 4 dimensions, including time, the formalism includes also the time dependence of the observables, i.e. the drift effects.

a. Applications. The most straightforward application of the mathematical machinery developed in this paper is the study of the parallax and the position drift effects in general relativity and cosmology. Formulas derived here express all possible observables in terms of the optical operators. These operators, in turn, are given by solutions of a matrix ODE with the curvature tensor along the line of sight playing the role of the input data. Given a single null geodesic in a spacetime, obtained exactly, perturbatively or by numerical integration, we may then integrate the appropriate matrix ODE’s. The solution defines then two bilocal optical operators, which encode all the relativistic light bending effects near the null geodesic. The parallax and the position drift (proper motion) can be then calculated easily given the details of the motions of the source and the observer, i.e. their momentary 4-velocities and the observer’s 4-acceleration.

This method of calculating the drift and the parallax is particularly useful in numerical relativity since the problem of position drift in a numerically evolved spacetime becomes a question of solving a number of linear ODE’s using geometric data collected along a null geodesic. The problem of the position drift and parallax in the context of cosmological distances has recently been considered by many authors in the context of so-called “real-time cosmology” [7–9]: it has been shown that the position and redshift drifts provide an additional set of observable data we may use to probe large-scale matter flows, inhomogeneities in the matter distributions and constraint this way cosmological models [7, 8, 10–16]. Moreover, numerical relativistic simulations are currently becoming one of the most important

tools of theoretical cosmology [17–27]. While the null geodesic tracking is a fairly standard problem in numerical relativity [28, 29], extracting the position drift effects is not. It is, of course, possible to do it using ray tracing together with the shooting method: we begin with one null geodesic connecting the source and the observer and then search by trial and error for another one, connecting observer’s and emitter’s worldline at a slightly later moment. This procedure is, however, rather cumbersome, while the results of this paper and [1] offer a relatively simple method to do it.

We also note that the formulation of geometric optics presented here offers a uniform theoretical approach to various optical phenomena connected with light propagation in a curved spacetime. In particular, since the gravitational lensing, the parallax effects and the position drift effects are all within reach of the framework described above, it offers the possibility to study general and non-perturbative relations between them, valid independently of the details of the spacetime geometry. As an example, in Section IV we use the optical operator-based approach to compare and contrast various definitions of parallax in the context of general relativity appearing in the literature and point out their relation to the position drift. We also apply it to study the behaviour of the gravitational lensing, the parallax and the drift effects in the vicinity of a caustic, noting their simultaneous blowup in the general case.

Finally, a less obvious application is connected with the long-standing problem of determining the spacetime geometry from the optical observations [30, 31]. Assume that a known physical process takes place on the “stage”, such as a type Ia supernova explosion, emission of gravitational waves by a binary system, or even less energetic ones like a simple motion of one or more luminous bodies. An observer in the “auditorium” region identifies it and performs a number of observations using various auxiliary observers contained within the “auditorium” region. It is assumed that the details of the positions and motions of those auxiliary observers with respect to a local inertial frame are known. The observations do not have to be momentary, in general, they will also involve the time variations of the standard observables. The observer then compares the results of all these observations with the (inferred) course of events in the “stage” region, as described in the emitter’s own frame. The results will obviously depend on the relative motions of the observers and the emitter and their motion with respect to the gravitational field. In standard astronomical or cosmological measurements, such as determining the redshift or luminosity distance, this dependence must be taken into account when interpreting the observational data. Nevertheless, as we will see, it is possible to craft certain observational strategies and define specific combinations of observables in which the dependence on momentary motions on both sides cancels out completely. These combinations depend only on the geometry of the spacetime between the two regions. Since in our formalism the dependence of the observables on the momentary motions is explicit, finding these combinations is rather straightforward. It is also easy to show that they can be expressed as functions of the two optical operators, which in turn constitute fairly simple functionals of the Riemann tensor. Therefore these combinations effectively probe the value of the curvature along the line of sight, allowing this way to discriminate between various models of the spacetime geometry or providing direct information about the tidal forces or the mass distribution within the connecting tube. Measurements of this kind may be called *direct optical measurements of curvature*. An example of such measurement, based on the notions of parallax distance and angular diameter distance, is described in Section V.

b. Limits of applicability. We assume that signal propagation can be treated within the geometric optics approximation. This means that we require the radiation wavelength to be much smaller than the size of the regions considered, and therefore much smaller than the curvature radius scale of the whole spacetime, and that the intensity of the radiation is small enough that its contribution to the stress-energy tensor does not disturb the underlying spacetime geometry [2, 3]. Wave effects can be then added as small, frequency-dependent corrections if necessary [32]. On top of that, we assume that the width of the connecting tube is small with respect to the curvature radius, which means that the first order geodesic equation approximation is valid for those geodesics connecting the two regions which are contained within the tube. This assumption holds if the curvature is small with respect to the width of the tube and roughly constant across each of its cross sections, although it may vary strongly and rapidly along the tube. The tube itself may be arbitrary long and it may also pass through strongly curved regions of spacetime.

We will focus in this paper on the case of regions positioned sufficiently far away that we may apply the distant observer approximation. Within this approximation, we assume that the relation between null tangent vector and the observed position of an object on the observer’s sky can be linearized around the fiducial geodesic and that the condition for geodesics to be null can be linearized as well. In a flat spacetime, these assumptions work very well when the emitter is positioned sufficiently far away from the observers as measured in the observer’s frame. In a non-flat spacetime, with strong lensing between the two regions considered, the applicability of this approximation is a more complicated issue and we discuss their limits of validity in more detail in Section III A. In most astrophysical applications of GR, we expect the distant observer approximation to work fairly well. Note also that the position drift formula (75) as a formula for the momentary proper motion holds independently of the distant observer approximation, as explained in Section IV D.

Note that we only linearize the equation for neighbouring geodesics and the expressions for observables in the three transverse directions, but not along the connecting tube. This means that all the curvature effects (lensing,

signal delays etc.) along the connecting tube are considered exactly, without any approximations. In particular, the formalism captures the inherently nonlinear way the curvature corrections accumulate as light passes through various regions of the spacetime.

Throughout this work, we assume that the signals are of electromagnetic nature. However, the results should also apply to the gravitational radiation as long as other assumptions listed above hold as well. Most results should also hold in any other, modified theory of gravity as long as the signal propagation follows the null geodesics of a Lorentzian metric, with the only exception of the results of Section V, in which we make use of the Einstein equations.

c. Structure of the paper. The paper is structured as follows. In the next section, we formulate the physical problem and introduce the necessary mathematical machinery, including the displacement vectors, direction deviation vectors and the first order geodesic deviation equation which connects those vectors in both regions. We define the bilocal geodesic operators obtained from the general solution of the geodesic deviation equation, explain their relation to the curvature tensor and prove a number of algebraic relations satisfied by them. Finally, we discuss the relation of the geometric objects introduced above to the results of measurements by arbitrary observers, defining also the notion of a semi-null frame connected to an observer. The main technical results of Section II, used later throughout the paper, consist of the linear relations (10)-(11) linking the displacement and direction deviation vectors in the “stage” and the “auditorium” regions, the direct relations between the bilocal geodesic operator and the curvature (12)-(21) and the equation for the apparent position on the observer’s sky (39).

In Section III we introduce first the distant observer approximation and note that under its assumptions we may pass to the appropriate quotient spaces when discussing the displacement and direction deviation vectors, getting rid this way of the gauge degrees of freedom. We then prove the two most important technical results of the paper, the time lapse formula (49) and the direction deviation equation (53), linking the behaviour of a perturbed null geodesic in the “stage” and the “auditorium” regions. We introduce for that purpose the two optical operators \mathcal{D} and m and derive their general relation to the Riemann curvature tensor.

Section IV contains most of the physical results of the paper, including the results about the relativistic corrections to the parallax. We begin by a clarification of various definitions of parallax in general relativity. We then discuss the relation of the Jacobi operator \mathcal{D} to the magnification matrix given by equation (65) (a standard result) and the relation of the emitter-observer asymmetry operator m to the parallax given by equation (71). We then rederive the equation (75) for the position drift using the optical operators and finally consider the most physically relevant case of an observer and an emitting body contained both in gravitationally bound systems undergoing a geodesic motion as a whole. We show that the time-dependent position drift with respect to the center of mass frame consists of the proper motion term given by the position drift formula plus the expected parallax terms, as in equation (77). Finally, we discuss the behaviour of the parallax, the image magnification and the position drift in the vicinity of a caustic.

In Section V we show how one can combine the data about the parallax and about the image magnifications and distortions into observables which do not depend on the momentary motions on both the emitter’s and the observer’s end of the fiducial geodesic. The combinations are defined by equations (78) and (80). We show how these observables can be related to the curvature and the matter content along the line of sight and propose also a simple measurement of the amount of gravitating matter along the line of sight. It is based on equation (86), relating one of those observables to a weighted integral of the stress-energy tensor along the line of sight for short distances.

We conclude with a Summary.

d. Notation. In this paper we will use both coordinate bases and bases related to special frames. Within these frames, we will need to separate out sets of 3 and 2 components. We, therefore, need to distinguish 4 types of indices in the paper. Greek letters μ, ν, \dots run from 0 to 3 and will denote components expressed in coordinate bases in the spacetime \mathcal{M} . Boldface greek indices $\boldsymbol{\mu}, \boldsymbol{\nu}, \dots$ also run from 0 to 3, but will be used for geometric objects expanded in frames. The boldface Latin indices $\boldsymbol{i}, \boldsymbol{j}, \dots$ run from 0 to 2 and denote the first three components in a frame. Finally the boldface capital latin indices $\boldsymbol{A}, \boldsymbol{B}, \dots$ denote the 1 and 2 components expressed in a frame. δ^μ_ν will denote the standard Kronecker delta with any set of indices.

Boldface letters will also be reserved for geometric objects in quotient spaces \mathcal{Q}_p and \mathcal{P}_p defined in Section III, while objects defined in tangent spaces will be denoted by standard letters.

Throughout the text, we assume the speed of light $c = 1$.

II. GEOMETRIC SETUP

a. Description of the physical problem. Consider two regions in the spacetime, the “stage” and the “auditorium”, separated by a large distance, but at the same time causally connected. Both regions by definition extend in both space and time, but we assume they are small enough that their local geometry can be treated as flat. On the other hand, the region connecting them is curved, although the curvature is small enough so that light propagation can be treated within the first order geodesic deviation equation. We assume that from every point in the “stage” region

it is possible to find a null geodesic which crosses the “auditorium” in the future. In this setup the communication between the two regions is possible only one way: light signals emitted from one of them, the “stage”, can reach later the other one, the “auditorium”, provided that they have been sent in the right direction.

We assume that each of the regions hosts one or more participants who travel through the spacetime along timelike worldlines (not necessary geodesic). The emitters in the “stage” region send time-dependent electromagnetic radiation at all directions, while the observers in the “auditorium” perform the standard astronomical observations, registering the energy and the direction from which they perceive the light coming at a given moment of their worldline, as well as their variations in their own proper time. The problem we will be concerned with in this paper can be phrased in the following way: knowing the trajectories of the emitters and the observers in their locally flat regions, for example with respect to a local inertial frame, describe the results of the observations made by the observers in the geometric optics approximation.

As we have already noted in the introduction, the results will depend on the motions of both the emitters and the observers and also on the properties of the spacetime between them. In this paper, in the spirit of [1], we would like to separate these dependencies. Our aim is to obtain expressions for the observables which contain the quantities describing the momentary motions on both sides (i.e. the 4-velocities, 4-accelerations and momentary positions with respect to the local inertial frames) as well as geometric objects describing the effects of gravitational fields on the light propagation and defined in a frame- and coordinate-independent way. The problem may seem at first sight impossible to tackle in full generality, without any assumptions about the spacetime geometry. Recall that in general relativity the light propagation is affected by gravity in many ways: variations of the gravitational field induce the gravitational light bending, causing also the Shapiro and the geometric delays in the signal arrivals. On top of that, the gravitational fields may also increase or decrease the energy of the photon. In a time-dependent gravitational field, all effects may also be time- and position-dependent, influencing this way the time variations of the observables recorded by the observers. In the most general approach, without any simplifying assumptions, the light propagation in the geometric optics approximation is governed by the null geodesic equation which contains the Christoffel symbols [2]. The Christoffel symbols, in turn, depend on the values of the metric tensor and its first derivatives. It seems then that without further assumptions about the metric little can be said about the relations between the motions of the participants in both regions and the results of observations.

Fortunately, the problem simplifies very much when we realize that all the information about the gravitational fields we need is actually the geometry inside the narrow tube connecting the two regions. This can be seen as follows: let us single out for the sake of convenience a pair of points, \mathcal{E} in the “stage” region and \mathcal{O} in the “auditorium”, connected by a null geodesic γ_0 representing a single ray of light. The two points will serve as a reference for the positions of objects in the corresponding regions, while the null geodesic γ_0 will play the role of the fiducial null geodesic for all other light rays. We then build a long, thin 4-dimensional tube around it, extending in both the 2 spatial dimensions as well as time. Its width should be comparable to the size of both regions. Now, one can show that as long as the connecting tube is narrow with respect to the typical curvature radius of the manifold, the geometry of the spacetime inside it is quite simple: it must be close to the flat space up to the leading order corrections proportional to the value of the Riemann tensor along γ_0 . This may seem slightly surprising at first, but we can see that clearly if we introduce the equivalent of the Fermi coordinates in the neighbourhood of γ_0 .

The Fermi coordinates around a timelike geodesic are well-known and described in many textbooks [33], but in the case of a null geodesic they are less known [4]. In short, given the null geodesic γ_0 and an appropriate parallel propagated frame along it one constructs coordinates (ξ^i, λ) , consisting of a coordinate λ agreeing with the affine parametrization of γ_0 and three transverse coordinates ξ^i , such that $\xi^i = 0$ corresponds to γ_0 . Unlike the timelike case, all three transverse coordinates cannot be made orthogonal to γ_0 . Nevertheless, the metric tensor near γ_0 can be expanded as a Taylor series in ξ^i just like in the “standard” Fermi coordinates [4]:

$$g_{\mu\nu} = C_{\mu\nu} + D(\mu, \nu) R_{\mu i j \nu}(\lambda) \xi^i \xi^j + O(|\xi|^3), \quad (1)$$

where $C_{\mu\nu}$ is a constant matrix representing the flat space in non-standard coordinates, $R_{\mu i j \nu}(\lambda)$ denotes the components of the Riemann tensor along γ_0 in the parallel propagated frame and $D(\mu, \nu)$ are irrelevant constant coefficients depending on the indices μ and ν . Thus the leading order term in the expansion turns out to be always constant, corresponding to a flat tube cut out of the Minkowski spacetime, plus subleading curvature corrections quadratic in the transverse coordinates. We have shown this way that while in a general coordinate system the fiducial geodesic may have a very complicated form, resulting in a complicated form of the metric in its neighbourhood, the Fermi coordinates can “unwind” the geodesic together with the nearby spacetime geometry, revealing that it has a fairly simple form of a flat tube with slight deformations seen in its outer parts.

In this setting *all* the light propagation effects (gravitational light bending, propagation delays etc.), taking place between the “stage” and the “auditorium”, can be understood in terms of the geometry of the connecting tube, given by (1), and its influence on null geodesics. Consequently, all the relativistic corrections to the light propagation effects should be *expressible as functionals of the only relevant ingredient of the spacetime geometry, i.e. the Riemann tensor*

along the line of sight. The geometry outside the connecting tube on the other hand, possibly very complicated and not expressible easily in terms of data defined along γ_0 , is irrelevant from the point of view of the problem. Let us stress again that the statements above hold for *arbitrary* geodesics and along their *whole* length, even if they are long enough to feel the global spacetime curvature or they pass through strongly curved regions. The global decomposition (1) of the narrow tube geometry works for those geodesics equally well as for short ones.

The main mathematical tool of this paper, applicable under the assumptions defined above to the problem of null geodesics connecting both regions, is the first order geodesic deviation equation (GDE). Recall that it is a linear ODE relating directly the behaviour of neighbouring geodesics to the Riemann tensor along a fiducial geodesic we use as a reference [34–37]. One may obtain it from the geodesic equation in metric (1) by neglecting the quadratic terms in the velocity deviations.

The geodesic deviation equation of the first and higher orders (and their generalizations) around a timelike geodesic, representing the relative motion of free falling bodies, has been discussed by many authors (see for example [34–36, 38–43]), but the null case has attracted less attention [1, 37, 44–46]. The general solution of the first order GDE around a null geodesics allows for studying the behaviour of the rays of light in a completely covariant manner, i.e. without invoking explicitly a coordinate system or a frame. This is consistent with the fully relativistic and geometric point of view we adopt in this paper: we delay the introduction of coordinate systems, frames or other auxiliary structures as much as possible in order to understand the “pure” geometric relations, valid independently of them.

b. Mathematical formulation of the problem. Let \mathcal{M} be the spacetime with a Lorentzian metric g , of signature $(-, +, +, +)$. We consider two regions of spacetime: the “stage” $N_{\mathcal{O}}$ and the “auditorium” $N_{\mathcal{E}}$. Both are 4-dimensional domains in the vicinity of two points $\mathcal{O} \in N_{\mathcal{O}}$ and $\mathcal{E} \in N_{\mathcal{E}}$ respectively. We assume that \mathcal{O} and \mathcal{E} are connected by a null geodesic γ_0 , called the *fiducial geodesic* (or the *optical axis* in non-relativistic optics literature), i.e. $\gamma_0(\lambda_{\mathcal{O}}) = \mathcal{O}$ and $\gamma_0(\lambda_{\mathcal{E}}) = \mathcal{E}$ for the values $\lambda_{\mathcal{O}}$ and $\lambda_{\mathcal{E}}$ of the affine parameter λ of γ_0 . For convenience we also assume that we have parametrized our geodesic backwards in time, i.e. $\lambda_{\mathcal{O}} < \lambda_{\mathcal{E}}$. As null geodesics do not have a preferred, normalized parametrization, we may always reparametrize γ_0 by an affine transformation without violating the assumptions above:

$$\lambda \rightarrow \lambda' = E \cdot \lambda + F, \quad (2)$$

with $E > 0$ and arbitrary F . Then the tangent vector to γ_0 transforms according to

$$l^\mu \rightarrow l'^\mu = \frac{1}{E} l^\mu. \quad (3)$$

In geometric optics we are interested only in the incidence relations between null geodesics and points, i.e. in the question whether or not a null geodesic passes through a given event and what null direction it follows at that moment. These relations are invariant with respect to reparametrization, because the value of the affine parameter at which the null geodesic intersects a given point carries no physical meaning. Therefore we may identify all null geodesics which share the same path and consider affine reparametrizations (2) gauge transformations from the point of view of geometric optics. They should therefore leave all physical observables invariant.

The regions $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ are assumed to be of the size of L much smaller than the characteristic curvature scale of the spacetime R_c , i.e. $L \ll R_c$. In this case, we may introduce in both regions locally flat coordinate systems, centered at \mathcal{O} and \mathcal{E} respectively, in which the metric tensor is the flat Minkowski metric up to quadratic terms in x^μ :

$$g_{\mu\nu} = \eta_{\mu\nu} + O(x^2).$$

The additional terms are due to the local spacetime curvature and therefore scale like $(L/R_c)^2$. We may consider them negligibly small because the size of both regions L is too small for any curvature effects to be directly detected by experiments performed within each region. Therefore we will effectively treat both $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ as flat. Stating the same in a more coordinate-invariant manner: with the curvature effects being negligible in regions of size L , we may simply identify the points in $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ with points in the tangent spaces $T_{\mathcal{O}}\mathcal{M}$ and $T_{\mathcal{E}}\mathcal{M}$ in the vicinity of \mathcal{O} , using, for example, the exponential map. Under this identification, the physical spacetime metric agrees with the flat metric on the appropriate tangent space up to quadratic terms in the distance from 0.

Consider now all geodesics connecting points from $N_{\mathcal{O}}$ with $N_{\mathcal{E}}$ contained in a 4-dimensional tube around γ_0 , sufficiently narrow so that we can apply the metric expansion (1) for its geometry. This means we can use the first order geodesic deviation equation for those geodesics. The geodesics are uniquely specified by giving an initial point $i_{\mathcal{O}}$ in $N_{\mathcal{O}}$ and the initial tangent vector $v_{\mathcal{O}}$ in $N_{\mathcal{O}}$. In locally flat coordinates they look like straight lines, although their propagation through the spacetime in between may more complicated, with details depending on the coordinate system, see Figure 1. They form an 8-parameter family of curves.

Now, let $l_{\mathcal{O}}^\mu$ be the tangent vector to γ_0 at \mathcal{O} and $l_{\mathcal{E}}^\mu$ at \mathcal{E} . We would like to parametrize the geodesics from the family considered by their deviation from the fiducial null geodesic γ_0 . Obviously we can parametrize every perturbed

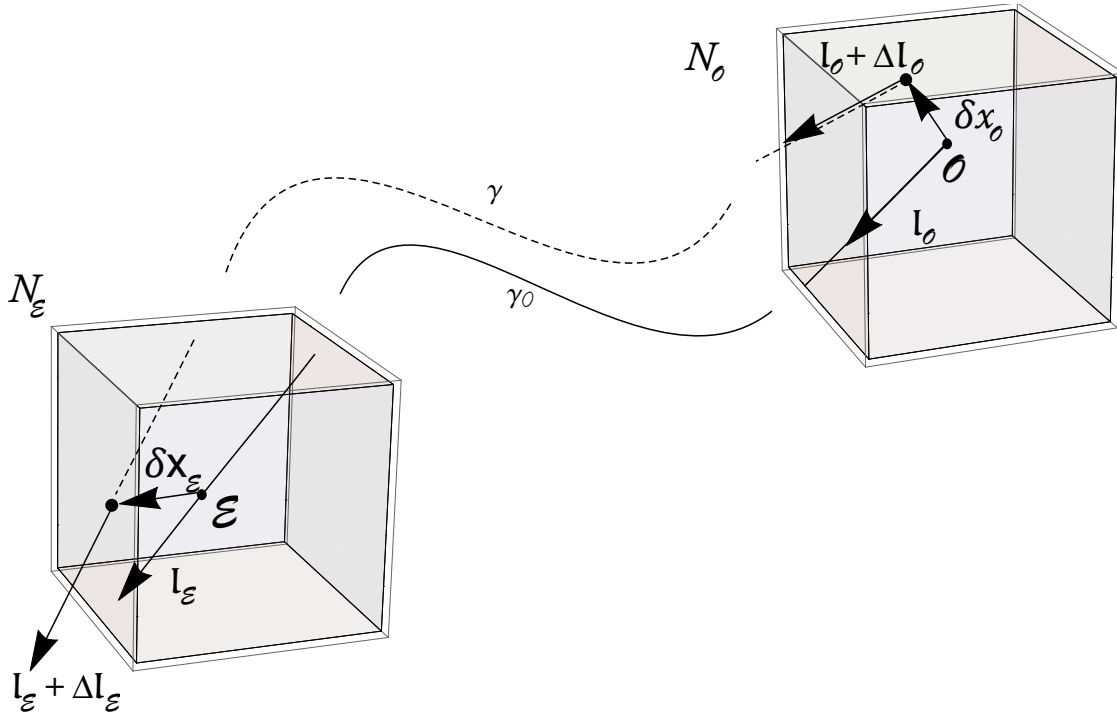


FIG. 1: Flat, distant regions $N_{\mathcal{E}}$ and $N_{\mathcal{O}}$ connected by a family of null geodesics. The fiducial null geodesic γ_0 connects events \mathcal{O} and \mathcal{E} . We consider other geodesics lying entirely within a narrow tube around γ_0 . The geodesics may be parametrized by a pair of vectors in the tangent space $T_{\mathcal{O}}\mathcal{M}$ giving the initial displacement $(\delta x_{\mathcal{O}})$ and the initial direction deviation $(\Delta l_{\mathcal{O}})$ with respect to γ_0 . Similar parametrization $(\delta x_{\mathcal{E}}, \Delta l_{\mathcal{E}})$ can also be used in $N_{\mathcal{E}}$, at point $\lambda = \lambda_{\mathcal{E}}$ along the displaced geodesic.

geodesic $\tilde{\gamma}$ by its initial position p at $\lambda = \lambda_{\mathcal{O}}$ and the tangent vector v^{μ} at p . With the point-vector the identification described above we may use simply the *initial displacement vector* $\delta x_{\mathcal{O}} \in T_{\mathcal{O}}\mathcal{M}$ as the parametrization of the initial position. In a locally flat coordinate system (y^{μ}) in $N_{\mathcal{O}}$ it is defined as

$$\delta x_{\mathcal{O}}^{\mu} = y^{\mu}(\tilde{\gamma}(\lambda_{\mathcal{O}})) - y^{\mu}(\mathcal{O}). \quad (4)$$

Note that $\delta x_{\mathcal{O}}$ is a fully 4-dimensional vector, because we consider here displacements in all 3 spatial directions and in time.

We now move on to v^{μ} . Fixing an arbitrary coordinate system in $N_{\mathcal{O}}$ we may use the difference between the components of v^{μ} and $l_{\mathcal{O}}^{\mu}$ to parametrize the tangent vector:

$$\delta l_{\mathcal{O}}^{\mu} = v^{\mu} - l_{\mathcal{O}}^{\mu} \quad (5)$$

in analogy with (4). However, $\delta l_{\mathcal{O}}^{\mu}$ defined this way is not a proper vector, because it does not transform like a vector under general coordinate system transformations. In other words, expression (5) defines, in general, different elements of the tangent space $T_{\mathcal{O}}\mathcal{M}$ when applied in different coordinate systems. As we know from elementary differential geometry, this is because we have subtracted here the components of vectors defined at two close but distinct points, p and \mathcal{O} . Of course, this can be fixed by adding an appropriate term involving the Christoffel symbols term: we define

$$\Delta l_{\mathcal{O}}^{\mu} = \delta l_{\mathcal{O}}^{\mu} + \Gamma^{\mu}_{\nu\sigma}(\mathcal{O}) l_{\mathcal{O}}^{\nu} \delta x_{\mathcal{O}}^{\sigma}, \quad (6)$$

where $\Gamma^{\mu}_{\nu\sigma}(\mathcal{O})$ are the Christoffel symbols at \mathcal{O} . The resulting expression $\Delta l_{\mathcal{O}}^{\mu}$ parametrizes the initial tangent vector equally well, but unlike the “bare” $\delta l_{\mathcal{O}}^{\mu}$ it is a proper vector. From the geometric point of view, equation (6) yields the *same* vector in $T_{\mathcal{O}}\mathcal{M}$ independently of the coordinate system chosen. It is obvious, though, that in locally flat coordinate systems like (y^{μ}) - and only in those - we have $\Delta l_{\mathcal{O}}^{\mu} = \delta l_{\mathcal{O}}^{\mu}$ and both definitions coincide. The vector $\Delta l_{\mathcal{O}}^{\mu}$,

which we will call the *initial direction deviation vector*, corresponds to the difference between v^μ parallel transported from p to \mathcal{O} and $l_\mathcal{O}^\mu$. The pair $(\delta x_\mathcal{O}^\mu, \Delta l_\mathcal{O}^\mu)$ defines a unique geodesic considered as a perturbation of γ_0 and will be referred to as *the displacement vectors*, see again Figure 1.

c. Remarks. Since the geodesics are supposed to be confined within the narrow tube all along γ_0 both the initial displacement and the initial direction deviation cannot be too large. Additionally, the condition for applicability of the first order GDE means that the gravitational lensing produces only a linear distortion of the image of all objects in $N_\mathcal{E}$, which excludes the possibility of multiple imaging for light rays contained within the tube.

The choice of the particular value $\lambda_\mathcal{O}$ of the affine parameter λ at which we parametrize the initial displacement and direction deviation is arbitrary: in principle we could choose any point and any value for that purpose. This choice is nevertheless consistent with the assumption that the geodesics are only linearly perturbed with respect to the fiducial one: we may expect that for slightly perturbed geodesics the endpoint given by $\lambda = \lambda_\mathcal{O}$ will lie very close to the corresponding endpoint of γ_0 , i.e. \mathcal{O} , and therefore within $N_\mathcal{O}$. Thus we can parametrize the position by a small displacement vector. The same reasoning applies to the other endpoint.

A. Geodesic deviation equation and bilocal geodesic operators

We will now consider the relation between the displacement vectors around \mathcal{O} and around \mathcal{E} for geodesics passing through $N_\mathcal{O}$ and $N_\mathcal{E}$. Assume that for $\lambda = \lambda_\mathcal{E}$ the deviated geodesic passes through $N_\mathcal{E}$ and let $\delta x_\mathcal{E}$ and $\Delta l_\mathcal{E}$ denote the displacement vector and the direction deviation vector respectively, see Figure 1. Assuming the geodesics deviate from γ_0 only by distances small with respect R_c , we may obtain these vectors by solving the ODE for the first order perturbation of the geodesic equation around γ_0 , i.e. the first order geodesic deviation equation (GDE):

$$\nabla_l \nabla_l \xi^\mu - R^\mu{}_{\alpha\beta\nu} l^\alpha l^\beta \xi^\nu = 0, \quad (7)$$

with the initial data

$$\xi^\mu(\lambda_\mathcal{O}) = \delta x_\mathcal{O}^\mu \quad (8)$$

$$\nabla_l \xi^\mu(\lambda_\mathcal{O}) = \Delta l_\mathcal{O}^\mu. \quad (9)$$

With this setup we obtain the displacements at the other end from the values of the solution for $\lambda = \lambda_\mathcal{E}$: $\delta x_\mathcal{E}^\mu = \xi^\mu(\lambda_\mathcal{E})$ and $\Delta l_\mathcal{E}^\mu = \nabla_l \xi^\mu(\lambda_\mathcal{E})$. The combination $R^\mu{}_{\alpha\beta\nu} l^\alpha l^\beta$ is often referred to as the *optical tidal matrix*.

Since the GDE is linear, the solution at $\lambda_\mathcal{E}$ must be a linear function of the initial data at \mathcal{O} :

$$\delta x_\mathcal{E}^\mu = W_{XX}{}^\mu{}_\nu \delta x_\mathcal{O}^\nu + W_{XL}{}^\mu{}_\nu \Delta l_\mathcal{O}^\nu \quad (10)$$

$$\Delta l_\mathcal{E}^\mu = W_{LX}{}^\mu{}_\nu \delta x_\mathcal{O}^\nu + W_{LL}{}^\mu{}_\nu \Delta l_\mathcal{O}^\nu, \quad (11)$$

with W_{XX} , W_{XL} , W_{LX} , W_{LL} being bilocal operators (also known as 2-point tensors [47] or bitensors [33, 36]), acting from $T_\mathcal{O}\mathcal{M}$ to $T_\mathcal{E}\mathcal{M}$. Together they form the resolvent operator, or the Wroński matrix [48] for the GDE, $W = W(\lambda_\mathcal{E}, \lambda_\mathcal{O})$. It is a linear mapping between vector sums of two copies of the tangent space, i.e.

$$W : T_\mathcal{O}\mathcal{M} \oplus T_\mathcal{O}\mathcal{M} \rightarrow T_\mathcal{E}\mathcal{M} \oplus T_\mathcal{E}\mathcal{M},$$

defined by the relation $W(\delta x_\mathcal{O}, \Delta l_\mathcal{O}) = (\delta x_\mathcal{E}, \Delta l_\mathcal{E})$. As noticed by Uzun [37], it is also a symplectic mapping, since in GR the ODE's for null geodesics can be formulated as a Hamiltonian system, in general as well as in the first order perturbation theory [48] [76].

In this paper the four operators W_{XX} , W_{XL} , W_{LL} and W_{LX} will be referred to as the *bilocal geodesic operators* (BGO). The notation we introduced for them highlights the fact that they constitute four parts of a larger geometric object. In the context of timelike geodesics the first two have already been introduced by DeWitt and Brehme [49] and Dixon [36, 50], under the name of *Jacobi propagators*, denoted by K and H (the definition of the latter involves often an additional prefactor). They can also be obtained by differentiating the Synge's worldfunction [36, 47, 50], but here we will not make any use of the worldfunction formalism. Recently the BGO's defined along a timelike geodesic have been used as a tool to study of the gravitational waves memory effect [43]. In the rest of the paper, we will focus exclusively on the null case and explore the relation of the BGO's to the spacetime geometry and to the optical observations made by observers in $N_\mathcal{O}$.

a. Bilocal geodesic operators and the Riemann curvature tensor. It follows easily from the geodesic deviation equation (7) and from equations (10)-(11) that the BGO's can be expressed as solutions to ODE's defined along γ_0

and involving the curvature tensor. Assume we fix a parallel-propagated frame e_μ along the null geodesic. Then we solve the following ODE for a 4-by-4 matrix-valued function $A^\mu{}_\nu(\lambda)$ with initial data at $\lambda_{\mathcal{O}}$:

$$\dot{A}^\mu{}_\nu - R^\mu{}_{\alpha\beta\sigma} l^\alpha l^\beta A^\sigma{}_\nu = 0 \quad (12)$$

$$A^\mu{}_\nu(\lambda_{\mathcal{O}}) = \delta^\mu{}_\nu \quad (13)$$

$$\dot{A}^\mu{}_\nu(\lambda_{\mathcal{O}}) = 0, \quad (14)$$

where dot denotes the derivative with respect to the affine parameter λ and $R^\mu{}_{\alpha\beta\sigma} l^\alpha l^\beta$ denotes the components of the optical tidal matrix in the parallel propagated frame. From the definition (10)-(11) we can prove that the two BGO's are given in terms of $A^\mu{}_\nu(\lambda)$ and its derivative at the emission point:

$$W_{XX}{}^\mu{}_\nu = A^\mu{}_\nu(\lambda_{\mathcal{E}}) \quad (15)$$

$$W_{LX}{}^\mu{}_\nu = \dot{A}^\mu{}_\nu(\lambda_{\mathcal{E}}). \quad (16)$$

The other two operators can be obtained from the solution of the same equation with different initial data, namely we take the matrix-valued function $B^\mu{}_\nu(\lambda)$ satisfying

$$\ddot{B}^\mu{}_\nu - R^\mu{}_{\alpha\beta\sigma} l^\alpha l^\beta B^\sigma{}_\nu = 0 \quad (17)$$

$$B^\mu{}_\nu(\lambda_{\mathcal{O}}) = 0 \quad (18)$$

$$\dot{B}^\mu{}_\nu(\lambda_{\mathcal{O}}) = \delta^\mu{}_\nu. \quad (19)$$

In that case we have

$$W_{XL}{}^\mu{}_\nu = B^\mu{}_\nu(\lambda_{\mathcal{E}}) \quad (20)$$

$$W_{LL}{}^\mu{}_\nu = \dot{B}^\mu{}_\nu(\lambda_{\mathcal{E}}). \quad (21)$$

Relations presented above clarify the dependence of the BGO's on the spacetime geometry. They show that they can be expressed as non-local functionals of the Riemann tensor along the line of sight. We stress here that even though the GDE and the matrix equations (12)-(19) are linear, the BGO's are *nonlinear* functionals of the curvature tensor along γ_0 . The reader may check that if we take two solutions of (12)-(14) or (17)-(19) corresponding to two different optical tidal tensor functions $R^\mu{}_{\alpha\beta\sigma} l^\alpha l^\beta(\lambda)$, then a linear combination of these solutions *does not* satisfy the same equations (12)-(14) or (17)-(19) for the same linear combination of the optical tidal tensor functions. The situation is analogous to the dependence of the evolution operator $U(t)$ on the Hamiltonian $H(t)$ in quantum mechanics: $U(t)$ is defined by a first order ODE and initial condition

$$i\hbar\dot{U}(t) = H(t)U(t) \quad (22)$$

$$U(0) = \mathbf{1}. \quad (23)$$

Now, adding a perturbation term to the Hamiltonian results a nonlinear change of the evolution operator, which can be expressed using the well-known path-ordered exponential formula [77].

The nonlinearity of the BGO's as functionals of curvature reflects the fact that our formalism captures all nonlinear effects of combined light bending at different points along the fiducial geodesic. Note also that although all four BGO's are effectively functionals of the *same* optical tidal tensor $R^\mu{}_{\alpha\beta\sigma} l^\alpha l^\beta(\lambda)$, they are in fact *different functionals* and therefore without any further assumptions regarding the curvature their values for any pair of points \mathcal{O} and \mathcal{E} should be treated as completely independent from each other.

b. Algebraic properties of the BGO's. In [1, 36] two general properties of the solutions of the GDE have been proved. We will show here that they immediately translate to corresponding two properties of the bilocal geodesic operators, which hold irrespective of the spacetime geometry or whether γ_0 is null or not.

Let ξ^μ denote a solution of the GDE (7). Then we have

$$\xi^\mu l_\mu = A + B \lambda,$$

where $A, B = \text{const.}$ This way we have defined 2 constants of motion for the GDE, namely $B = \nabla_l \xi^\mu l_\mu$ and $A = \xi^\mu l_\mu - B \lambda$.

The second property is that the expression

$$\xi^\mu = (C + D \lambda) l^\mu \quad (24)$$

with $C, D = \text{const}$ is always a solution and, since the first order GDE is linear, it can be added to any other solution without affecting equation (7). Both properties are easy to verify using equation (7).

Consider the solution (24) at \mathcal{O} and \mathcal{E} : we have $\delta x_{\mathcal{O}}^{\mu} = (C + \lambda_{\mathcal{O}} D) l_{\mathcal{O}}^{\mu}$, $\Delta l_{\mathcal{O}}^{\mu} = D l_{\mathcal{O}}^{\mu}$ and $\delta x_{\mathcal{E}}^{\mu} = (C + \lambda_{\mathcal{E}} D) l_{\mathcal{E}}^{\mu}$, $\Delta l_{\mathcal{E}}^{\mu} = D l_{\mathcal{E}}^{\mu}$. We substitute these equations to (10)-(11) and assuming the resulting relations must hold for all C and D we get

$$W_{XX}{}^{\mu}{}_{\nu} l_{\mathcal{O}}^{\nu} = l_{\mathcal{E}}^{\mu} \quad (25)$$

$$W_{LX}{}^{\mu}{}_{\nu} l_{\mathcal{O}}^{\nu} = 0 \quad (26)$$

$$W_{XL}{}^{\mu}{}_{\nu} l_{\mathcal{O}}^{\nu} = (\lambda_{\mathcal{E}} - \lambda_{\mathcal{O}}) l_{\mathcal{E}}^{\mu} \quad (27)$$

$$W_{LL}{}^{\mu}{}_{\nu} l_{\mathcal{O}}^{\nu} = l_{\mathcal{E}}^{\mu}. \quad (28)$$

The second set of relations can be obtained from the conservation of the two constants of motion defined above. Namely, for any initial data $\delta x_{\mathcal{O}}^{\mu}$ and $\Delta l_{\mathcal{O}}^{\mu}$ the values of A and B need to remain equal in \mathcal{O} and \mathcal{E} . This means that

$$\begin{aligned} l_{\mathcal{O}\mu} \Delta l_{\mathcal{O}}^{\mu} &= l_{\mathcal{E}\mu} \Delta l_{\mathcal{E}}^{\mu} \\ l_{\mathcal{O}\mu} \delta x_{\mathcal{O}}^{\mu} - \lambda_{\mathcal{O}} l_{\mathcal{O}\mu} \Delta l_{\mathcal{O}}^{\mu} &= l_{\mathcal{E}\mu} \delta x_{\mathcal{E}}^{\mu} - \lambda_{\mathcal{E}} l_{\mathcal{E}\mu} \Delta l_{\mathcal{E}}^{\mu}. \end{aligned}$$

We again make use of (10)-(11) in order to express $\delta x_{\mathcal{E}}^{\mu}$ and $\Delta l_{\mathcal{E}}^{\mu}$ by $\delta x_{\mathcal{O}}^{\mu}$ and $\Delta l_{\mathcal{O}}^{\mu}$. The resulting equations turn out to be equivalent to the following 4 relations:

$$l_{\mathcal{E}\mu} W_{XX}{}^{\mu}{}_{\nu} = l_{\mathcal{O}\nu} \quad (29)$$

$$l_{\mathcal{E}\mu} W_{LX}{}^{\mu}{}_{\nu} = 0 \quad (30)$$

$$l_{\mathcal{E}\mu} W_{XL}{}^{\mu}{}_{\nu} = (\lambda_{\mathcal{E}} - \lambda_{\mathcal{O}}) l_{\mathcal{O}\nu} \quad (31)$$

$$l_{\mathcal{E}\mu} W_{LL}{}^{\mu}{}_{\nu} = l_{\mathcal{O}\nu}. \quad (32)$$

Finally, let us note that two of the BGO's undergo rescalings under the affine reparametrizations of the fiducial null geodesic γ_0 . Namely, under the reparametrization (2) we have the following transformation law:

$$W_{XX} \rightarrow W'_{XX} = W_{XX} \quad (33)$$

$$W_{XL} \rightarrow W'_{XL} = E \cdot W_{XL} \quad (34)$$

$$W_{LX} \rightarrow W'_{LX} = \frac{1}{E} \cdot W_{LX} \quad (35)$$

$$W_{LL} \rightarrow W'_{LL} = W_{LL}. \quad (36)$$

B. Observed position on the sky and semi-null frames

Let $u_{\mathcal{O}}^{\mu}$ be the 4-velocity of an observer and l^{μ} the past-pointing, null tangent vector to a null geodesic passing through \mathcal{O} . The direction from which the observer sees the light coming is defined as a normalized, spatial vector, orthogonal to $u_{\mathcal{O}}^{\mu}$, pointing in the same direction as l^{μ} :

$$r^{\mu} = \frac{1}{l_{\sigma} u_{\mathcal{O}}^{\sigma}} l^{\mu} + u_{\mathcal{O}}^{\mu}. \quad (37)$$

Formula (37) defines an observer-dependent mapping

$$V(u_{\mathcal{O}}, \cdot) : \mathcal{N}_{\mathcal{O}}^{-} \rightarrow \text{Dir}(u_{\mathcal{O}})$$

from the set of past-oriented null vectors $\mathcal{N}_{\mathcal{O}}^{-} = \{X \in T_{\mathcal{O}}\mathcal{M} | X^{\mu} X_{\mu} = 0, X^0 < 0\}$ to the observer's sphere of directions, i.e. the set of normalized, purely spatial vectors for the observer, i.e. $\text{Dir}(u_{\mathcal{O}}) = \{X \in T_{\mathcal{O}}\mathcal{M} | X^{\mu} X_{\mu} = 1, u_{\mathcal{O}}^{\mu} X_{\mu} = 0\}$ [1, 51, 52]. The space $\mathcal{N}_{\mathcal{O}}^{-}$ does not have a well-defined metric and, therefore, by using it, it is not possible to calculate the angular distance between points on the sky: one really needs to pass to the observer-dependent space $\text{Dir}(u_{\mathcal{O}})$ in order to do that. It turns out that this introduces the dependence of the angle measurements between apparent positions on the sky from the observer's 4-velocity. This is commonly referred to as the *light aberration effect*, or *stellar aberration* and may be explained by a relative tilt of the spheres of directions (and the whole simultaneity planes) for observers with different 4-velocities, see [53] for a more detailed discussion involving the historical background.

In the context of relativistic geometric optics it is customary to introduce a frame, or vierbein, connected with the observer. The standard approach is to use the Sachs orthonormal frame, consisting of $u_{\mathcal{O}}^{\mu}$, the direction vector

r^μ and two perpendicular, spatial vectors e_A^μ spanning the Sachs screen space [54–56]. We have found out that it is more convenient to use a related, but slightly modified frame, which we will call the *semi-null frame* (SNF). The frame consists of $u_\mathcal{O}^\mu$, the same two perpendicular, spatial vectors e_A^μ and the null vector l^μ instead of r^μ . It is not orthonormal and we can check that the products of the basis vectors read

$$\begin{aligned} l^\mu l_\mu &= 0 \\ e_A^\mu l_\mu &= 0 \\ u_\mathcal{O}^\mu e_{A\mu} &= 0 \\ e_A^\mu e_{B\mu} &= \delta_{AB} \\ u_\mathcal{O}^\mu u_{\mathcal{O}\mu} &= -1 \\ l^\mu u_{\mathcal{O}\mu} &= Q > 0, \end{aligned}$$

where we have introduced the notation Q for the product of l^μ and $u_\mathcal{O}^\mu$, a constant, but so far undetermined.

We remind the reader that in the subsequent calculations the frame indices will be denoted by boldface letters, with capital latin indices $\mathbf{A}, \mathbf{B}, \dots$ running over the spatial components $\mathbf{1}$ and $\mathbf{2}$, small latin indices i, j, \dots running over $\mathbf{0}, \mathbf{1}$ and $\mathbf{3}$ and the boldface Greek indices $\boldsymbol{\mu}, \boldsymbol{\nu}, \dots$ running over all 4 dimensions from $\mathbf{0}$ to $\mathbf{3}$.

The position on the observer's sky r_0^μ , determined by the fiducial geodesic γ_0 , will serve as the point of reference for all other points on the sky considered here. Namely, let k^μ be another past-oriented null vector, corresponding to another source of light, and let $\tilde{r}^\mu = \frac{1}{k_\sigma u_\mathcal{O}^\sigma} k^\mu + u_\mathcal{O}^\mu$ be the corresponding direction vector. If the position of the second source lies on the same hemisphere as r_0^μ (we will assume that throughout the paper) then it can be uniquely determined from the transversal components of \tilde{r}^μ in the semi-null frame of the observer, denoted by $\tilde{r}^{\mathbf{A}}$, $\mathbf{A} = 1, 2$. We will, therefore, use these components as the main variables describing of positions on the observer's sky. Note that for a source which lies close to r^μ we have a simple, direct relation between $\tilde{r}^{\mathbf{A}}$ and the angular coordinates on the sky $\delta\theta^{\mathbf{A}}$ centered at r^μ , directed along e_1^μ and e_2^μ and expressed in radians. Namely, we have

$$\delta\theta^{\mathbf{A}} \approx \tilde{r}^{\mathbf{A}} \quad (38)$$

for $\delta\theta^{\mathbf{A}} \ll 1$. The relation for larger angles requires the use of the standard trigonometric formulae.

In the next sections, we will need to compare the positions on the sky of various objects as registered by observers with different 4-velocities and at different points in $N_\mathcal{O}$. In a general spacetime, this is not a trivial task, because, as we mentioned, the position on the sky is a vector in the observer-dependent space of directions. Two problems arise here:

1. how do we compare position vectors at different points,
2. how do we compare directions on the sky registered by observers boosted with respect to each other since the notion of a spatial vector is different for each of them.

It turns out that, owing to the flatness of $N_\mathcal{O}$, this is possible using the reference direction given by the null vector $l_\mathcal{O}$.

Since the region $N_\mathcal{O}$ is effectively flat in our approximation, we may simply identify the tangent spaces at all points with $T_\mathcal{O}\mathcal{M}$ using the parallel propagation. This is possible independently of the paths connecting points with \mathcal{O} we may choose for that purpose. Thus the problem of comparing vectors at different points is solved. We also introduce, consistently with the approach above, a parallel propagated SNF $(u_\mathcal{O}^\mu, e_{\mathbf{A}}^\mu, l_\mathcal{O}^\mu)$ from \mathcal{O} throughout the whole region $N_\mathcal{O}$. This way any vector or tensor can be directly compared componentwise with a corresponding object at another point.

From now on we assume that all equations are written using this type of parallel frame at $N_\mathcal{O}$ and a similar one at $N_\mathcal{E}$. Following Section II A, we will write $l^\mu = l_\mathcal{O}^\mu$ and from (5) and (6) we have simply $k^\mu = l_\mathcal{O}^\mu + \Delta l_\mathcal{O}^\mu$ for the other null geodesic. Then it is easy to see that for the observer $u_\mathcal{O}^\mu$ we have

$$\tilde{r}^{\mathbf{A}} = \frac{\Delta l_\mathcal{O}^{\mathbf{A}}}{u_{\mathcal{O}\sigma} (l_\mathcal{O}^\sigma + \Delta l_\mathcal{O}^\sigma)} = \frac{\Delta l_\mathcal{O}^{\mathbf{A}}}{u_{\mathcal{O}\sigma} l_\mathcal{O}^\sigma} \left(1 - \frac{\Delta l_\mathcal{O}^{\mathbf{0}}}{u_{\mathcal{O}\sigma} l_\mathcal{O}^\sigma} + \Delta l_\mathcal{O}^{\mathbf{3}} \right)^{-1} \quad (39)$$

in the SNF.

Consider now an observer with a different 4-velocity U^μ . Obviously he or she uses a different screen space, but we may introduce a SNF $(U^\mu, f_{\mathbf{A}}^\mu, l_\mathcal{O}^\mu)$ such that the spatial vectors are aligned along $e_{\mathbf{A}}^\mu$, i.e. $f_{\mathbf{A}}^\mu = e_{\mathbf{A}}^\mu + C_{\mathbf{A}} l_\mathcal{O}^\mu$, with appropriate C_1 and C_2 [1]. This way we can compare also the positions on the sky of observers boosted with respect to each other: both $u_\mathcal{O}^\mu$ and U^μ may use the fiducial null vector $l_\mathcal{O}^\mu$ as providing the reference direction on their skies

and the screen vectors $e_{\mathbf{A}}^{\mu}$ and $f_{\mathbf{A}}^{\mu}$ as the two perpendicular directions on the celestial sphere used for defining the two-dimensional angular distance from the fiducial direction. The two spatial components of \tilde{r} , i.e. $\tilde{r}^{\mathbf{A}}$, evaluated for both observers and expressed in the corresponding SNF's, may now be used to compare the registered directions on the sky among them.

The construction can be repeated at the emitter's end of the null geodesic. Given the emitter's momentary 4-velocity $u_{\mathcal{E}}^{\mu}$ and the tangent vector $\tilde{k}^{\mu} = l_{\mathcal{E}}^{\mu} + \Delta l_{\mathcal{E}}^{\mu}$ we can define the *viewing direction vector* s^{μ} , pointing in the spatial direction (in emitter's frame) from which a given observer watches the events in $N_{\mathcal{E}}$:

$$s^{\mu} = - \left(\frac{1}{\tilde{k}_{\sigma} u_{\mathcal{E}}^{\sigma}} \tilde{k}^{\mu} + u_{\mathcal{E}}^{\mu} \right). \quad (40)$$

III. INFINITESIMAL DISPLACEMENT FORMULAS

BGO's are fundamental objects describing the properties of all geodesics in the vicinity of γ_0 . However, from the point of view of geometric optics, they contain too much information. Apart from null geodesics they also describe the spatial and timelike geodesics in the vicinity of γ_0 . On top of that, they distinguish differently parametrized null geodesics sharing the same path, which is indistinguishable in geometric optics. The affine parametrization carries no physical information and therefore the formalism effectively contains two gauge degrees of freedom corresponding to the affine reparametrizations of the null geodesics. We would like to isolate them in our formalism and focus on the remaining, physical degrees of freedom.

The condition for the displaced geodesic to remain null reads

$$g_{\mu\nu} (l_{\mathcal{O}}^{\mu} + \Delta l_{\mathcal{O}}^{\mu})(l_{\mathcal{O}}^{\nu} + \Delta l_{\mathcal{O}}^{\nu}) = 0. \quad (41)$$

It needs to be imposed only once along the geodesic, in this case at $\lambda = \lambda_{\mathcal{O}}$. We would also like to identify null geodesics sharing the same path. This means that we identify the initial data for which the initial tangent vector is proportional, while the initial points differ only by a vector proportional to the tangent. We identify therefore the initial data pairs of the form

$$\begin{pmatrix} \delta x_{\mathcal{O}}^{\mu} \\ \Delta l_{\mathcal{O}}^{\mu} \end{pmatrix} \sim \begin{pmatrix} \delta x_{\mathcal{O}}^{\mu} + C_1 (l_{\mathcal{O}}^{\mu} + \Delta l_{\mathcal{O}}^{\mu}) \\ \Delta l_{\mathcal{O}}^{\mu} + C_2 (l_{\mathcal{O}}^{\mu} + \Delta l_{\mathcal{O}}^{\mu}) \end{pmatrix} \quad (42)$$

for some non-vanishing constants C_1 and C_2 .

A. Distant observer approximation

In the rest of this paper, we will assume that we work within the validity regime of the relativistic distant observer approximation (DOA) or the relativistic counterpart of the paraxial approximation. Relativistic DOA is straightforward to explain mathematically, but it is more difficult to understand its physical meaning and its limits of applicability. Mathematically it boils down to linearizing all equations and relations involved in the displacement variables $\delta x_{\mathcal{O}}^{\mu}$ and $\Delta l_{\mathcal{O}}$, thereby neglecting all quadratic and higher terms, i.e. $\delta x_{\mathcal{O}}^{\alpha} \delta x_{\mathcal{O}}^{\beta}$, $\Delta l_{\mathcal{O}}^{\alpha} \Delta l_{\mathcal{O}}^{\beta}$ as well as the cross terms $\delta x_{\mathcal{O}}^{\alpha} \Delta l_{\mathcal{O}}^{\beta}$. This is consistent with the use of the first order GDE for the linearized relations between the deviation vectors δx and direction deviations Δl at $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$, ignoring this way higher order effects in light propagation. Physically this is equivalent to two distinct types of approximation:

1. *Flat light cones approximation (FLA)*. We assume we may neglect higher order terms in $\Delta l_{\mathcal{O}}$ in equation (41), in this way obtaining the null condition in the linearized form

$$\Delta l_{\mathcal{O}}^{\mu} l_{\mathcal{O}\mu} = 0. \quad (43)$$

In Section IV C we will show that this approximation is equivalent to assuming that the light cones originating on one end of \mathcal{O} degenerate to flat null hypersurfaces on the other end. In other words, the regions $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ are small enough that the bending of the surface of the light cone with apex in the opposite region is negligible.

2. *Parallel rays approximation (PRA)*. We assume we can drop the $\Delta l_{\mathcal{O}}^0$ and the $\Delta l_{\mathcal{O}}^3$ term in equation (39). This way we effectively linearize the whole expression in the direction vector deviation:

$$\tilde{r}^{\mathbf{A}} = \frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}}. \quad (44)$$

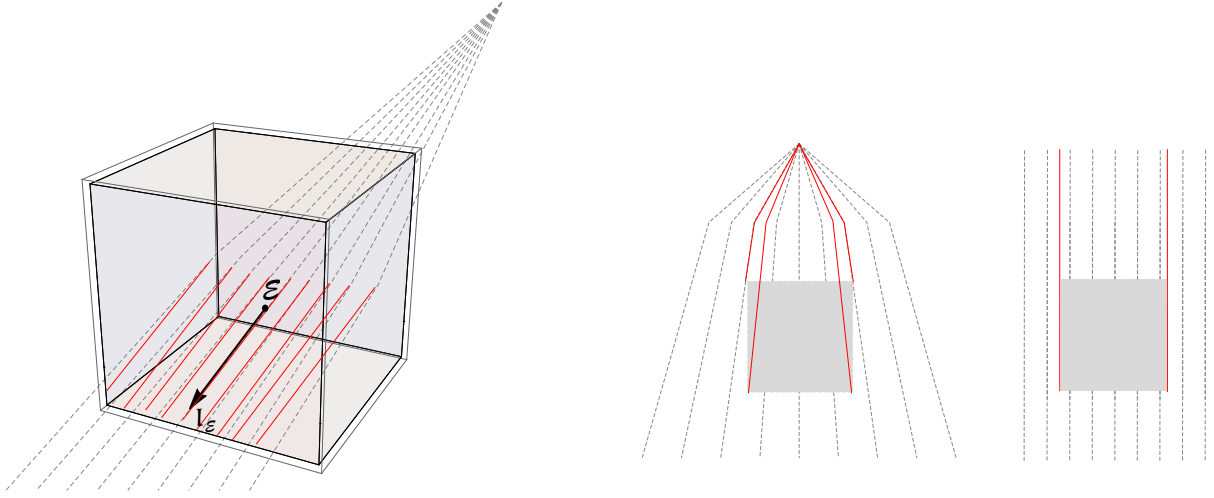


FIG. 2: In the parallel light rays approximation, we assume that the regions $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ are small enough and at the same time sufficiently separated that we can treat the light rays of null geodesics passing through them as parallel when discussing the relation of the direction deviation vector at \mathcal{O} to the displacement vectors at both ends, see the Figure on the left-hand side. This amounts to neglecting all possible perspective distortions of three-dimensional images and allows to consider only flat, two-dimensional projections to the screen space, see the Figure on the right-hand side.

This is equivalent to linearizing the whole mapping $V(u_{\mathcal{O}}, \cdot)$ around $l_{\mathcal{O}}$. It is straightforward to verify using equations (25)-(32) that this way we neglect the dependence of the position on the sky on the null SNF components $\delta x_{\mathcal{O}}^{\mathbf{3}}$ and $\delta x_{\mathcal{E}}^{\mathbf{3}}$, leaving just the transversal $\mathbf{1}$ and $\mathbf{2}$ as well as the timelike components $\mathbf{0}$ of the position deviation vectors. This, in turn, means that we treat all null rays considered within $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ as effectively parallel to each other (and in turn also to the original $l_{\mathcal{O}}$), neglecting the small change of the transversal position of the light rays between the front and the back of $N_{\mathcal{O}}$ due to the direction deviation, see Figure 2. Indeed, for sufficiently small $N_{\mathcal{E}}$ and $N_{\mathcal{O}}$, we may neglect the convergence or divergence of null light rays when discussing the relation between the direction deviation and displacement vectors. From the observational point of view, it is important to note that in PRA we neglect all perspective distortions of the images as perceived by the observers in $N_{\mathcal{O}}$.

We will now consider the limits of applicability of both approximations in a concrete physical situation. We will begin by the flat case and then move to the case of a general spacetime with an arbitrary metric.

a. Limits of applicability - flat case. In the Minkowski space we have simply $W_{XX}^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu}$ and $W_{XL}^{\mu}{}_{\nu} = (\lambda_{\mathcal{E}} - \lambda_{\mathcal{O}}) \delta^{\mu}{}_{\nu}$. Let $x_{\mathcal{E}}^{\mu}$ and $x_{\mathcal{O}}^{\mu}$ denote the coordinates of \mathcal{E} and \mathcal{O} respectively and let $d_{\mathcal{O}}$ denote the spatial distance between \mathcal{O} and \mathcal{E} in the observer's frame, given by $d_{\mathcal{O}} = (x_{\mathcal{E}}^{\mu} - x_{\mathcal{O}}^{\mu}) u_{\mathcal{O}\mu}$. The reader may verify that in that case the following expansions for the dimensionless combinations of the $\Delta l_{\mathcal{O}}$ components are valid:

$$\begin{aligned} \frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} &= O\left(\frac{L}{d_{\mathcal{O}}}\right) \\ \Delta l_{\mathcal{O}}^{\mathbf{3}} &= O\left(\frac{L}{d_{\mathcal{O}}}\right) \\ \frac{\Delta l_{\mathcal{O}}^{\mathbf{0}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} &= O\left(\frac{L^2}{d_{\mathcal{O}}^2}\right). \end{aligned}$$

(note that the component $\Delta l_{\mathcal{O}}^{\mathbf{3}}$ is by definition dimensionless, unlike the other two which require dividing by $l_{\mathcal{O}\mu} u_{\mathcal{O}}^{\mu}$). We may substitute these expansions to (39) and (41); it follows then easily that both FLA and PRA are equivalent to neglecting the subleading, quadratic terms in the $\frac{L}{d_{\mathcal{O}}}$ expansion, while keeping the linear ones. Also, the angular size of the region $N_{\mathcal{E}}$ on the observer's sky expressed in radians, estimated by applying the expansion above to (44), scales linearly like $\frac{L}{d_{\mathcal{O}}}$. This justifies the use of equation (38) for the angular variables $\delta\theta^{\mathbf{A}}$ determining the position in the sky.

Obviously, the approximations work well as long as the spatial distance between both regions is much larger than their size. This justifies the name distant observer approximation. Note, however, that the applicability of both PRA

and FLA depends not only on the position of the points \mathcal{O} and \mathcal{E} but also on the observer's 4-velocity $u_{\mathcal{O}}$. Even if they work well for a given $u_{\mathcal{O}}$ they may fail for another, strongly boosted observer, because in his or her frame the spatial distance $d_{\mathcal{O}}$ will be much smaller due to the Lorentz contraction. This, in turn, will lead to a relative increase of the subleading terms in the $\frac{L}{d_{\mathcal{O}}}$ expansion and thus to the increase of the perspective distortions and the light cone bending effects. On the other hand, we point out that this problem can always be cured by shrinking the regions $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ appropriately and thus decreasing L . This analysis should also apply to almost flat spacetimes and weak lensing.

b. Limits of applicability - general case. In a completely general spacetime drawing the precise limits of applicability of the FLA and PRA is much more difficult. Guided by the results for the flat case we assume that the following dimensionless combinations of components of $\Delta l_{\mathcal{O}}$, expressed in the SNF, are small:

$$\frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} = \frac{1}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} W_{XL}^{-1\mathbf{A}}{}_{\nu} (\delta x_{\mathcal{E}}^{\nu} - W_{XX}{}^{\nu}{}_{\sigma} \delta x_{\mathcal{O}}^{\sigma}) \ll 1 \quad (45)$$

$$\Delta l_{\mathcal{O}}^{\mathbf{3}} = W_{XL}^{-1\mathbf{3}}{}_{\nu} (\delta x_{\mathcal{E}}^{\nu} - W_{XX}{}^{\nu}{}_{\sigma} \delta x_{\mathcal{O}}^{\sigma}) \ll 1. \quad (46)$$

These equations can be viewed as conditions for the BGO's W_{XX} and W_{XL} , the observer's 4-velocity $u_{\mathcal{O}}$ and the size L of both domains determining the scale of the terms $\delta x_{\mathcal{O}}$ and $\delta x_{\mathcal{E}}$. Substituting these relations to equation (41) expressed in the SNF yields an expansion for the remaining component of $\Delta l_{\mathcal{O}}$:

$$\frac{\Delta l_{\mathcal{O}}^{\mathbf{0}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} = O\left(\left(\frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}}\right)^2\right).$$

It is quadratic, therefore, negligible in comparison with the other three components. Neglecting $\Delta l_{\mathcal{O}}^{\mathbf{0}}$, on the other hand, is precisely equivalent to the FLA defined by equation (43).

As for the PRA, we note that with conditions (45)-(46) we have

$$\tilde{r}^{\mathbf{A}} = \frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} + O\left(\frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} \cdot \Delta l_{\mathcal{O}}^{\mathbf{3}}\right).$$

Again keeping the leading, linear order is equivalent to the PRA condition (44). Just like in the flat case, for a given spacetime and fixed \mathcal{E} and \mathcal{O} the conditions (45)-(46) are sensitive to the 4-velocity of the observer. But in contrast to the flat case, there is no a priori relation between the BGO's and the distance between the two points \mathcal{O} and \mathcal{E} , because in general both W_{XX} and W_{XL} may take any value limited only by conditions (25), (27), (29) and (31). In fact, with very strong lensing the applicability of the DOA for fixed L and observer can even begin to *decrease* further down γ_0 as we move away from \mathcal{O} . We may illustrate this with a simple example: with very strong lensing along the way, the beam centered at \mathcal{O} may become strongly convergent or divergent far down γ_0 , introducing this way strong perspective distortions observable for regions of size L and violating this way the assumptions of the PRA. However, just like before, the applicability of both approximations can always be restored if we narrow down $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ appropriately.

Summarizing, the applicability of the DOA in a spacetime with an arbitrary metric and for any given observer is a subtle problem and needs to be tested on a case by case basis, for example, using conditions (45)-(46). In astrophysics, it is assumed they hold automatically because of the extreme ratio between the distances between \mathcal{O} and \mathcal{E} and their sizes (measured in a typical observer's frame, for example, the CMB rest frame), but if the influence of the curvature is strong or the observer is sufficiently boosted this must be done with care.

In principle it is possible to use the non-linearized relations (41) and (39) for the observables, taking into account this way the perspective effects and the effects of the bending of the lightcone surfaces, but mathematical consistency requires then to include also the nonlinear terms in the relations between the deviations of the geodesics in $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$, for example with the help of the higher order GDE's [34–36]. This is significantly more cumbersome than the formalism presented here and we leave this for future studies.

c. Quotient spaces. Within the regime of validity of the PRA we may simplify the equivalence relation of (42) to

$$\begin{pmatrix} \delta x_{\mathcal{O}}^{\mu} \\ \Delta l_{\mathcal{O}}^{\mu} \end{pmatrix} \sim \begin{pmatrix} \delta x_{\mathcal{O}}^{\mu} + C_1 l_{\mathcal{O}}^{\mu} \\ \Delta l_{\mathcal{O}}^{\mu} + C_2 l_{\mathcal{O}}^{\mu} \end{pmatrix} \quad (47)$$

i.e. we identify the initial data for which the position and directional deviations only differ by a multiple of $l_{\mathcal{O}}^{\mu}$. From (25)-(28) it is easy to see that adding this type of terms in $N_{\mathcal{O}}$ leads to a similar change in the final data:

$$\begin{pmatrix} \delta x_{\mathcal{E}}^{\mu} \\ \Delta l_{\mathcal{E}}^{\mu} \end{pmatrix} \sim \begin{pmatrix} \delta x_{\mathcal{E}}^{\mu} + D_1 l_{\mathcal{E}}^{\mu} \\ \Delta l_{\mathcal{E}}^{\mu} + D_2 l_{\mathcal{E}}^{\mu} \end{pmatrix} \quad (48)$$

with constants D_1 and D_2 related to C_1 and C_2 .

This invariance leads to the following idea: instead of considering the optical operators as acting from one tangent space to another one we may consider them as mappings between the appropriate quotient spaces $\mathcal{Q}_O = T_O\mathcal{M}/l_O$ and $\mathcal{Q}_E = T_E\mathcal{M}/l_E$, see Figure 3. Namely, let \mathcal{Q}_E be the quotient of $T_E\mathcal{M}$ by the equivalence relation $X^\mu \sim Y^\mu$ iff $X^\mu = Y^\mu + cl_E^\mu$ for any real number c . \mathcal{Q}_O can be defined in an analogous way, with l_E replaced by l_O and $T_E\mathcal{M}$ replaced by $T_O\mathcal{M}$.

Within these 3-dimensional spaces we also consider the 2-dimensional subspaces orthogonal to l_O^μ and l_E^μ respectively, i.e. $\mathcal{P}_O = l_O^\perp/l_O$ and $\mathcal{P}_E = l_E^\perp/l_E$, see again Figure 3. They will be referred to as the *perpendicular spaces*, see again Figure 3. Unlike \mathcal{Q}_O and \mathcal{Q}_E , the perpendicular spaces \mathcal{P}_O and \mathcal{P}_E inherit a positive definite metric q from the Lorentzian spacetime metric g [1]: let \mathbf{X} and \mathbf{Y} be two vectors in \mathcal{P}_E and let X and Y be any two corresponding vectors in $T_E\mathcal{M}$, i.e. $\mathbf{X} = [X]$ and $\mathbf{Y} = [Y]$, where $[\cdot]$ denotes the linear operation of taking the equivalence class of a vector in $T_E\mathcal{M}$ with respect to the relation \sim defined above. The reader may check that the formula $q(\mathbf{X}, \mathbf{Y}) = g(X, Y)$ defines the same value of the scalar product of \mathbf{X} and \mathbf{Y} irrespective of the choice of the equivalence class representatives X and Y .

The angles and distances calculated using q correspond to the angles and distances measured by any observer on points projected down to the plane perpendicular to the direction of observation (the Sachs's screen space) along the null direction of light propagation [1]. This is a reformulation of the well-known Sachs shadow theorem [54, 55] in terms of the quotient spaces. This fact gives the geometry of the perpendicular spaces an explicitly observer-invariant meaning and we explore it later in the paper in order to separate the dependence of observables on the observer's and emitter's frame and on the spacetime geometry.

We end this subsection by noting that both quotient spaces have a very simple interpretation in terms of the vector components in a SNF. Namely, they correspond to vectors with a "forgotten" fourth component corresponding to $\boldsymbol{\mu} = \mathbf{3}$. The mapping $[\cdot]$ has then a simple form of $(X^0, X^1, X^2, X^3) \rightarrow (X^0, X^1, X^2)$. Vectors in any perpendicular subspace \mathcal{P} along γ_0 have additionally vanishing first component, i.e. $(0, X^1, X^2)$.

The reader may check that in the GDE (7) expressed in the SNF the equations for the first three components decouple from the fourth one and that vectors with vanishing $\boldsymbol{\mu} = \mathbf{0}$ component form a subspace of solutions. This way we see that the quotient spaces defined above are compatible with the properties of the first order GDE around a null geodesic.

B. Time lapse formula

We multiply equation (10) by $l_{E\mu}$ from both sides and make use of equations (29), (31) to obtain the following relation:

$$l_{E\mu} \delta x_E^\mu = l_{O\mu} \delta x_O^\mu + (\lambda_E - \lambda_O) \Delta l_O^\mu l_{O\mu}.$$

Comparing this formula with (43) we see that the perturbed geodesic is null in the DOA iff

$$l_{E\mu} \delta x_E^\mu = l_{O\mu} \delta x_O^\mu. \quad (49)$$

We will refer to this equation as the *time lapse formula*. We have just proved that within the FLA any two points in N_O and N_E respectively can be connected by a null geodesic iff their deviation vectors satisfy the linear condition (49). This is the first important result of this paper, because it gives immediately the relation between the time lapse in N_E and the time lapse in N_O , as registered by the observers, and formulated in a coordinate- and frame-independent way. Namely, for any observer whose worldline intersects N_O we may relate his or her proper time to the "null time" $\chi = l_{O\mu} \delta x_O^\mu$ (for an effectively flat space this is a simple special relativity problem). Equation (49) shows then that at a given moment the observer can see only the events in N_E which lie on the null hypersurface given by $l_{E\mu} \delta x_E^\mu = \chi$. This a simple and elegant way to take into account the Rømer delay between the regions N_O and N_E due to the finite light speed.

Geometrically, the conditions $l_{O\mu} \delta x_O^\mu = \text{const}$ and $l_{E\mu} \delta x_E^\mu = \text{const}$ define foliations of N_O and N_E by families of null hypersurfaces, see Figure 4. We can, therefore, rephrase the result of this section as follows: the null vectors l_E and l_O define two foliations by null hypersurfaces in the two locally flat regions. For observers located on a leaf in N_O only the points lying on the corresponding leaf in N_E can be reached by a null geodesic. It is also clear that while the null tangent vectors l_O^μ and l_E^μ are defined only up to a common rescaling according to (3), this ambiguity does not affect the two foliations or the correspondence between their leaves.

The two null foliations at two ends of γ_0 are in fact the degenerate families of light cones centered at the opposite ends of γ_0 . If we pick a point p in N_O then its past light cone in N_E will degenerate to a flat hypersurface due to the large distance between the two regions and their small size. Moreover, the future light cone of *any* point on that

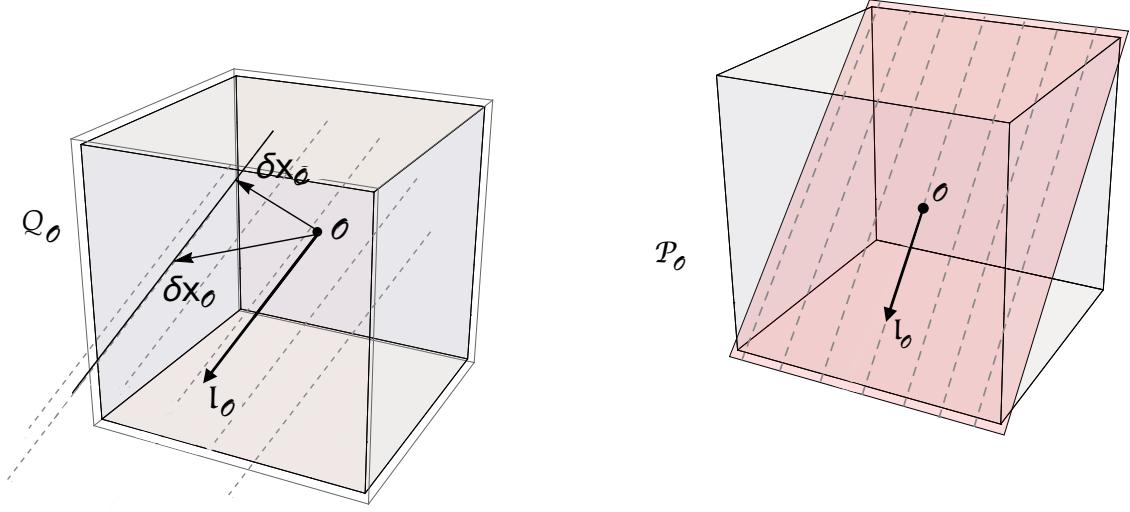


FIG. 3: Geometry of the quotient spaces \mathcal{Q}_O , \mathcal{P}_O and their counterparts on the other side. Elements of \mathcal{Q}_O correspond to the vectors in $T_O\mathcal{M}$, or points in N_O , identified if they are separated by a multiple of l_O . Geometrically it is the space of null straight lines in N_O parallel to γ_0 . The two-dimensional subspace \mathcal{P}_O corresponds to null straight lines parallel to γ_0 which additionally lie on the null hyperplane orthogonal to l_O .

null hypersurface will degenerate in N_O to the same null hypersurface containing p : in the FLA the light cones with apices contained within a single foliation leaf look exactly the same on the other end of γ_0 . The same argument works with the role of the emission and observation point reversed. These observations explain the name we have coined for flat light cone approximation (43). Physically, the flatness of light cones means that the Rømer delay is in a good approximation independent of the perpendicular displacements of the observer and the source because of the large distance between them.

We may pull back formula (49) to the quotient spaces \mathcal{Q}_O and \mathcal{Q}_ε . Note that the products on both sides are insensitive to adding multiples of l_O^μ and l_ε^μ to the displacement vectors. We may therefore consider $g(\cdot, l_O)$ and $g(\cdot, l_\varepsilon)$ as mappings defined on the respective quotient spaces and write

$$g([\delta x_O], l_O) = g([\delta x_\varepsilon], l_\varepsilon). \quad (50)$$

C. Direction variation formula

We will now recast formula (11) applied to deviations satisfying (43) in a more convenient form. We begin by rewriting it as follows:

$$W_{XL}{}^\mu{}_\nu \Delta l_O^\nu = \delta x_\varepsilon^\mu - W_{XX}{}^\mu{}_\nu \delta x_O^\nu \quad (51)$$

Equation (51) can be pulled back to the quotient spaces \mathcal{Q}_O and \mathcal{Q}_ε respectively.

The operator W_{XL} defines an operator acting on the perpendicular space \mathcal{P}_O with the image in \mathcal{P}_ε . Namely, given a vector $\mathbf{Y} \in \mathcal{P}_O$ we may pick any vector $Y \in T_O\mathcal{M}$ such that $[Y] = \mathbf{Y}$ and define

$$\mathcal{D}(\mathbf{Y}) = [W_{XL}(Y)].$$

It is straightforward to verify using (25)-(32) that the resulting vector in \mathcal{Q}_ε does not depend on the choice of the representative Y of the equivalence class and that the resulting vector is always perpendicular to l_ε^μ . Thus we have

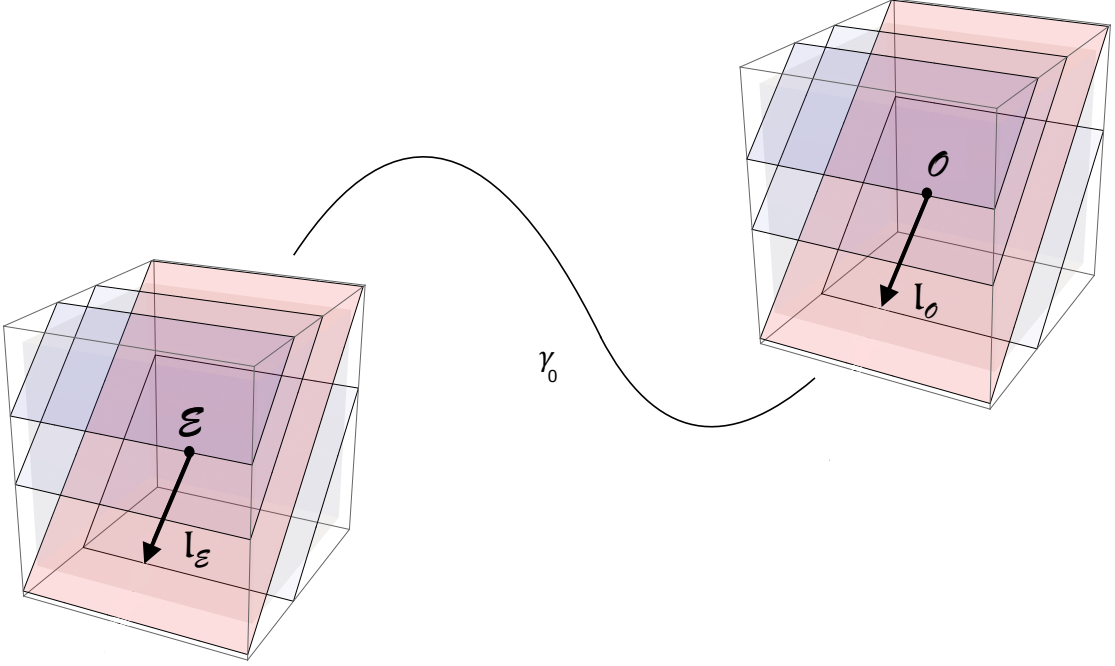


FIG. 4: Null tangent vectors $l_{\mathcal{O}}$ at \mathcal{O} and $l_{\mathcal{E}}$ at \mathcal{E} define two corresponding foliations by null hypersurfaces orthogonal to $l_{\mathcal{O}}$ and $l_{\mathcal{E}}$ respectively. In the flat light cones approximation, only the pairs of points lying on the corresponding leaves can be connected by a null geodesic. This requirement restricts the events in $N_{\mathcal{E}}$ from which an observer in $N_{\mathcal{O}}$ can register signals at a given instant of time.

defined a linear mapping

$$\mathcal{D} : \mathcal{P}_{\mathcal{O}} \rightarrow \mathcal{P}_{\mathcal{E}}.$$

This mapping can be represented by a 2-by-2 matrix and is known in the relativistic geometric optics as the Jacobi map [1, 54]. The Jacobi map as defined here depends on the parametrization we have chosen for the fiducial geodesic γ_0 . Under an affine reparametrization (2)–(3) it rescales according to the formula $\mathcal{D} \rightarrow E \cdot \mathcal{D}$.

Similar reasoning can be applied to W_{XX} : given $\mathbf{Z} \in \mathcal{Q}_{\mathcal{O}}$ we may take any $Z \in T_{\mathcal{O}}\mathcal{M}$ such that $[Z] = \mathbf{Z}$ and define

$$\mathcal{W}(\mathbf{Z}) = [W_{XX}(Z)]. \quad (52)$$

We check using (25)–(32) that this defines a linear mapping between the quotient spaces

$$\mathcal{W} : \mathcal{Q}_{\mathcal{O}} \rightarrow \mathcal{Q}_{\mathcal{E}}.$$

\mathcal{W} can be used directly, but since we are interested in the curvature effects on the light rays we have found it convenient to separate out the “non-flat” contribution to the operator, related directly to the curvature. In a flat space we have $W_{XX}{}^{\mu}{}_{\nu} = T^{\mu}{}_{\nu}$, where T is the parallel transport operator from \mathcal{O} to \mathcal{E} . Let \mathcal{T} denote the pullback of T to the space $\mathcal{Q}_{\mathcal{E}}$, i.e. mapping $\mathcal{T} : \mathcal{Q}_{\mathcal{O}} \rightarrow \mathcal{Q}_{\mathcal{E}}$ given by the formula

$$\mathcal{T}(\mathbf{Z}) = [T(Z)],$$

where $\mathbf{Z} \in \mathcal{Q}_{\mathcal{O}}$ and $Z \in T_{\mathcal{O}}\mathcal{M}$ is a vector such that $[Z] = \mathbf{Z}$. With this setup we may define

$$m = \mathcal{W} - \mathcal{T}.$$

The operator m will be referred to as the *emitter-observer asymmetry operator*. While the domain of m consists of the whole space $\mathcal{Q}_\mathcal{O}$ its image is automatically perpendicular to $l_\mathcal{E}^\mu$ – this follows from (29) and the elementary property of the parallel transport $l_{\mathcal{E}\mu} T^\mu_\nu = l_{\mathcal{O}\nu}$. Thus the emitter-observer asymmetry operator is a mapping

$$m : \mathcal{Q}_\mathcal{O} \rightarrow \mathcal{P}_\mathcal{E}.$$

Since the dimension of $\mathcal{Q}_\mathcal{O}$ is 3, while the dimension of $\mathcal{P}_\mathcal{E}$ is 2, the operator m is always degenerate. Unlike the Jacobi map, it is insensitive to affine reparametrizations of the fiducial null geodesic γ_0 , i.e. $m \rightarrow m$ if $\lambda \rightarrow \lambda' = A\lambda + B$. m has already appeared in [1], although without a full discussion of its properties. The pair of bilocal operators \mathcal{D} and m will be referred to as the *optical operators*.

We may rewrite the pullback of the whole equation (51) using the optical operators:

$$\mathcal{D}([\Delta l_\mathcal{O}]) = [\delta x_\mathcal{E}] - [\delta \hat{x}_\mathcal{O}] - m([\delta x_\mathcal{O}],$$

where $\delta \hat{x}_\mathcal{O} \equiv \mathcal{T}(\delta x_\mathcal{O})$ is a shorthand notation for the parallel transport from \mathcal{O} to \mathcal{E} . In the final step we can use the linearity of the projection $[\cdot]$ to the quotient space $\mathcal{Q}_\mathcal{E}$ to put both deviation vector inside a common square bracket and obtain the *direction deviation equation*:

$$\mathcal{D}([\Delta l_\mathcal{O}]) = [\delta x_\mathcal{E} - \delta \hat{x}_\mathcal{O}] - m([\delta x_\mathcal{O}]) \quad (53)$$

Note that as long as (49) is satisfied the combination $\delta x_\mathcal{E}^\mu - \delta \hat{x}_\mathcal{O}^\mu$ is perpendicular to $l_\mathcal{E}^\mu$ and the first term on the right hand side of (53) is a vector in $\mathcal{P}_\mathcal{E}$, even though each position deviation vector individually does not need to be perpendicular to $l_\mathcal{O}$ or $l_\mathcal{E}$. This is of course in full agreement with the fact that the left-hand side and the second term on the right-hand side are automatically in $\mathcal{P}_\mathcal{E}$. On the other hand, we see that (53) is impossible to satisfy when applied to a pair of deviation vectors not satisfying the time lapse formula.

Equations (50) and (53) describe completely how observers in $N_\mathcal{O}$ register signals coming from events in $N_\mathcal{E}$: (50) defines which events in $N_\mathcal{E}$ are visible from which points in $N_\mathcal{O}$ while (53) yields the direction from which the observers perceive the light coming. The optical properties of the spacetime between those regions are therefore completely contained in the following geometric structures:

1. The pair of null vectors $l_\mathcal{O}$ and $l_\mathcal{E}$, given up to a common rescaling. They define three important ingredients of the geometry: $l_\mathcal{O}$ gives the null direction from which an observer in \mathcal{O} sees the light coming from \mathcal{E} , this way providing the reference direction for all other observations. $l_\mathcal{E}$, on the other hand, gives null direction from which he or she effectively observes the region $N_\mathcal{E}$. Finally, both vectors together define the two null foliations in $N_\mathcal{O}$ and $N_\mathcal{E}$ relating the time lapse in both regions.
2. The Jacobi map \mathcal{D} acting between the perpendicular spaces on the two ends of the null geodesic (also defined up to the same rescaling), encoding the effects of gravitational lensing between \mathcal{O} and \mathcal{E} .
3. The parallel transport operator T along γ_0 , preserving the metric.
4. The emitter-observer asymmetry map m from $\mathcal{Q}_\mathcal{O}$ to $\mathcal{P}_\mathcal{E}$. It is related to the parallax effects and the apparent position drift. Its function will be explained in detail in Section IV C.

All these objects exist independently of any other structures in M , in particular independently from the choice of the coordinate systems, observers and frames in $N_\mathcal{O}$ and $N_\mathcal{E}$. The last two operators can be in fact be combined into \mathcal{W} from equation (52), but we have opted to consider them separately because m turns out to have a particularly elegant geometric interpretation.

Geometrically the pair of equations (50) and (53) can be seen as a linearized transformation between the two ways in which we may parametrize null geodesics passing through $N_\mathcal{O}$ and $N_\mathcal{E}$. On the one hand, we can parametrize them by giving the initial point $\delta x_\mathcal{O}^\mu$ and the initial direction, specified by the direction deviation vector $\Delta l_\mathcal{O}^\mu$. In the DOA and for geodesics considered without parametrization all we need is the equivalence class $[\delta x_\mathcal{O}]$, i.e. a vector in the 3-dimensional space $\mathcal{Q}_\mathcal{O}$, and the equivalence class of $\Delta l_\mathcal{O}^\mu$, also in $\mathcal{Q}_\mathcal{p}$. The null condition (41) restricts the choice of $[\Delta l_\mathcal{O}]$ to the 2-dimensional subspace $\mathcal{P}_\mathcal{O}$. The space of geodesics we consider is therefore 5-dimensional.

On the other hand, we may also parametrize the null geodesics by the endpoints $\delta x_\mathcal{E}$ and $\delta x_\mathcal{O}$, or – more precisely – by their equivalence classes in $\mathcal{Q}_\mathcal{O}$ and $\mathcal{Q}_\mathcal{E}$ respectively. In the second parametrization, we need to impose the linear condition given by the time lapse formula (50). The resulting dimension of the null geodesic family is therefore again $3 + 3 - 1 = 5$. The direction deviation vector at \mathcal{O} is then given by (53) up to an irrelevant multiple of $l_\mathcal{O}^\mu$.

D. Optical operators from the Riemann curvature tensor

We will show that the two optical mappings are effectively linear functionals of the spacetime curvature tensor along γ_0 , given by solutions of linear ODE's. Recall that W_{XX} and W_{XL} expressed in a parallel propagated SNF can be obtained from the curvature tensor directly via (12)-(16).

Given any SNF $(u^\mu, e_A^\mu, l_\mathcal{O}^\mu)$ we can construct a corresponding frame $(\mathbf{u}, \mathbf{e}_A)$ in $\mathcal{Q}_\mathcal{O}$ by simply taking $\mathbf{u} = [u]$, $\mathbf{e}_A = [e_A]$ and a frame (\mathbf{e}_A) in $\mathcal{P}_\mathcal{O}$. By parallel propagating the SNF and repeating this procedure we obtain similar parallel propagated frames $(\hat{u}^\mu, \hat{e}_A^\mu, \hat{l}_\mathcal{O}^\mu)$, $(\hat{\mathbf{u}}, \hat{\mathbf{e}}_A)$ and $(\hat{\mathbf{e}}_A)$, in $T_p\mathcal{M}$, \mathcal{Q}_p and \mathcal{P}_p respectively, defined now at all points p along γ_0 . Note that the parallel propagation $\hat{l}_\mathcal{O}^\mu$ of $l_\mathcal{O}^\mu$ along γ_0 is simply the tangent vector l^μ , because the tangent vector along a geodesic is always parallel propagated. Moreover, at \mathcal{E} we have $\hat{l}_\mathcal{O}^\mu = l_\mathcal{E}^\mu$.

We will now prove a direct relation between the optical operators expressed in the frame $(\hat{\mathbf{u}}, \hat{\mathbf{e}}_A)$ and submatrices of W_{XX} and W_{XL} expressed in the SNF. Let $\xi_0 \in \mathcal{P}_\mathcal{O}$, with the decomposition $\xi_0 = \xi_0^A e_A$. Take the vector ξ_0 in $T_\mathcal{O}\mathcal{M}$ given by $\xi_0 = \xi_0^A e_A$ as the corresponding vector such that $\xi_0 = [\xi_0]$. Applying the definition we get

$$\mathcal{D}(\xi_0) = [W_{XL}(\xi_0)] = [W_{XL}(\xi_0^B e_B)] = \xi_0^B [W_{XL}(e_B)].$$

On the other hand we know that

$$[W_{XL}(e_B)] = [W_{XL}{}^0{}_B \hat{u} + W_{XL}{}^A{}_B \hat{e}_A + W_{XL}{}^3{}_B l].$$

The first term vanishes because of (31): since e_B is orthogonal to $l_\mathcal{O}$, then $W_{XL}(e_B)$ must be orthogonal to $l_\mathcal{E}$. It follows that $W_{XL}{}^0{}_B = 0$ in the semi-null frame. The last term $[W_{XL}{}^3{}_B l]$ on the other hand vanishes because at \mathcal{E} we have $l = l_\mathcal{E}$ and the equivalence class of $l_\mathcal{E}$ in $\mathcal{Q}_\mathcal{E}$ is by definition 0. We are thus left with

$$\mathcal{D}(\xi_0) = W_{XL}{}^A{}_B \xi_0^B \hat{e}_A$$

or

$$\mathcal{D}^A{}_B = W_{XL}{}^A{}_B \quad (54)$$

in the aforementioned frames. Thus the Jacobi operator turns out to be a two-by-two submatrix of W_{XL} , independently of the choice of the SNF. Similar arguments prove m expressed in $(\hat{\mathbf{u}}, \hat{\mathbf{e}}_A)$ satisfies

$$m^A{}_i = W_{XX}{}^A{}_i - T^A{}_i, \quad (55)$$

where i runs through 0, 1, 2, W_{XX} is expressed in the SNF and T is the parallel transport operator in that frame:

$$T^A{}_i = \delta^A{}_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

We will use the results proved above to derive linear ODE's which relate the optical operators directly to the curvature. We note that in the ODE (17) for W_{XL} the equations for the submatrix (54) decouple from the rest of the equations. Namely, the relevant components of the ODE's read

$$\ddot{B}^A{}_B - R^A{}_{\mu\nu 0} l^\mu l^\nu B^0{}_B - R^A{}_{\mu\nu C} l^\mu l^\nu B^C{}_B - R^A{}_{\mu\nu 3} l^\mu l^\nu B^3{}_B = 0.$$

We know that $R^A{}_{\mu\nu 3} l^\mu l^\nu = R^A{}_{333} = 0$ because of the symmetries of the Riemann tensor. On the other hand, $B^0{}_B$ must vanish, because $B^\mu{}_\nu$ must satisfy (31) all along γ_0 . We are thus left with the following equation for $\mathcal{D}^A{}_B = B^A{}_B$ along γ_0 :

$$\ddot{\mathcal{D}}^A{}_B - R^A{}_{\mu\nu C} l^\mu l^\nu \mathcal{D}^C{}_B = 0 \quad (56)$$

with the initial data

$$\mathcal{D}^A{}_B(\lambda_\mathcal{O}) = 0 \quad (57)$$

$$\dot{\mathcal{D}}^A{}_B(\lambda_\mathcal{O}) = \delta^A{}_B. \quad (58)$$

The value of the solution at the endpoint \mathcal{E} gives the Jacobi map in the frame $(\hat{\mathbf{e}}_A)$ irrespective of the choice of the corresponding SNF. The derivation of the ODE for m proceeds in a similar way. We begin by noting that in equation (12) the ODE's for the components $A^A{}_i$ decouple from the rest:

$$\ddot{A}^A{}_i - R^A{}_{\mu\nu 0} l^\mu l^\nu A^0{}_i - R^A{}_{\mu\nu C} l^\mu l^\nu A^C{}_i - R^A{}_{\mu\nu 3} l^\mu l^\nu A^3{}_i = 0.$$

The fourth term vanishes again because $R^A_{\mu\nu\beta} l^\mu l^\nu = 0$ due to the symmetries of the Riemann. The second one survives, but it can be simplified using the algebraic properties of W_{XX} . Namely, $A^0_\nu = \delta^0_\nu$ because of (15) and (29). We get therefore the following ODE for W_{XX} in a SNF:

$$\begin{aligned} \ddot{W}_{XX}{}^A{}_i - R^A_{\mu\nu C} l^\mu l^\nu W_{XX}{}^C{}_i &= R^A_{\mu\nu 0} l^\mu l^\nu \delta^0_i \\ W_{XX}{}^A{}_i(\lambda_{\mathcal{O}}) &= \delta^A_i \\ \dot{W}_{XX}{}^A{}_i(\lambda_{\mathcal{O}}) &= 0. \end{aligned}$$

In the final step we rewrite the equations above replacing the components W_{XX} with the corresponding components of m as variables. The relation between the variables is given by (55). The ODE's take the form of

$$\begin{aligned} \ddot{m}^A{}_i - R^A_{\mu\nu C} l^\mu l^\nu m^C{}_i &= R^A_{\mu\nu i} l^\mu l^\nu & (59) \\ m^A{}_i(\lambda_{\mathcal{O}}) &= 0 & (60) \\ \dot{m}^A{}_i(\lambda_{\mathcal{O}}) &= 0. & (61) \end{aligned}$$

The value of the solution at $\lambda_{\mathcal{E}}$ yields m in the parallel-propagated SNF.

We immediately note that the ODE is an inhomogeneous version of (56) with two more components involved. The inhomogeneity is proportional to the Riemann tensor and the initial data is vanishing at $\lambda_{\mathcal{O}}$. Thus, in a flat space, the resulting operator always vanishes, unlike the Jacobi map. A non-vanishing mapping m is therefore always an effect of the spacetime curvature.

Just like the four BGO's, m and \mathcal{D} are two different functionals of the Riemann tensor (or, more precisely, the optical tidal matrix $R^\alpha_{\mu\nu\beta} l^\mu l^\nu$) along the fiducial null geodesic, given by the solutions two different matrix ODE's. Therefore their values are in general unrelated to each other.

E. Remarks

a. Direction deviation formula in a semi-null frame. The derivation of the direction variation formula (53) has been presented in an abstract and covariant manner, highlighting this way the coordinate- and frame-invariance of the formalism. It is nevertheless very instructive to rewrite it in a parallel-propagated SNF. It takes the form of

$$\mathcal{D}^A{}_B \Delta l_{\mathcal{O}}^B = (\delta x_{\mathcal{E}} - \delta \hat{x}_{\mathcal{O}})^A - m^A{}_i \delta x_{\mathcal{O}}^i. \quad (62)$$

b. The perpendicular part of the emitter-observer asymmetry operator. We have already seen when discussing the Jacobi operator that the components corresponding to two spatial directions perpendicular to u tend to decouple from the other two in the GDE. This suggests that a similar independence may be present in the spatial submatrix of m . In order to show it we define the *perpendicular part of the emitter-observer asymmetry operator* $m_{\perp} : \mathcal{P}_{\mathcal{O}} \rightarrow \mathcal{P}_{\mathcal{E}}$ as the restriction of m to the subspace $\mathcal{P}_{\mathcal{O}}$:

$$m_{\perp} = m|_{\mathcal{P}_{\mathcal{O}}}$$

In a SNF m_{\perp} can be represented by a 2-by-2 submatrix of m . The reader may also check that the ODE's (59)-(61) for its components $m_{\perp}{}^A{}_B \equiv m^A{}_B$ indeed decouple from the equations for the remaining two components $m^A{}_0$. We emphasize that while the perpendicular part m_{\perp} of m is defined independently of any observer, the splitting of m into the perpendicular part and the timelike components $m^0{}_A$ requires fixing an observer with a 4-velocity vector u .

c. The geodesic deviation equation as an ODE in the quotient spaces. The ODE systems (56)-(58) and (59)-(61) can also be interpreted as equations defined directly in the quotient spaces \mathcal{Q}_p and \mathcal{P}_p [1]. Namely, it is possible to define the covariant derivative ∇_l as a differentiation in the quotient spaces and pull back the optical tidal matrix to \mathcal{Q}_p . This way we avoid specifying a whole SNF including an observer u^μ . This interpretation highlights observer-independence of the spatial components of the operators considered.

d. The emitter-observer asymmetry operator extended to the full tangent space. It is possible to extend the domain of m to the whole tangent space, defining $m_{ext}(X) = m([X])$, i.e. with identically vanishing $m_{ext}(l)$. This is the way m has been defined in [1]. The extended operator satisfies a similar ODE of type

$$\begin{aligned} \ddot{m}_{ext}{}^A{}_\beta - R^A_{\mu\nu C} l^\mu l^\nu m_{ext}{}^C{}_\beta &= R^A_{\mu\nu\beta} l^\mu l^\nu \\ m_{ext}{}^A{}_\beta(\lambda_{\mathcal{O}}) &= 0 \\ \dot{m}_{ext}{}^A{}_\beta(\lambda_{\mathcal{O}}) &= 0. \end{aligned}$$

IV. PARALLAX, POSITION DRIFT AND THE PHYSICAL MEANING OF THE OPTICAL OPERATORS

In this section, we will discuss the physical meaning of the optical operators \mathcal{D} and m , and in particular, we will point out their relation to the parallax measurements, but before that, we need to define what precisely we mean by parallax in the context of general relativity.

A. Parallax in general relativity

The notion of parallax is straightforward to grasp in a flat space and in non-relativistic context, but its generalization to general relativity is more ambiguous [9], with various researchers using different definitions [9, 57–61]. We will, therefore, begin this section by a short clarification what one can mean by parallax in the context of general relativity.

In the broadest possible sense parallax is the difference of the apparent position of an object on the celestial sphere when regarded from at least two different points of view. The difference may be due to many effects, including the gravitational light bending. Note that the definition above requires a method of comparing the celestial spheres defined at a different point of the spacetime, and possibly registered by different observers. In GR this already introduces a great deal of ambiguity, as there exist infinitely many ways to identify points in two distinct celestial spheres. The problem is made even more complicated by the fact that the positions on the sky, defined as vectors in the sphere of direction $\text{Dir}(u_{\mathcal{O}})$, depend on the observer's 4-velocity $u_{\mathcal{O}}$ via the aberration effect, see equation (37). This means that the direction identification is also non-trivial among two observers at the same event but with different 4-velocities u and v . While infinitely many such identifications are possible, it is reasonable to require that they preserve the metric structure of the celestial sphere, i.e. the angle measured between any two points on the sky should be invariant under the identifications.

The definitions of parallax can be classified according to the way we compare the directions on the sky at different events and according to how we select the points between which we make the comparison. In the literature the possibilities have been considered:

a. Parallax with respect to the local inertial frame. If the region of spacetime in which we consider the measurements can be considered locally flat, like $N_{\mathcal{O}}$, then we can employ the parallel transport of vectors for the purpose of direction identification. Physically this assumption means that we use *local physical phenomena* and the *local geometry* to define the notion of parallel directions on the celestial spheres among all nearby observers. Assume we fix an observer with a given 4-velocity $u_{\mathcal{O}}$ at one of the observation points, say \mathcal{O} , and a compatible SNF $(u_{\mathcal{O}}, e_{\mathbf{A}}, l_{\mathcal{O}})$. Its parallel transport $(\hat{u}_{\mathcal{O}}, \hat{e}_{\mathbf{A}}, \hat{l}_{\mathcal{O}})$ defines corresponding frames at all points in $N_{\mathcal{O}}$. This way we have defined a way to compare the directions on the sky for all nearby observers whose 4-velocities are equal to $\hat{u}_{\mathcal{O}}$: the spatial vectors $\hat{e}_{\mathbf{A}}$ and the projection of the null direction $\hat{l}_{\mathcal{O}}$ define 3 orthogonal directions on the spheres of direction $\text{Dir}(\hat{u}_{\mathcal{O}})$ at each point. These, in turn, allow for angle-preserving identification of all other points among all spheres $\text{Dir}(\hat{u}_{\mathcal{O}})$.

In practice, the observations at different points will most likely be performed by observers with different 4-velocities $v \neq \hat{u}_{\mathcal{O}}$. A simple way around that is to assume we perform the corresponding boost of the celestial sphere right after the observations, subtracting this way the effects of the Bradley aberration and reconstructing this way the sky looks at a given point for a fictitious observer with 4-velocity $\hat{u}_{\mathcal{O}}$. The change of the apparent position of an object in that reference frame will be referred to as *the parallax with respect to a local inertial frame*.

b. Classic parallax. In general, the observations of a distant object performed at different events will register light emitted in different moments along the emitter's worldline. This means that the result of observations will not only depend on the spacetime geometry and the observer's frame, but also on the motion of the emitter. One may, however, devise a measurement in which many observers will deliberately measure the emitter's position at carefully chosen moments so that all of them register signals emitted exactly at the same moment \mathcal{E} . This can be achieved by appropriate timing of the observations. In the DOA this corresponds to performing observations within a single null hypersurface $\delta x_{\mathcal{O}}^{\mu} l_{\mathcal{O}\mu} = \text{const}$. The observers can be displaced in two perpendicular directions $e_{\mathbf{A}}$, but note the displacement in the direction of $l_{\mathcal{O}}$ does not involve any parallax in the DOA. In this measurement we are considering the parallax of a *single spacetime event* \mathcal{E} , therefore this notion of parallax seems to be the closest in spirit to the non-relativistic understanding of the term. Following Räsänen [9] it will be called the *classic parallax* (in [58, 60] it is referred to as the *trigonometric parallax*, and in [59] it is simply called *parallax* or *triangulation*). Since the position measurement is made by comparison with parallel-propagated basis it is a special case of the parallax with respect to the local inertial frame. This type of measurement is by definition completely independent of the motion of the emitting body, i.e. its momentary 4-velocity, 4-acceleration etc., what matters is only its exact position at a single instant of time.

c. Single worldline parallax. The measurements of the classic parallax must be made at points separated by spacelike vectors. They require therefore using more than one physical observer separated by sufficiently large dis-

tances. In astronomy, this is usually not feasible. Instead, we rely on observations performed along the worldline of a *single* observer, positioned for example on the Earth or on a spacecraft [9, 62]. This observation will certainly happen on different null hypersurfaces, involving, therefore, light emitted at different moments along the emitter’s world line. The results of observations will therefore certainly contain a contribution from the motion of the source of light.

The comparison of the position at different moments may nevertheless proceed as before: knowing the details of the motion of the observer in the local inertial frame (for example the knowing the Earth’s orbit in the barycentric reference system) we may subtract any effects of the Bradley aberration, Rømer delays and possibly the local effects of light bending due to the Sun or massive planets and consider the “pure” parallax with respect to a local inertial frame [62, 63]. The resulting effects of parallax will be referred to as the *single worldline parallax*. Again it is a special case of the parallax with respect to the local inertial frame.

This is obviously a different type of measurement than the classic parallax. Nevertheless, it is possible to infer about the classic parallax from a single worldline measurement under certain additional assumptions and we will discuss the exact relation between these two types of measurement later in this Section.

d. Position drift. For a single observer with a given worldline there is another natural way of identifying directions on the observer’s spheres of directions at different times, one that does not involve an artificial boosting of the sky in order to subtract the aberration effects. Instead of a parallel-transported basis, we may simply consider the Fermi-Walker transport of vectors in the observer’s the sphere of directions $\text{Dir}(u_{\mathcal{O}})$ along the worldline as our “fixed directions on the sky” [1, 64]. The Fermi-Walker transport preserves the metric structure of the sphere of directions, just like the parallel transport does in a fixed orthonormal frame [65, 66]. Fixed directions defined this way by an observer correspond physically to directions given by a system of gyroscopes carried by the observer during his or her non-inertial motion [66, 67]. The variation of the apparent position of an object defined this way will be referred to as the *position drift* [1]. It obviously depends on the details of the motion of both the observer and the emitter, but unlike the single worldline parallax, it also depends on the observer’s 4-acceleration [1, 61, 64].

e. Relative parallax. The notions of parallax defined above use the properties of the local inertial or non-inertial frame and the spacetime’s local geometry in order to define the reference directions for measuring the variations of the apparent positions. Therefore, they require not only measuring the apparent position of an object at two distinct points but also comparing it to fixed directions defined in the local inertial or non-inertial frame. This may be rather complicated to do in practice and, therefore, in astronomy, where one considers instead just the relative changes of the apparent positions of images of two or more objects with respect to each other, without any reference to the notion of fixed directions. The observables, in this case, are given by the variations of angles between the images. Measurements of this kind are by far the simplest to perform since they only require a single telescope, without a system gyroscopes or other devices defining fixed directions across the observers’ region $N_{\mathcal{O}}$. This type of parallax may be called the *relative parallax*. Among all possible definitions, this one is the closest to the way parallax is normally measured in astronomy: the positions of sources and their time variations are expressed with respect to the Solar System’s barycentric reference frame, in which the non-rotating axes are determined by the apparent positions of a number of selected extragalactic radio or optical sources [62] rather than any local physical phenomena. Therefore, in the end, the parallax measurements by Gaia or Hipparcos rely only on the relative positions on the celestial sphere of many sources. In the context of relativistic astrophysics and cosmology, relative parallax has been introduced recently (not under that name) by Räsänen, in his theoretical work about the parallax of distant quasars measured by the Gaia mission [9].

Note that it is very easy to obtain the relative parallax from the parallax with respect to the local inertial frame or the position drift. Since both methods of identifying directions on the sky preserve the metric structure of the sphere of directions the rate of change of the angle between any two sources can be easily expressed via their momentary positions and their variations using simple trigonometry: let \vec{r}_1 and \vec{r}_2 be normalized, spatial 3-vectors pointing towards the apparent positions of two sources in $\text{Dir}(u_{\mathcal{O}})$. Then the angle α between their images is given by their scalar product via $\cos \alpha = \vec{r}_1 \cdot \vec{r}_2$. Therefore a small variation of α is expressible in terms of the variations of \vec{r}_1 and \vec{r}_2 , defined by the parallax with respect to a local inertial frame or by the position drift: $-\delta\alpha \sin \alpha = \delta\vec{r}_1 \cdot \vec{r}_2 + \vec{r}_1 \cdot \delta\vec{r}_2$. However, going in the opposite direction is not possible: the change of relative angles between images of an arbitrarily large number of sources is not enough to recover the values of the image drifts with respect to the local inertial frame. This is because passing to the relative parallax involves the loss of information about the rigid rotation of the whole celestial sphere with respect to the fixed directions in the local inertial frame. This would manifest itself by a secular, average rotation the images of the faraway sources like quasars when compared to the Solar System’s local non-rotating frame. In the standard cosmological model it is most likely very small due to negligible vorticity on large scales, but note that it is by definition unobservable using only the relative parallax measurements. The (unphysical) Goedel metric provides a simple and elegant example of this effect, as discussed in [9].

In the considerations above we have assumed the region $N_{\mathcal{O}}$ to be strictly flat. It is possible however to extend the applicability of all definitions to a more physically relevant case when the metric is flat plus small, localized Newtonian or post-Newtonian perturbations due to the presence of local masses. One may then use the fictitious, background

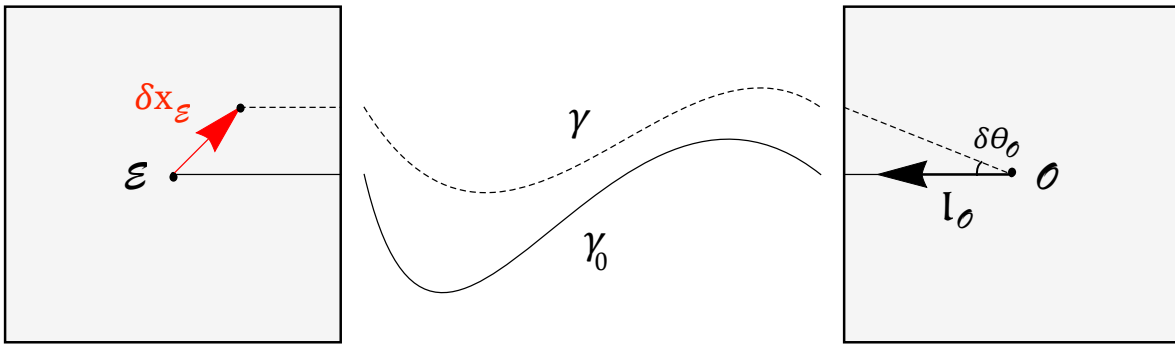


FIG. 5: A view of the projection of the null hypersurfaces containing \mathcal{E} and \mathcal{O} . The angle at which the observer sees another point in an extended emitter's cross-section is determined by $\delta x_{\mathcal{E}}$, the Jacobi operator and the observer's 4-velocity.

flat metric to define fixed directions in $N_{\mathcal{O}}$ along the same lines. This would correspond to using the inertial frame defined far away from the local masses as a reference for the parallax. Note, however, that in order to obtain the “pure” parallax from the observations of a source one would need to subtract all the local gravitational effects such as the light bending due to the presence of the local masses [62].

B. Jacobi operator and the image distortions

We now move on to the discussion of the role of the optical operators. We begin by explaining the physical meaning of the Jacobi operator. The material presented here consists standard results (see for example the review paper [54]), but will serve as an introduction to the results of the next sections in which explain the physical significance of the other optical operator m .

Consider a single observer performing an observation at \mathcal{O} ($\delta x_{\mathcal{O}} = 0$) and a body of finite size passing through \mathcal{E} , see Figure 5. Obviously $\delta x_{\mathcal{O}}^{\mu} l_{\mathcal{E}\mu} = 0$ and from the time lapse formula (49) we see that the observer registers light from various points of the body emitted at the moment given by the condition $\delta x_{\mathcal{E}}^{\mu} l_{\mathcal{E}\mu} = 0$. Thus when we pass to the quotient space we have $[\delta x_{\mathcal{E}}] \in \mathcal{P}_{\mathcal{E}}$. Therefore equation (62) takes the form of

$$\Delta l_{\mathcal{O}}^A = \mathcal{D}^{-1A}{}_B \delta x_{\mathcal{E}}^B. \quad (63)$$

The equation above yields a linear relation between the deviation of the direction of light propagation at \mathcal{O} and the displacement of a point on the emitting body's cross-section from \mathcal{E} . This latter can be related to the physical distance from \mathcal{E} on the emitter's screen space as measured in the body's own frame, regardless of the emitter's momentary motion: we noted that the distances between points on the cross-section, as measured in the body's reference frame, are given by the distances evaluated in $\mathcal{P}_{\mathcal{E}}$ using its internal metric, irrespective of the body's 4-velocity $u_{\mathcal{E}}^{\mu}$.

a. Magnification matrix. We would like to relate the distances on the emitter's screen space to the angles measured at the observer's sky. This requires taking into account the Bradley aberration effects. According to (38) and (44) we have

$$\delta \theta_{\mathcal{O}}^A = \frac{1}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} \mathcal{D}^{-1A}{}_B \delta x_{\mathcal{E}}^B \quad (64)$$

with $u_{\mathcal{O}}^{\sigma}$ being the observer's 4-velocity. The combination

$$M^A{}_B = \frac{1}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} \mathcal{D}^{-1A}{}_B \quad (65)$$

is called the *magnification matrix* in the gravitational lensing theory (the definition here relates the perpendicular displacements on the source plane directly to angles on the sky rather than the position on the image plane, but these two are related by a simple rescaling). It has the dimension of $1/L$, where L denotes length. It depends on the observer's motion via the aberration effect: observers with different 4-velocities observe the celestial sphere transformed by a conformal mapping with respect to each other. This transformation makes certain parts of the sky

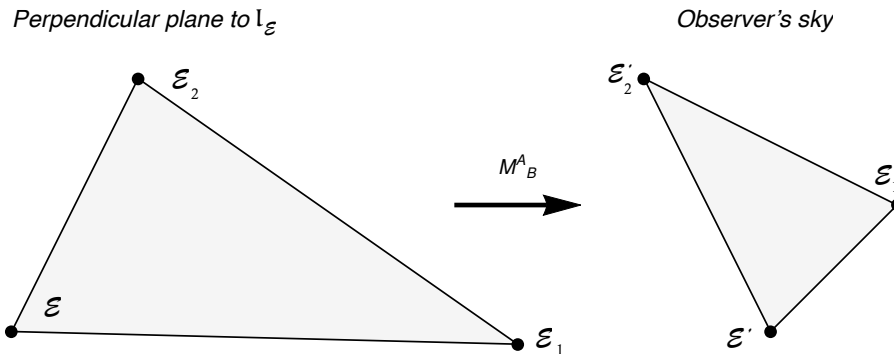


FIG. 6: A point-like luminous object at \mathcal{E} and two other point-like luminous objects nearby, forming a triangle on the cross-section plane in the direction of the propagation of light. The angular diameter distance between \mathcal{O} and \mathcal{E} is obtained as the square root of the ratio between the area of the triangle on the cross-section plane and the stereographic angle occupied by the triangle with vertices at the apparent positions of the 3 objects at the observer's sky.

appear larger or smaller depending on the observer's 4-velocity but without any shape distortions of small objects. In (65) this dependence is encoded in the $(l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma})^{-1}$ factor in front of the Jacobi matrix. The magnification matrix for weak lensing in a perturbed FLRW Universe, together with corrections due to proper motions, has been calculated in [68].

b. Angular diameter distance. The magnification matrix gives explicitly the relation between perpendicular distances on the emitter's side and in his/her frame and the angles on the observer's sky. Its determinant measures the relation of a light emitting body's cross-sectional area $A_{\mathcal{E}}$ to the stereographic angle $\Omega_{\mathcal{E}}$ it takes up on the sky. The ratio between these areas is used to define the *angular diameter distance* to the body, also known as the *area distance*:

$$D_{ang} = \sqrt{\frac{A_{\mathcal{E}}}{\Omega_{\mathcal{E}}}}. \quad (66)$$

This quantity is related to the magnification matrix via

$$D_{ang} = |\det M^{\mathbf{A}}_{\mathbf{B}}|^{-1/2}. \quad (67)$$

Equivalently, we can express the angular diameter distance directly via the Jacobi map:

$$D_{ang} = (l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}) |\det \mathcal{D}^{\mathbf{A}}_{\mathbf{B}}|^{1/2}. \quad (68)$$

Just like $M^{\mathbf{A}}_{\mathbf{B}}$, D_{ang} depends on the observer's motion via the aberration effects, but not on the emitter's motions. It is related to the luminosity distance D_{lum} via the Etherington's formula

$$D_{lum} = D_{ang}(1+z)^2,$$

see [54, 69], the latter republished as [70]. The off-diagonal part of $M^{\mathbf{A}}_{\mathbf{B}}$ contains the information about the image distortions [6].

A simple physical interpretation of the angular diameter distance can be explained as in Figure 6. Assume we observe from the point \mathcal{O} the light emitted by three pointlike objects at the events \mathcal{E} , \mathcal{E}_1 and \mathcal{E}_2 , all lying on the same null hypersurface $l_{\mathcal{E}\mu} \delta x_{\mathcal{E}}^{\mu} = 0$. Their projections from a triangle of area $A_{\mathcal{E}}$ on the cross-sectional plane perpendicular to $l_{\mathcal{E}}$ (this value is independent of the choice of the observer in \mathcal{E} due to the Sachs shadow theorem). On the other hand, their images form a triangle on the observer's sky, covering the stereographic angle of $\Omega_{\mathcal{E}}$. In a spacetime with strong lensing the image may be subject to a strong deformation, including a rescaling, shear, and rotation, but the angular diameter distance is defined simply by the ratio of $A_{\mathcal{E}}$ and $\Omega_{\mathcal{E}}$ via (66).

C. Emitter-observer asymmetry operator and the classic parallax

Consider now the opposite situation: we have a single point source emitting light while passing through the point \mathcal{E} and a number of observers o_1, o_2, \dots , displaced with respect to each other and such that o_1 passes through \mathcal{O} . Each

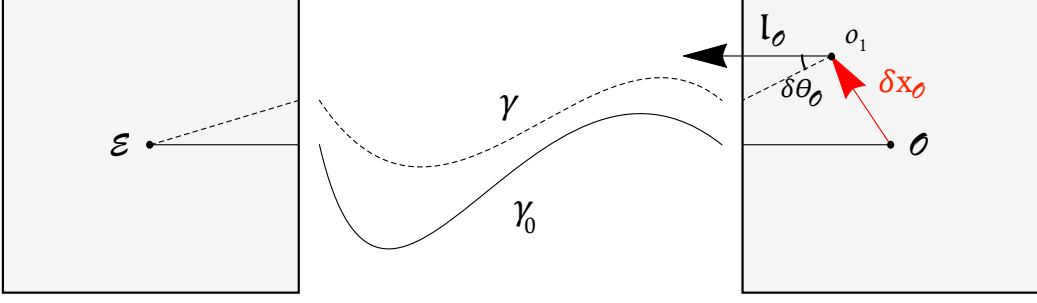


FIG. 7: A view of the projection of the null hypersurfaces containing \mathcal{E} and \mathcal{O} . The angle at which the displaced observer sees the point source at \mathcal{E} as compared to the one at \mathcal{O} , i.e. the classical parallax, is determined by the Jacobi operator, the emitter-observer asymmetry operator, and the observer's 4-velocity.

observer o_i will register the emitter's apparent position e_i on their celestial spheres. We assume that the observers register the emitter's position on the sky using light emitted only at point \mathcal{E} , i.e. we have $\delta x_{\mathcal{E}}^{\mu} = 0$. From (49) this means that they all register the position of the source *at the moment they cross the null hypersurface* $\delta x_{\mathcal{O}}^{\mu} l_{\mathcal{O}\mu} = 0$. For simplicity we also assume that they are comoving: their 4-velocities at the moment of measurement are exactly the same, i.e. in locally flat coordinates they are all equal to a fixed $u_{\mathcal{O}}^{\mu}$. This way we do not need to consider the aberration effects when comparing the results of their measurements.

Since $\delta x_{\mathcal{O}}^{\mu}$ must be orthogonal to $l_{\mathcal{O}}^{\mu}$ for all observers at the moment of observation we have $[\delta x_{\mathcal{O}}] \in \mathcal{P}_{\mathcal{O}}$ for the equivalence class of their displacement vectors. From (62) we get the following relation between the observers' displacement vectors and the direction deviation:

$$\mathcal{D}^A_B \Delta l_{\mathcal{O}}^B = -\delta \hat{x}_{\mathcal{O}}^A - m_{\perp}^A_B \delta \hat{x}_{\mathcal{O}}^B = -(\delta^A_B + m_{\perp}^A_B) \delta \hat{x}_{\mathcal{O}}^B. \quad (69)$$

We see that the equation (69) differs from (63) by the sign at the position displacement term and the presence of a term involving the perpendicular part m_{\perp} of the emitter-observer asymmetry operator m . In order to elucidate its physical meaning, we will now compare the situation when m_{\perp} vanishes (for example because the spacetime is flat) and when it does not.

In a flat space, where m_{\perp} vanishes between \mathcal{E} and \mathcal{O} , (69) is formally identical to (63) with $\delta x_{\mathcal{E}}$ replaced by $-\delta \hat{x}_{\mathcal{O}}$. In other words, a perpendicular displacement of the observer $[\delta x_{\mathcal{O}}]$ is precisely equivalent to a displacement of the emitter in the opposite direction, i.e. $[\delta x_{\mathcal{E}}] = -[\delta \hat{x}_{\mathcal{O}}]$. The notion of opposite direction, used here for displacement vectors defined at two different endpoints of the fiducial null geodesic γ_0 , is defined by the parallel transport along it.

These conclusions follow already from (62), where, in the absence of m , any value of the LHS can be attributed to either $\delta \hat{x}_{\mathcal{O}}$ or $-\delta \hat{x}_{\mathcal{E}}$ at will. This is easy to understand geometrically if we realize that in the flat space the perpendicular displacement vectors and the null geodesics form in this case a thin but long parallelogram, see Figure 8. The two null geodesics corresponding to the displacements on two ends of γ_0 are initially parallel, so the angles between them and the direction of the fiducial geodesic must be the same.

However, if the curvature is present along γ_0 then the property of exact equivalence between the position displacements at the two ends of the null geodesic may be broken by a term proportional to $m_{\perp}^A_B$. From (62) it is easy to see that the value of LHS cannot simply be shifted from $\delta x_{\mathcal{E}}$ to $-\delta \hat{x}_{\mathcal{O}}$ since the operator m_{\perp} spoils the symmetry. This observation justifies the name emitter-observer asymmetry operator, proposed in Section III C.

One possible geometric interpretation of the linear operator m_{\perp} is thus that it is the obstruction for the existence of the thin and long parallelogram discussed above, made of the two initially parallel, long null geodesics, and two vectors $\delta x_{\mathcal{O}}$ and $-\delta \hat{x}_{\mathcal{O}}$. If m_{\perp} happens to vanish between \mathcal{O} and \mathcal{E} then the parallelogram exists. It follows then that the problem of parallax is effectively equivalent to the problem of image distortion discussed in the previous section despite all gravitational lensing which may happen along the way, see again Figure 8. If m_{\perp} does not vanish, then the parallelogram does not exist in general and the equivalence of opposite displacements on both sides of γ_0 is broken. In this case (69) yields

$$\Delta l_{\mathcal{O}}^A = -\mathcal{D}^{-1A}_C (\delta^C_B + m_{\perp}^C_B) \delta \hat{x}_{\mathcal{O}}^B.$$

Recall now that we have assumed that all observers are comoving with the observer \mathcal{O} , i.e. their 4-velocities are equal

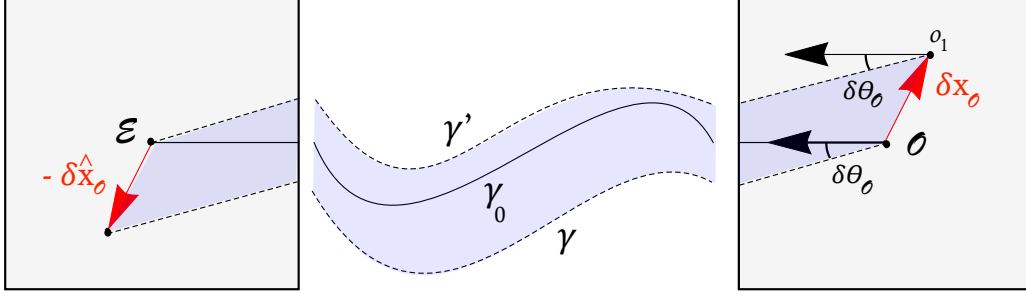


FIG. 8: If the perpendicular part of the emitter-observer asymmetry operator m_{\perp} vanishes, it is possible to draw a thin and long parallelogram made of two, initially parallel null geodesics γ and γ' , the displacement vector $\delta x_{\mathcal{O}}$ and its parallel transport $-\delta\hat{x}_{\mathcal{O}}$. γ_0 plays the role of its diagonal. In this case determining the parallax angle for an observer displaced by $\delta x_{\mathcal{O}}$ is equivalent to the problem of the angular size of an object extending from \mathcal{E} to $-\delta\hat{x}_{\mathcal{O}}$.

to $u_{\mathcal{O}}^{\mu}$. Then the position on the celestial sphere of what a displaced observer will measure is

$$\delta\theta_{\mathcal{O}}^{\mathbf{A}} = -\frac{1}{u_{\mathcal{O}}^{\sigma} l_{\sigma}} \mathcal{D}^{-1\mathbf{A}}_{\mathbf{C}} (\delta^{\mathbf{C}}_{\mathbf{B}} + m_{\perp}^{\mathbf{C}}_{\mathbf{B}}) \delta\hat{x}_{\mathcal{O}}^{\mathbf{B}}. \quad (70)$$

This is a linear relation between the perpendicular displacement of the observer and the change of apparent position as measured for a source at a single instant of the source's time, defined with respect to the local geometry, i.e. the classic parallax in the terminology of Section IV A.

Note that if the 4-velocities of the observers o_1, o_2, \dots are not exactly equal we need to take into account the aberration effects when comparing the registered positions on the sky between the observer. These effects will add a 4-velocity-dependent term on top of the linear term from (70). Calculating this term is a standard special relativity problem which we leave to the reader. Here we prefer to assume comoving observers in order to isolate the dependence of the position on the sky on the observer's displacement.

a. Parallax matrix. In analogy with the magnification matrix from (65) we can introduce the observer-dependent *parallax matrix*

$$\Pi^{\mathbf{A}}_{\mathbf{B}} = \frac{1}{u_{\mathcal{O}}^{\sigma} l_{\sigma}} \mathcal{D}^{-1\mathbf{A}}_{\mathbf{C}} (\delta^{\mathbf{C}}_{\mathbf{B}} + m_{\perp}^{\mathbf{C}}_{\mathbf{B}}). \quad (71)$$

The parallax matrix relates perpendicular distances on the observer's side and two-dimensional angles measuring the observed position on the sky in comparison with the position observed by o_1 at \mathcal{O} . Namely, the equation (70) for the classic parallax takes the form of

$$\delta\theta_{\mathcal{O}}^{\mathbf{A}} = -\Pi^{\mathbf{A}}_{\mathbf{B}} \delta\hat{x}_{\mathcal{O}}^{\mathbf{B}}.$$

$\Pi^{\mathbf{A}}_{\mathbf{B}}$ is independent of $u_{\mathcal{E}}$ because its definition relies on the observations from various points of view of the light emitted by the source in a single moment. Therefore what matters for the observation is only the exact position of the emitter at the moment of observation \mathcal{E} , while the rest of its worldline, which for short times can be approximated by the first two terms in the Taylor expansion, given by the momentary 4-velocity $u_{\mathcal{E}}^{\mu}$ and 4-acceleration $w_{\mathcal{E}}^{\mu}$, is irrelevant. On the other hand, the parallax matrix depends on the spacetime geometry and on the observer's 4-velocity. The dependence on the geometry is via the curvature tensor along γ_0 because in (71) $\Pi^{\mathbf{A}}_{\mathbf{B}}$ is expressed as a function of the optical operators, themselves functionals of the Riemann tensor. The dependence on $u_{\mathcal{O}}^{\mu}$ enters only via the aberration effects, just like in $M^{\mathbf{A}}_{\mathbf{B}}$. Thus we have

$$\Pi^{\mathbf{A}}_{\mathbf{B}} \equiv \Pi^{\mathbf{A}}_{\mathbf{B}} \left(R^{\mu}_{\nu\alpha\beta}, u_{\mathcal{O}}^{\mu} \right).$$

Just like $M^{\mathbf{A}}_{\mathbf{B}}$, the parallax matrix has the dimension of $1/L$.

b. Parallax distance. In astronomy determining the parallax is one of the standard methods of measuring the distances to objects up to few kiloparsecs [71, 72]. The method relies again on the flat space formula for the parallax matrix: in a flat space we have $\Pi^{\mathbf{A}}_{\mathbf{B}} = d^{-1} \delta^{\mathbf{A}}_{\mathbf{B}}$, where d is the spatial distance between \mathcal{O} and \mathcal{E} measured in the

observer's frame. If we include the relativistic effects of light bending $\Pi^{\mathbf{A}}_{\mathbf{B}}$ is not guaranteed to be proportional to the unit matrix any more. This leads to the dependence of the parallax effect on the orientation of the baseline $\delta x_{\mathcal{E}}^{\mu}$ [9]. It is therefore reasonable to try to extract an angle-averaged quantity out of $\Pi^{\mathbf{A}}_{\mathbf{B}}$. In [9, 60] the following definition has been proposed: the parallax distance is defined to be proportional to the inverse of the expansion of the congruence of null geodesics originating from \mathcal{E} at \mathcal{O} . In the language of this paper it is simply inverse of arithmetic average of elements of the diagonal, see [1]:

$$\tilde{D}_{par} = \frac{2}{\text{Tr}(\Pi^{\mathbf{A}}_{\mathbf{B}})}.$$

This is a possible generalization, but in this paper we would like to put forward another approach, analogous to the one used in the standard definition of the angular diameter distance D_{ang} . We define the parallax distance using the determinant of the parallax matrix

$$D_{par} = |\det \Pi^{\mathbf{A}}_{\mathbf{B}}|^{-1/2},$$

or equivalently

$$D_{par} = (l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}) |\det \mathcal{D}^{\mathbf{A}}_{\mathbf{B}}|^{1/2} |\det (\delta^{\mathbf{A}}_{\mathbf{B}} + m_{\perp}^{\mathbf{A}}_{\mathbf{B}})|^{-1/2}. \quad (72)$$

The reader may check that in a flat space this definition yields again the right answer, but in a non-flat space it averages the results over directions in a different way than the standard one. Let π_1 and π_2 be the two roots of the characteristic equation of $\Pi^{\mathbf{A}}_{\mathbf{B}}$, possibly real or possibly complex and conjugate to each other. Then $\tilde{D}_{par}^{-1} = \frac{1}{2} (\pi_1 + \pi_2)$ while $D_{par}^{-1} = \sqrt{|\pi_1 \pi_2|}$. Thus the latter definition is equivalent to the inverse of the *geometric* average of the moduli of the roots, while the former uses the inverse of their *arithmetic* average. As a consequence we see that both values coincide if no shear is present between \mathcal{O} and \mathcal{E} : $\Pi^{\mathbf{A}}_{\mathbf{B}}$ is proportional to the unit matrix in that case and both roots are equal.

We will now explain the physical interpretation of the new definition proposed above, highlighting this way the analogy to the notion of the angular diameter distance. Consider an observer o_1 at \mathcal{O} and two additional observers o_2 and o_3 , displaced with respect to \mathcal{O} in two different directions (not necessary perpendicular to each other) and comoving with o_1 . They all measure the apparent position of the emitter when it passes through \mathcal{E} and combine their results of observation on a single celestial sphere S^2 by identifying points on their celestial spheres corresponding to parallel directions in the sense of locally flat coordinates in $N_{\mathcal{O}}$. The result will be a triple of close points on S^2 corresponding the three observations and forming a solid triangle T_1 . On the other hand, we may also consider a physical triangle T_2 these observers form when we project their displacement vectors to the screen space perpendicular to $l_{\mathcal{O}}$, see Figure 9. Let $A_{\mathcal{O}}$ denote the area of the triangle T_2 (again it is independent of the observer's o_1 4-velocity due to the Sachs theorem) and let $\Omega_{\mathcal{O}}$ denote the solid angle taken up by T_1 . Then we have

$$D_{par} = \sqrt{\frac{A_{\mathcal{O}}}{\Omega_{\mathcal{O}}}}.$$

The analogy with the definition of the angular distance is now evident: the definition of D_{par} is equivalent with D_{ang} , but with the displacements considered on the observer's side of γ_0 instead of the emitter's side. As we will see in Section V, the compatibility of both definitions opens up the possibility of defining a new observable measuring directly the spacetime curvature.

Summarizing, the curvature along the line of sight produces two types of effect: the gravitational lensing, which modifies the Jacobi map, resulting in the (de)magnification and distortion of images seen by observers, and introducing asymmetry between the displacements of the two end of the null geodesic. Both effects are independent and both affect the classic parallax.

D. Position drift formula

Before we consider more realistic models of measurements of the parallax we will rederive the general position drift formula for the momentary rate of change of the apparent position of a source at \mathcal{E} as observed by an observer in \mathcal{O} in any spacetime. The formula relates the position drift (or proper motion) to the optical operators between \mathcal{O} and \mathcal{E} and to the quantities describing the momentary motions both the emitter and the observer, namely the emitter's 4-velocity $u_{\mathcal{E}}^{\mu}$, the observer's 4-velocity $u_{\mathcal{O}}^{\mu}$ and the observer's 4-acceleration $w_{\mathcal{O}}^{\mu}$. It has already been presented in [1], but the derivation there is rather involved. The derivation using the formalism developed here on the other hand is conceptually simpler and computationally rather straightforward.

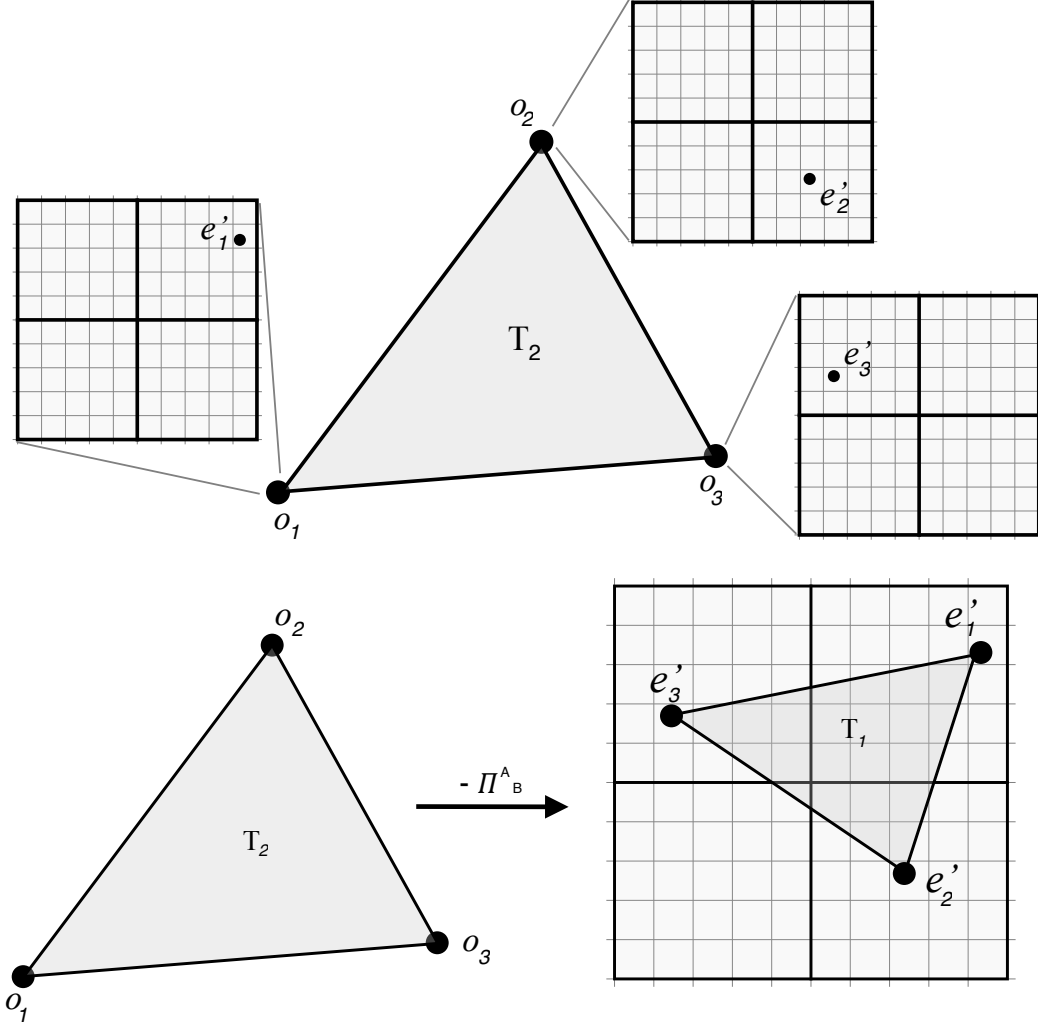


FIG. 9: A single, point-like luminous source at \mathcal{E} observed by an observer o_1 at \mathcal{O} as well as two comoving observers o_2 and o_3 displaced with respect to o_1 , performing their measurements on the same null hypersurface. The projection of their positions to the screen space perpendicular to $l_{\mathcal{O}}$ yields a triangle T_2 of area $A_{\mathcal{O}}$. On the other hand, superimposing the registered positions of the sources by the three observers yields a triangle T_1 on the “combined” celestial sphere, whose solid angle area is denoted by $\Omega_{\mathcal{O}}$.

Let $\tau_{\mathcal{E}}$ and $\tau_{\mathcal{O}}$ denote the proper time as measured by the emitter and observer respectively. Additionally, let $\tau_{\mathcal{E}} = \tau_{\mathcal{O}} = 0$ when both objects cross \mathcal{E} and \mathcal{O} respectively, i.e. at the moment of emission and observation. After a short while we have $\delta x_{\mathcal{O}}^{\mu} = u_{\mathcal{O}}^{\mu} \delta \tau_{\mathcal{O}}$ and $\delta x_{\mathcal{E}}^{\mu} = u_{\mathcal{E}}^{\mu} \delta \tau_{\mathcal{E}}$ in the leading, linear order in each of the proper times. The time lapse formula (49) yields

$$u_{\mathcal{E}}^{\mu} l_{\mathcal{E}\mu} \delta \tau_{\mathcal{E}} = u_{\mathcal{O}}^{\mu} l_{\mathcal{O}\mu} \delta \tau_{\mathcal{O}}.$$

This relation can be turned into a formula for the emitter’s time lapse as registered by the observer and compared with his/her proper time lapse:

$$\frac{\delta \tau_{\mathcal{E}}}{\delta \tau_{\mathcal{O}}} = \frac{u_{\mathcal{O}}^{\sigma} l_{\mathcal{O}\sigma}}{u_{\mathcal{E}}^{\rho} l_{\mathcal{E}\rho}}.$$

The ratio on the right hand side is obviously related to the redshift defined as the relative difference between the

photon energy as measured by \mathcal{E} and \mathcal{O} :

$$z = \frac{u_{\mathcal{E}}^{\rho} l_{\mathcal{E}\rho}}{u_{\mathcal{O}}^{\sigma} l_{\mathcal{O}\sigma}} - 1.$$

Therefore we see that (49) is equivalent to the following relation between the redshift z defined by photon energy change and the rate of the emitter's time lapse the the observer's time lapse:

$$\delta\tau_{\mathcal{E}} = \frac{1}{1+z} \delta\tau_{\mathcal{O}}$$

(derived earlier in [1, 51, 73]).

We can now evaluate the position drift. After substituting the position deviation and simple manipulations the direction variation formula (62) yields

$$\frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{\delta\tau_{\mathcal{O}}} = \mathcal{D}^{-1\mathbf{A}}{}_{\mathbf{B}} \left(\left(\frac{1}{1+z} u_{\mathcal{E}} - \hat{u}_{\mathcal{O}} \right)^{\mathbf{B}} - m^{\mathbf{B}}{}_{\mathbf{i}} u_{\mathcal{O}}^{\mathbf{i}} \right). \quad (73)$$

The ratio on the left hand side is simply the covariant derivative of the null tangent vector along the observer's worldline evaluated at \mathcal{O} and pulled back to $\mathcal{P}_{\mathcal{O}}$:

$$\frac{\Delta l_{\mathcal{O}}^{\mathbf{A}}}{\delta\tau_{\mathcal{O}}} = (u_{\mathcal{O}}^{\nu} \nabla_{\nu} l_{\mathcal{O}}) \Big|_{\mathcal{O}}^{\mathbf{A}}$$

This can be related to the parallax with respect to the local inertial frame by simple rescaling according to (38) and (44). For a non-geodesic observer, we may also obtain the expression for the position drift. By definition we need to evaluate the Fermi-Walker derivative of r^{μ} [1, 64], given by

$$\delta_{\mathcal{O}} r^{\mu} = (u_{\mathcal{O}}^{\nu} \nabla_{\nu} r^{\mu}) + (-u_{\mathcal{O}}^{\mu} w_{\mathcal{O}\nu} + w_{\mathcal{O}}^{\mu} u_{\mathcal{O}\nu}) r^{\nu}, \quad (74)$$

with the transverse components of the term $u_{\mathcal{O}}^{\nu} \nabla_{\nu} r^{\mu}$ given by (44) and (73). $\delta_{\mathcal{O}} r^{\mathbf{A}}$ corresponds to the position drift measured with respect to inertially dragged fixed directions. Combining (74) and (73) yields

$$\delta_{\mathcal{O}} r^{\mathbf{A}} = \frac{1}{l_{\mathcal{O}\sigma} u_{\mathcal{O}}^{\sigma}} \mathcal{D}^{-1\mathbf{A}}{}_{\mathbf{B}} \left(\left(\frac{1}{1+z} u_{\mathcal{E}} - \hat{u}_{\mathcal{O}} \right)^{\mathbf{B}} - m^{\mathbf{B}}{}_{\mathbf{i}} u_{\mathcal{O}}^{\mathbf{i}} \right) + w_{\mathcal{O}}^{\mathbf{A}} \quad (75)$$

for the only 2 non-vanishing components of $\delta_{\mathcal{O}} r^{\mu}$ in a SNF. The last term is the perpendicular component of the observer's 4-acceleration. It corresponds to the special relativistic effect of the position drift due to the drift of the aberration [1, 9, 61]. Its influence on the drift of the positions of sources at cosmological distances has been recently discussed in [61]. For a longer discussion of the position drift formula and its physical and astrophysical consequences see [1], here we will just briefly look at the role of the emitter-observer asymmetry operator.

First, consider the situation in which $u_{\mathcal{E}}^{\mu}$ and the parallel-transported $\hat{u}_{\mathcal{O}}^{\mu}$ differ only by a component along the line of sight. In this case, it is easy to see that the perpendicular component of the 4-velocity difference $\frac{1}{1+z} u_{\mathcal{E}} - \hat{u}_{\mathcal{O}}$ vanishes, so the first term in (75) does not contribute to the drift. In the absence of $m^{\mathbf{A}}{}_{\mathbf{i}}$, this means no drift seen by the observer. However, if the emitter-observer asymmetry operator is not 0 the observer can perceive a "curvature-induced" position drift even when both 4-velocities $u_{\mathcal{E}}$, $\hat{u}_{\mathcal{O}}$ and the null vector $l_{\mathcal{E}}$ lie on a single 2-plane and the first term in (75) vanishes. This type of drift is proportional to the timelike component of $m^{\mathbf{A}}{}_{\mathbf{i}}$, i.e. $m^{\mathbf{A}}{}_{\mathbf{i}} u_{\mathcal{O}}^{\mathbf{i}}$. It is independent of the velocity of the emitter's radial motion with respect to the observer.

Secondly, we note that in (75) operator m appears once again in a term which breaks the symmetry between the emitter and observer. This is not so easy to see at first inspection, because, unlike (62), the formula (75) even without the m term does not seem antisymmetric with respect to the exchange of $u_{\mathcal{E}}$ and $u_{\mathcal{O}}$ at first glance because of the $\frac{1}{1+z}$ factor in front of $u_{\mathcal{E}}$. This is because in the derivation above we have made a choice which ties the formula to the observer's frame: namely, we relate the rate of change of the apparent position to the lapse of the *observer's* proper time. This choice introduces an asymmetry between $u_{\mathcal{O}}$ and $u_{\mathcal{E}}$ in (75) at the level of pure special relativity, even without any GR effects present. Nevertheless, the symmetry breaking role of m can again be seen in the following example.

Consider an emitter at \mathcal{E} and an observer at \mathcal{O} for whom not only the perpendicular components of $\frac{1}{1+z} u_{\mathcal{E}}^{\mu} - \hat{u}_{\mathcal{O}}^{\mu}$ vanish, but actually both 4-velocities coincide. This means that both objects are at rest with respect to each other with the comparison made using the parallel transport along γ_0 . Assume as before that the observer is geodesic

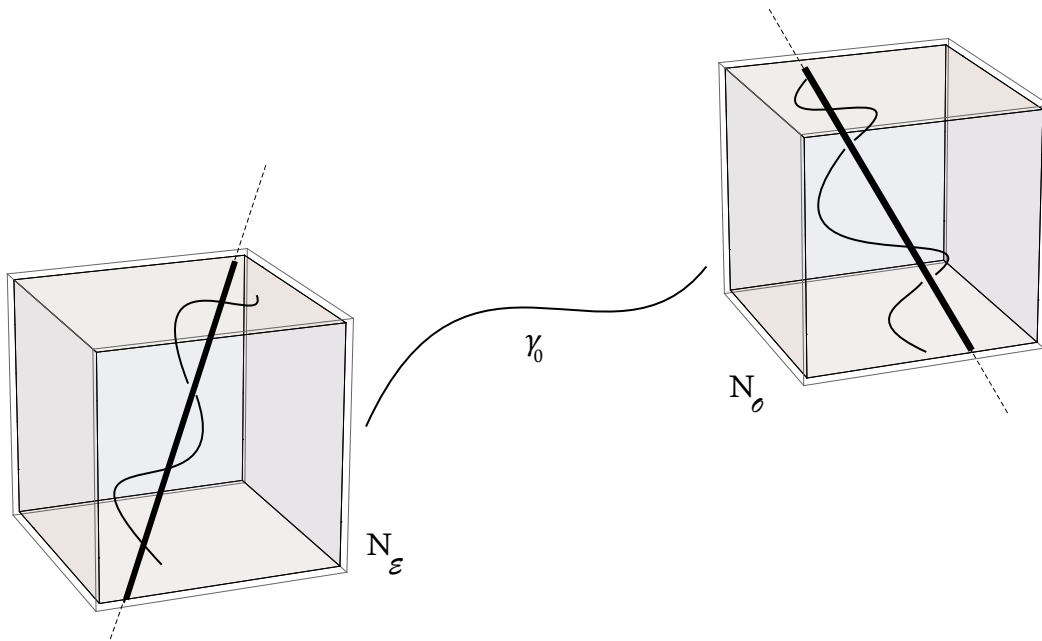


FIG. 10: The observer and the emitter in gravitationally bound systems, undergoing short-time orbital motions around their free-falling barycenters.

($w_{\mathcal{O}}^{\mu} = 0$). It follows then that $z = 0$ and the first term in (75) vanishes. Just like in the previous example, in the absence of m the curvature-induced drift vanishes, but additionally the whole situation is symmetric with respect to boosting the emitter and observer in the following sense: consider another emitter $\tilde{u}_{\mathcal{E}}^{\mu}$ passing through \mathcal{E} , boosted with respect to $u_{\mathcal{E}}^{\mu}$ in a direction orthogonal to the line of sight (orthogonality determined in the $u_{\mathcal{E}}$ reference frame). This new emitter will exhibit drift according to the observer, although in general not in the direction related to the parallel transported direction of its motion. On the other hand, the reader may check that exactly the *same* position drift $\delta_{\mathcal{O}} r^A$ can be induced by considering a free-falling, moving *observer*, boosted with respect to $u_{\mathcal{O}}^{\mu}$ with the same velocity but in the opposite direction (again in the sense of parallel transport) and observing the unboosted emitter. Just like the displacement equivalence on both ends of γ_0 noted in Section IV C, this boost equivalence property is lost whenever $m_{\perp} \neq 0$.

We would like to remark that even though we have used the flat light cones approximation and the parallel light rays approximation, the formula derived here is valid without any restrictions regarding the distance between \mathcal{O} and \mathcal{E} or the bilocal geodesic operators. This is because the position drift is by definition the first derivative of the position on the sky and the derivative is always obtained by linearization of all relations involved, including the null condition (41) and the position on the sky formula (39).

E. Single worldline parallax in a general situation

Finally, we move on to discussing a more realistic model of parallax observations. Assume that both the observer and the emitter are located in gravitationally bound systems whose barycenters undergo a geodesic motion (free fall) with a good approximation. The gravitational fields binding the systems are assumed to be so weak that the light bending they induce is negligible – if it is not then they may be introduced later as small, perturbative corrections, for example using the PPN formalism [62]. We assume here that the motion of the observer and the emitter in their respective local inertial frames can be well approximated by short-time, non-relativistic orbital motions around the barycenters, superimposed on top of a secular motion of the barycenters with constant 4-velocities $U_{\mathcal{O}}$ and $U_{\mathcal{E}}$ respectively, see Figure 10. The characteristic timescale of this short-time motion is assumed to be smaller than the size of $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$, i.e. L . Assume that points \mathcal{O} and \mathcal{E} lie on the worldlines of the respective barycenters and let $t_{\mathcal{E}}$ and $t_{\mathcal{O}}$ denote the proper time in the appropriate barycentric reference system (the barycentric coordinate times in the astronomical terminology), defined such that $t_{\mathcal{E}} = 0$ at \mathcal{E} and $t_{\mathcal{O}} = 0$ at \mathcal{O} . Then the momentary position of the

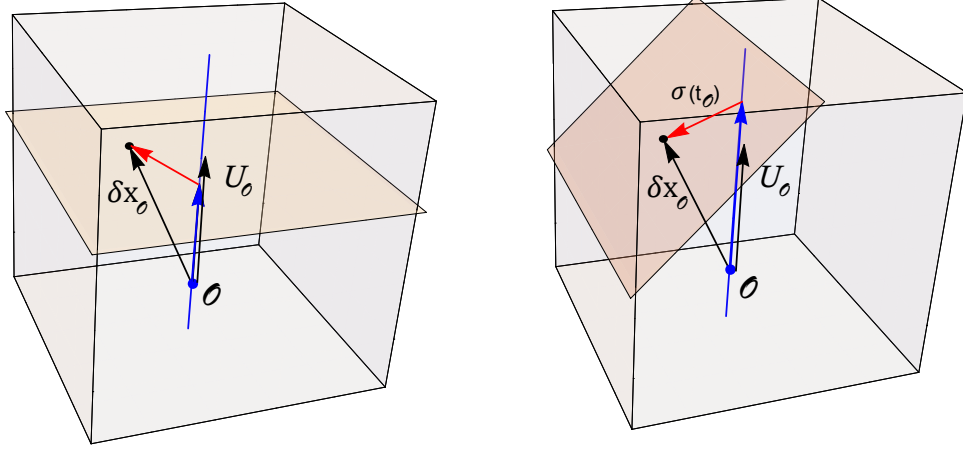


FIG. 11: The standard way to describe the position of a point mass $\delta x_{\mathcal{O}}$ is to decompose the position vector into the tangent and perpendicular components w.r.t. the barycenter 4-velocity $U_{\mathcal{O}}$ (on the left). Here we use a modified decomposition: $\delta x_{\mathcal{O}}$ decomposed into the part parallel to $U_{\mathcal{O}}$ and σ orthogonal to $l_{\mathcal{O}}$. The decompositions differ by the tilt of the 3-plane containing the second component.

observer can be decomposed according to

$$\delta x_{\mathcal{O}}^{\mu} = U_{\mathcal{O}}^{\mu} t_{\mathcal{O}} + \sigma^{\mu}(t_{\mathcal{O}}),$$

where the momentary position vector $\sigma^{\mu}(t_{\mathcal{O}})$ is assumed to be orthogonal to $l_{\mathcal{O}}^{\mu}$ rather than $U_{\mathcal{O}}^{\mu}$, i.e. $\sigma_{\mu} l_{\mathcal{O}}^{\mu} = 0$ (vector σ^{μ} can be spacelike or null). Similar decomposition can be used for the emitter:

$$\delta x_{\mathcal{E}}^{\mu} = U_{\mathcal{E}}^{\mu} t_{\mathcal{E}} + \rho^{\mu}(t_{\mathcal{E}})$$

with the condition $\rho_{\mu} l_{\mathcal{E}}^{\mu} = 0$. In $N_{\mathcal{O}}$ this decomposition is effectively equivalent to introducing a null time coordinate $v_{\mathcal{O}}$, consistent with the barycentric coordinate time $t_{\mathcal{O}}$ at the barycenter, but whose gradient is proportional to $l_{\mathcal{O}\mu}$, and then splitting the momentary displacement vector into the timelike component and the other component lying on the null surface of constant $v_{\mathcal{O}}$ (Fig. 11). This is, in turn, equivalent to the decomposition of $\delta x_{\mathcal{O}}$ in an SNF with $U_{\mathcal{O}}^{\mu}$ as the first basis vector. The purpose of this operation in both $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ is to take into account the Rømer delays on the observer's and the emitter's side. Note that the observer's and the emitter's proper times need to be related their respective barycentric coordinate times, but this is a fairly simple special relativity problem [62].

We now consider the apparent position of the emitter on the observer's sky. From the time lapse formula (49) we get the relation between the barycentric time variables $t_{\mathcal{O}}$ and $t_{\mathcal{E}}$, calculated at the barycenters:

$$t_{\mathcal{E}} = \frac{l_{\mathcal{O}\mu} U_{\mathcal{O}}^{\mu}}{l_{\mathcal{E}\nu} U_{\mathcal{E}}^{\nu}} t_{\mathcal{O}} = \frac{1}{1+z} t_{\mathcal{O}},$$

where z is the redshift between the two barycenter frames. Then from the direction deviation formula in the SNF (62) we get

$$\mathcal{D}^A_B \Delta l_{\mathcal{O}}^B = \left(\left(\frac{1}{1+z} U_{\mathcal{E}} - \hat{U}_{\mathcal{O}} \right)^A - m^A_i U_{\mathcal{O}}^i \right) t_{\mathcal{O}} + (\rho - \hat{\sigma})^A - m_{\perp}^A_B \sigma^B. \quad (76)$$

Assume now the observer measures the emitter's apparent position on the sky along his or her worldline, but subtracting the effects of aberration due to his or her motion with respect to the barycenter. In the terminology of Section IV A this amounts to the parallax with respect to the local inertial frame, connected with the free-falling barycenter, and given by $U_{\mathcal{O}}$. The apparent position with respect to $l_{\mathcal{O}}$ can be obtained from (76) combined with (38) and (44). After rearranging the terms and applying the definitions of the magnification matrix, the parallax matrix and the position drift formula (75) we obtain

$$\delta\theta^{\mathbf{A}} = \delta_{\mathcal{O}} r^{\mathbf{A}} t_{\mathcal{O}} + M^{\mathbf{A}}{}_{\mathbf{B}} \rho^{\mathbf{B}}(t_{\mathcal{O}}) - \Pi^{\mathbf{A}}{}_{\mathbf{B}} \hat{\sigma}^{\mathbf{B}}(t_{\mathcal{O}}). \quad (77)$$

In this formula $\delta_{\mathcal{O}} r^{\mathbf{A}}$ in the first term is the position drift rate calculated for a fictitious emitter-observer pair in which both are located at the free-falling barycenters of their gravitationally bound systems:

$$\delta_{\mathcal{O}} r^{\mathbf{A}} = \frac{1}{l_{\mathcal{O}\mu} U_{\mathcal{O}}^{\mu}} \mathcal{D}^{-1\mathbf{A}}{}_{\mathbf{B}} \left(\left(\frac{1}{1+z} U_{\mathcal{E}} - \hat{U}_{\mathcal{O}} \right)^{\mathbf{B}} - m^{\mathbf{B}}{}_{\mathbf{i}} U_{\mathcal{O}}^{\mathbf{i}} \right).$$

$M^{\mathbf{A}}{}_{\mathbf{B}}$ and $\Pi^{\mathbf{A}}{}_{\mathbf{B}}$ are the magnification and the parallax matrix in the observer's barycenter reference frame. We can see that the total effect is a sum of a secular drift caused by the motions of the barycenters of both gravitationally bound systems (proper motion), an oscillatory term due to the emitter's orbital motions and another oscillatory term due to the observer's orbital motion. The latter two terms depend only on the perpendicular components of the deviation vectors σ^{μ} and ρ^{μ} . This is a straightforward consequence of the PRA which neglects the perspective distortions. Note however that in the presence of curvature they are multiplied by two different matrices: the magnification matrix and the parallax matrix, differing by a curvature correction proportional to $m_{\perp}{}^{\mathbf{A}}{}_{\mathbf{B}}$.

If the emitter's short-time motion term is negligible with respect to the third term then the result is simply a sum of the proper motion and the classic parallax term [63, 74]. In the most important case for modern astronomy, i.e. the question of the solar parallax $\sigma^{\mathbf{A}}(t_{\mathcal{O}})$ has a very precisely known functional form with an annual periodicity and can be easily disentangled from the first, secular term [74]. This way we may measure the parallax matrix in the Solar System's barycenter frame using the standard astrometric observations performed along a timelike worldline of the Earth-based observatory or a spacecraft. Note however that the applicability of this procedure relies on the assumptions that underlie the analysis above: that the emitter does not undergo short-scale motions of similar time scale, the gravitational field is such that the curvature is roughly constant across a connecting tube of the size of the Earth's orbit, the distortions due to the light bending from nearby masses can be disregarded or subtracted and we know sufficiently well the short-scale motion of the observer around the Solar System's barycenter. The analysis above applies also to the case when the source is positioned at cosmological distances and we need to take into account the non-flat geometry between the source and the observer, or when the image has undergone strong lensing.

F. Magnification, parallax and position drift near a caustic

We can now use formulas (65), (71) and (75) to discuss the behaviour of the magnification, classical parallax and the position drift near a typical caustic. Recall that on a caustic the Jacobi map becomes degenerate in at least one direction. Assume now $\mathcal{D}^{\mathbf{A}}{}_{\mathbf{B}}$ is degenerate along a single direction $n^{\mathbf{A}}$, i.e. $\mathcal{D}^{\mathbf{A}}{}_{\mathbf{B}} n^{\mathbf{B}} = 0$. This means that an infinitesimally small image undergoes a formally infinite distortion along $n^{\mathbf{B}}$ and an infinite magnification. Consider now the parallax matrix: as long as the curvature along the line of sight is bounded, the emitter-observer asymmetry operator m_{\perp} should stay finite. Moreover we may expect the combination $\delta^{\mathbf{A}}{}_{\mathbf{B}} + m_{\perp}{}^{\mathbf{A}}{}_{\mathbf{B}}$ to be an invertible operator in a generic case. Now, since (71) contains the inverse of the Jacobi map, we see that unless the term $\delta^{\mathbf{A}}{}_{\mathbf{B}} + m_{\perp}{}^{\mathbf{A}}{}_{\mathbf{B}}$ happens by chance to be degenerate along the very same direction $n^{\mathbf{A}}$ the parallax matrix will blow up as well. In that case there obviously must be a direction $\bar{n}^{\mathbf{A}}$ in $\mathcal{P}_{\mathcal{O}}$ such that the parallax for an observer displaced along it is formally infinite. It is given by $\bar{n}^{\mathbf{A}} = C (\delta^{\mathbf{A}}{}_{\mathbf{B}} + m_{\perp}{}^{\mathbf{A}}{}_{\mathbf{B}})^{-1} n^{\mathbf{B}}$, with C being a positive normalization factor.

We can also obtain a similar conclusion works for the position drift. Looking at (75) we note that unless the 4-velocities of the observer and the emitter happen to be aligned in a special way, such that the combination $[(1+z)^{-1} u_{\mathcal{E}} - \hat{u}_{\mathcal{O}}] - m([u_{\mathcal{O}}])$ has a vanishing component along $n^{\mathbf{A}}$, the position drift becomes infinite as well.

Summarizing, we have just proved that near a caustic the magnification matrix, the parallax matrix and the value of the position drift, registered for a generic observer and emitter pair, will blow up simultaneously as measured on the observer's sky. Of course for real sources, of small but finite size, and for observational instruments of finite resolution, the measured values of the drift, parallax, and magnification will be large but finite. Nevertheless, in a generic situation, all of those effects become amplified *simultaneously* when the emitter approaches a caustic.

V. MOMENTARY MOTIONS-INDEPENDENT OBSERVABLES

As an example of an application of the formalism presented above, we will show that by combining the data about the classical parallax and the image distortion and magnification it is possible to define quantities which are entirely insensitive to the momentary 4-velocities of both the observer and the emitter. In other words, we will introduce momentary motions-independent observables, which measure the geometry of the spacetime between the emission point \mathcal{E} and the observation point \mathcal{O} , encoded in the optical operators.

Recall that the magnification matrix M^A_B and the parallax matrix Π^A_B do not depend on the emitter's 4-velocity $u_{\mathcal{E}}$, but they do depend on the observer's 4-velocity $u_{\mathcal{O}}$ because of the stellar aberration effect. Consider now the combination

$$w_{\perp}^A_B = (M^{-1})^A_C \Pi^C_B \quad (78)$$

calculated in a SNF. From (65) and (71) we see that the formula above defines an $u_{\mathcal{O}}$ -independent quantity, i.e. a combination of observables depending only on the curvature along the line of sight and not on the kinematical quantities describing the motions:

$$w_{\perp}^A_B \equiv w_{\perp}^A_B(R^{\mu}_{\nu\alpha\beta}).$$

Geometrically this matrix defines a frame-independent operator $w_{\perp} : \mathcal{P}_{\mathcal{O}} \rightarrow \mathcal{P}_{\mathcal{O}}$. Unlike M and Π it is dimensionless. From (65) and we see that it has a simple expression in terms of the optical operators

$$w_{\perp}^A_B = \delta^A_B + m_{\perp}^A_B. \quad (79)$$

Therefore its deviation from the unit matrix may serve as a measure of the spacetime curvature.

The value of the determinant of w_{\perp} , calculated in a SNF, is of particular interest. We define the dimensionless, scalar parameter μ by

$$\mu = 1 - \det w_{\perp}^A_B, \quad (80)$$

or equivalently

$$\mu = 1 - \frac{\det \Pi^A_B}{\det M^A_B}. \quad (81)$$

Just like w_{\perp} , it depends only on the curvature along the optical axis, but not on the motions of the observer and the emitter. Using (68), (72) and (81) we show that μ can be expressed via the parallax and the angular diameter distance:

$$\mu = 1 - \sigma \frac{D_{ang}^2}{D_{par}^2}, \quad (82)$$

where $\sigma = \pm 1$ determines the sign of the second term. σ depends on sign of the determinants of the magnification and the parallax matrices:

$$\sigma = \text{sgn} \det M^A_B \cdot \text{sgn} \det \Pi^A_B.$$

In short, we take the minus sign in the second term of (82) (i.e. $\sigma = 1$) if both matrices are orientation-preserving (i.e. the image the observer sees is not flipped and neither is the dependence of the parallax on position deviation) or both are negative (i.e. if the observer sees an inverted image and at the same time the linear dependence of the parallax on the position has inverted parity with respect to the standard one) and the plus sign ($\sigma = -1$) if only one of them is flipped. The reader may check that for sufficiently small perturbation of the null geodesics by curvature both determinants should be positive, so we have

$$\mu = 1 - \frac{D_{ang}^2}{D_{par}^2} \quad (83)$$

for sufficiently short distances and/or sufficiently weak bending of light rays between \mathcal{O} and \mathcal{E} . Note that there are no simple relations analogous to (82), (80) and (79) for the other parallax distance \tilde{D}_{par} .

We can see that μ measures the deviation of the metric from the flat one by comparing the parallax and the angular diameter distances measured to an object positioned far away. Obviously, both definitions must give the same

answer in a flat spacetime, i.e. $\mu = 0$ in the Minkowski space, but non-vanishing curvature along γ_0 gives rise to the asymmetry between the observer and the emitter, as we have discussed in Section IV C. This, in turn, can make the two optical methods of determining the distance inequivalent, giving rise to $\mu \neq 0$.

It is very instructive to consider μ in the case of fairly small curvature along the line of sight. We assume we may use the first order perturbation theory in the GDE, effectively treating the curvature tensor as a small perturbation. In that case we obtain from (59)

$$\ddot{m}_\perp^A{}_B \equiv \ddot{m}^A{}_B = R^A{}_{\mu\nu B} l^\mu l^\nu,$$

in the leading order. After imposing the initial conditions (60)-(61) we obtain the solution as an iterated integral:

$$m_\perp^A{}_B \approx \int_{\lambda_0}^{\lambda_\mathcal{E}} d\lambda_1 \int_{\lambda_0}^{\lambda_1} d\lambda_2 R^A{}_{\mu\nu B}(\lambda_2) l^\mu l^\nu.$$

We also linearize (80) around $\mu = 0$, obtaining

$$\mu \approx -m_\perp^A{}_A = -\text{tr } m_\perp.$$

Thus the leading order contribution to μ has also the form of an iterated integral

$$\mu \approx - \int_{\lambda_0}^{\lambda_\mathcal{E}} d\lambda_1 \int_{\lambda_0}^{\lambda_1} d\lambda_2 R^A{}_{\mu\nu A} l^\mu l^\nu. \quad (84)$$

Finally we apply the standard decomposition of the Riemann tensor into the Weyl tensor $C^\mu{}_{\nu\alpha\beta}$ and the Ricci tensor $R_{\mu\nu}$. We note first that the trace over the screen space of the Riemann tensor contracted twice with the null vector l^μ , or the optical tidal matrix, is equal to the full trace of the Riemann tensor contracted in the same way, i.e.

$$R^A{}_{\mu\nu A} l^\mu l^\nu = R^\alpha{}_{\mu\nu\alpha} l^\mu l^\nu = -R_{\mu\nu} l^\mu l^\nu.$$

Thus the Weyl tensor does not contribute to the integral (84). We also notice that for a null vector l^μ we have $R_{\mu\nu} l^\mu l^\nu = G_{\mu\nu} l^\mu l^\nu$, where $G_{\mu\nu}$ is the Einstein tensor.

In the final step we apply the Einstein field equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

contracted with $l^\mu l^\nu$. Again we see that because l^μ is null the cosmological constant Λ does not contribute to the integral, so we obtain

$$\mu \approx 8\pi G \int_{\lambda_0}^{\lambda_\mathcal{E}} d\lambda_1 \int_{\lambda_0}^{\lambda_1} d\lambda_2 T_{\mu\nu}(\lambda_2) l^\mu l^\nu \quad (85)$$

in the leading order in the curvature. Note that although we have used a parallel-propagated SNF to derive it, this formula is valid in any basis, including every coordinate basis. We see that for small curvature effects the Weyl tensor, carrying the information about tidal forces and gravitational waves along the line of sight, and the cosmological constant drop out from the integral, leaving only the dependence on the matter content along γ_0 and in its vicinity. The iterated integral in (85) can be converted into a single integral of the same expression with a linear weight function:

$$\mu \approx 8\pi G \int_{\lambda_0}^{\lambda_\mathcal{E}} T_{\mu\nu}(\lambda) l^\mu l^\nu (\lambda_\mathcal{E} - \lambda) d\lambda. \quad (86)$$

The proof of equivalence of (85) and (86) proceeds via the integration by parts of (86), with the linear function $(\lambda_\mathcal{E} - \lambda)$ undergoing differentiation and the stress-energy tensor term being integrated.

Finally, the reader may verify that all expressions for μ are invariant with respect to affine reparametrizations of γ_0 given by (2)-(3).

a. Applications. We will sketch now a simple application of the result above. Since the formulas (85) or (86) relate μ , a quantity potentially measurable using optical observations, with the amount of matter along the line of sight, we may use them to devise a purely optical method of determining the spacetime matter distribution. Consider a very precise, momentary measurement of the size of the image of a small object of known physical size at \mathcal{E} , performed by a telescope at \mathcal{O} from very far away. We assume that the DOA holds in this configuration. At the same moment,

we need to perform equally precise measurements of its classic parallax from at least two other, non-colinear points nearby, displaced orthogonally to the direction of light propagation. The measurement is done by comparing the two-dimensional position on the sky of the source, seen by the two displaced and comoving auxiliary observers, with the source's position recorded by the telescope \mathcal{O} . Since both auxiliary observers are displaced strictly orthogonally to the direction of propagation of light at \mathcal{O} , performing the measurements simultaneously in the observer's frame will yield the measurement of parallax of a single event along the emitter's worldline. Therefore what we measure this way is indeed the classic parallax in the terminology of Section IV A.

We assume that the physical size of the emitter (in its own frame) and the positions of the auxiliary observers with respect to the central one (in the observer's frame) are known very accurately. From these data we can determine with high precision the matrices Π^A_B and M^A_B along a null geodesic using directly the relations (65) and (71). Then from (81) we obtain μ . This way we have effectively weighed the whole matter content in the spacetime along the line of sight: from (85) we see that this method determines the amount of any kind of matter between the observer and the emitter.

We note that the method sketched above seems to be well-suited for a space-based mission. Due to the insensitivity of μ to the momentary motions on both sides, there is absolutely no need to know the relative motions or the precise distance *between* the observer and the emitter. It is only the distances and velocities *within* the group of observers measured *in the observer's own frame*, as well as the emitter's size, measured *in its own frame*, which are used in this measurement and which need to be determined with high precision. The exact shape of the emitter is also irrelevant since it only provides a background image. The angular size of this image and its apparent shift as observed from 2 other points carry the information about the spacetime curvature along the line of sight.

The measurement is also highly selective regarding the matter it takes into account. Masses located off the optical axis may introduce a measurable image displacement due to the gravitational light bending as well as image distortions by their tidal fields. Nevertheless, their influence on the value of μ is negligible. This is because distant masses may only influence the results of the observations via the tidal effects encoded in the Weyl curvature tensor along the line of sight. This influence, however, drops out of the trace in (84), leaving just the integral of the energy density *on* the optical axis. Thus, at least within the range of applicability of the approximations from this paper, the measurement of μ effectively cuts out a thin tube around the fiducial geodesic through the matter distribution and neglects any influence of the gravitational field sources lying outside it. On the other hand, note that repeated measurements along *different* null geodesics may provide a tomography-like method to determine a map of the matter content of a given region. A more detailed discussion of the parameter μ and its applications will be provided in a separate paper [75].

Comparing the parallax distance D_{par} or \tilde{D}_{par} and the angular diameter distance D_{ang} (or the closely related luminosity distance D_{lum}) as a method of determining the spacetime geometry has a long history in relativistic cosmology, beginning with McCrea [57]. Weinberg [58] noticed that in a perfectly homogeneous FLRW model comparing the parallax distance and the luminosity distance as functions of the redshift allows one to obtain the spatial curvature, which is impossible to determine by the luminosity distance observations alone. Rosquist [60] derived a differential relation between the two distances valid under the assumption of no shear and noticed that the comparison of both distances yields information about an otherwise unobservable component of the spacetime metric in the observational coordinates introduced by Ellis, Nel, Maartens, Stoeger and Whitman [30]. Kasai [59] considered the parallax distance \tilde{D}_{par} in an FLRW model with first-order perturbations, comparing of the results to the expressions for the luminosity distance D_{lum} in the same setting. Finally, Räsänen [9] proposed an FLRW consistency condition based on the comparison of the two types of distance measures. None of these works, however, mentions the independence of μ from momentary motions or its direct relation to the curvature and the matter content along the line of sight.

VI. SUMMARY AND REMARKS

We have presented a general, geometric approach to the problem of geometric optics in general relativity. It concerns the problems of observations of luminous objects or outbursts of radiation, contained within a small “stage” region $N_{\mathcal{E}}$, from a large distance by observers contained in another small “auditorium” region $N_{\mathcal{O}}$. Both regions are assumed to be small enough to be considered effectively flat and light propagation is treated within the geometric optics approximation. The approach works under rather general assumptions and should, therefore, apply to observations within the Solar System, the parallax measurements within the Galaxy, as well as cosmological observations.

The problem of observation is divided into two separate problems: the question of propagation of light through the inhomogeneous spacetime between the regions and the problem of the dependence of the results of observations from the motions in both regions. The second problem lies within the range of applicability of special relativity and is fairly easy to formulate in a geometric way. The first problem is considered using the first order geodesic deviation equation around a known, fiducial null geodesic, or the optical axis. In this formulation, the behavior of geodesics near the fiducial one is determined by a second-order linear ODE with the curvature playing the role of one of the

coefficients. The problem of light propagation, considered within the distant observer approximation, turns out to be fairly simple to reformulate in a frame- and observer-independent way: the tangent vector the fiducial geodesic defines two corresponding foliations of the two regions $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ by null hypersurfaces. Only points lying on the corresponding leaves of the foliations in $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$ can be connected by null geodesics (equation (49)). The action of curvature on the light propagation, on the other hand, is completely encoded in two optical operators, the well-known Jacobi operator \mathcal{D} and the emitter-observer asymmetry operator m , defined in Section III C. Both are most conveniently defined as bilocal operators acting from a quotient space on the observer's side to an appropriate quotient space at the emitter's end. Both can be expressed as functionals of the curvature along the line of sight (equations (56)-(58) and (59)-(61)) and they do not depend on the coordinate systems, bases, observers or any other structures defined on the manifold. m quantifies the difference between the direction variations measured by an observer due to perpendicular displacements of the null geodesic endpoints in $N_{\mathcal{O}}$ and $N_{\mathcal{E}}$. Unlike \mathcal{D} , m vanishes in a flat space and therefore it measures directly the impact of the curvature on the optical observations. This makes m and quantities derived from it excellent probes of the curvature along the line of sight.

We have then shown how one can state within this framework the problems of the parallax, the position drift (rate of change of the apparent position on the sky) and the gravitational lensing in a frame- and observer-independent way. All observables (i.e. the positions on the observers' sky and their rates of change in the observers' proper time) can be obtained from the two optical operators and the data characterizing the motions of the emitter and observer: their momentary 4-velocities, their displacement with respect to the fiducial null geodesic and the 4-acceleration of the observer with respect to his or her local inertial frame. In the resulting formulas, the effects of spacetime geometry on the light propagation and the effects of motions on both ends of the null geodesics are clearly separated. The geometric machinery allowed also to compare and contrast various definitions of parallax appearing in the relativistic literature and discuss relations between them. Additionally, we managed to show that in a generic situation the parallax and drift effects blow up along with the magnification of the image of the source as it passes through a caustic. The underlying reason is that the expressions for those effects involve the inverse of the Jacobi map, which by definition becomes degenerate at a caustic.

We remind the reader that the kinematic quantities appearing in the formalism, i.e. the 4-velocities, 4-accelerations and displacements of the observers and emitters, are defined *with respect to the local inertial frames* in the “stage” and “auditorium” regions. Those frames represent the local non-rotating reference frames falling freely in the large-scale gravitational fields. Therefore in the formalism presented here, the results do not depend explicitly on the potentially complicated details of motions of the observers and the emitters with respect to any external masses generating the gravitational fields: all that matters is their motions expressed in the locally flat coordinates defined within their respective neighbourhoods. This is, of course, a consequence of the GR equivalence principle applied to geometric optics: within our approximation, the dependence of observables on the momentary motions is a consequence of purely special relativistic effects. The SR effects, of course, cannot depend on the local details of the large-scale gravitational field. Thus the only possible dependence on the gravitational field generated by external objects enters via the curvature tensor along the line of sight appearing in the formulas for the optical operators.

The reformulation of the geometric optics in terms of the optical operators yielded finally an unexpected result. It turns out that by combining the full data about the parallax of a faraway object, contained in the parallax matrix (71), and the data about the magnification and distortion of its image, given by the magnification matrix (65), we may define observables which are entirely independent of the motions of both the observer and the source, given by equation (78). In other words, neither the relative motion of the observer and the emitter, difficult to determine for very long distances nor the motion of any of them with respect to a local inertial frame has any influence on their values. This is in a stark contrast to the standard observables like the redshift or luminosity distance, which always depend additionally at least on the 4-velocity of the observer [1]. The new observables probe exclusively the spacetime geometry between the regions $N_{\mathcal{E}}$ and $N_{\mathcal{O}}$. The simplest of them, i.e. the ratio between the suitably defined parallax distance and the angular diameter distance (equations (81)-(82)), contains information about the amount of matter along the line of sight for short distances, as seen in equation (86).

Acknowledgments

The work was supported by the National Science Centre, Poland (NCN) via the SONATA BIS programme, grant No 2016/22/E/ST9/00578 for the project “*Local relativistic perturbative framework in hydrodynamics and general relativity and its application to cosmology*”. The authors would also like to thank Andrzej Krasiński, Nezihe Uzun, Eleonora Villa, Thomas Buchert and Edward Malec for useful discussions and comments. MK would also like to thank the Max Planck Institute for Gravitational Physics in Potsdam (Albert-Einstein-Institut) for hospitality and

support.

-
- [1] M. Korzyński and J. Kopiński, *Journal of Cosmology and Astroparticle Physics* **2018**, 012 (2018), URL <http://stacks.iop.org/1475-7516/2018/i=03/a=012>.
- [2] R. A. Isaacson, *Phys. Rev.* **166**, 1263 (1968), URL <https://link.aps.org/doi/10.1103/PhysRev.166.1263>.
- [3] R. A. Isaacson, *Phys. Rev.* **166**, 1272 (1968), URL <https://link.aps.org/doi/10.1103/PhysRev.166.1272>.
- [4] M. Blau, D. Frank, and S. Weiss, *Class. Quant. Grav.* **23**, 3993 (2006), hep-th/0603109.
- [5] J. Wambsganss, *Living Reviews in Relativity* **1**, 12 (1998), ISSN 1433-8351, URL <https://doi.org/10.12942/lrr-1998-12>.
- [6] P. Schneider, C. S. Kochanek, and J. Wambsganss, *Gravitational Lensing: Strong, Weak and Micro*, vol. 33 of *Saas-Fee Advanced Course* (Springer-Verlag Berlin Heidelberg, 2006).
- [7] C. Quercellini, M. Quartin, and L. Amendola, *Phys. Rev. Lett.* **102**, 151302 (2009), 0809.3675.
- [8] C. Quercellini, L. Amendola, A. Balbi, P. Cabella, and M. Quartin, *Physics Reports* **521**, 95 (2012), ISSN 0370-1573, real-Time Cosmology, URL <http://www.sciencedirect.com/science/article/pii/S0370157312002633>.
- [9] S. Räsänen, *Journal of Cosmology and Astroparticle Physics* **2014**, 035 (2014), URL <http://stacks.iop.org/1475-7516/2014/i=03/a=035>.
- [10] M. Fontanini, E. J. West, and M. Trodden, *Phys. Rev.* **D80**, 123515 (2009), 0905.3727.
- [11] C. Quercellini, P. Cabella, L. Amendola, M. Quartin, and A. Balbi, *Phys. Rev.* **D80**, 063527 (2009), 0905.4853.
- [12] A. Krasinski and K. Bolejko (2012), [Int. J. Mod. Phys.D22,0013(2013)], 1212.4697.
- [13] A. Krasinski, *Phys. Rev.* **D84**, 023510 (2011), 1105.1299.
- [14] A. Krasinski, *Phys. Rev.* **D86**, 064001 (2012), 1205.6341.
- [15] A. Krasinski and K. Bolejko, *Phys. Rev.* **D83**, 083503 (2011), 1007.2083.
- [16] A. Krasinski and K. Bolejko, in *Proceedings, 13th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories (MG13): Stockholm, Sweden, July 1-7, 2012* (2015), pp. 922–934.
- [17] E. Bentivegna and M. Korzyński, *Class. Quant. Grav.* **29**, 165007 (2012), 1204.3568.
- [18] C.-M. Yoo, H. Okawa, and K.-i. Nakao, *Phys. Rev. Lett.* **111**, 161102 (2013), 1306.1389.
- [19] E. Bentivegna and M. Korzyński, *Class. Quant. Grav.* **30**, 235008 (2013), 1306.4055.
- [20] C.-M. Yoo and H. Okawa, *Phys. Rev.* **D89**, 123502 (2014), 1404.1435.
- [21] E. Bentivegna and M. Bruni, *Phys. Rev. Lett.* **116**, 251302 (2016), 1511.05124.
- [22] J. T. Giblin, J. B. Mertens, and G. D. Starkman, *Phys. Rev. Lett.* **116**, 251301 (2016), 1511.01105.
- [23] J. B. Mertens, J. T. Giblin, and G. D. Starkman, *Phys. Rev.* **D93**, 124059 (2016), 1511.01106.
- [24] J. Adamek, D. Daverio, R. Durrer, and M. Kunz, *Nature Phys.* **12**, 346 (2016), 1509.01699.
- [25] J. Adamek, D. Daverio, R. Durrer, and M. Kunz, *JCAP* **1607**, 053 (2016), 1604.06065.
- [26] H. J. Macpherson, P. D. Lasky, and D. J. Price, *Phys. Rev.* **D95**, 064028 (2017), 1611.05447.
- [27] H. J. Macpherson, P. D. Lasky, and D. J. Price, *Astrophys. J.* **865**, L4 (2018), 1807.01714.
- [28] F. H. Vincent, E. Gourgoulhon, and J. Novak, *Class. Quant. Grav.* **29**, 245005 (2012), 1208.3927.
- [29] E. Bentivegna, M. Korzyński, I. Hinder, and D. Gerlicher, *JCAP* **1703**, 014 (2017), 1611.09275.
- [30] G. Ellis, S. Nel, R. Maartens, W. Stoeger, and A. Whitman, *Physics Reports* **124**, 315 (1985), ISSN 0370-1573, URL <http://www.sciencedirect.com/science/article/pii/0370157385900304>.
- [31] G. Fanizza, M. Gasperini, G. Marozzi, and G. Veneziano, *JCAP* **1508**, 020 (2015), 1506.02003.
- [32] A. I. Harte, *Gen. Rel. Grav.* **51**, 14 (2019), 1808.06203.
- [33] E. Poisson, A. Pound, and I. Vega, *Living Reviews in Relativity* **14**, 7 (2011), ISSN 1433-8351, URL <https://doi.org/10.12942/lrr-2011-7>.
- [34] S. L. Bazański, *Annales de L'Institut Henri Poincare Section (A) Physique Theorique* **27**, 115 (1977).
- [35] S. L. Bazański, *Annales de L'Institut Henri Poincare Section (A) Physique Theorique* **27**, 145 (1977).
- [36] J. Vines, *Gen. Rel. Grav.* **47**, 59 (2015), 1407.6992.
- [37] N. Uzun (2018), 1811.10917.
- [38] F. A. E. Pirani, *Acta Phys. Polon.* **15**, 389 (1956).
- [39] P. Szekeres, *Journal of Mathematical Physics* **6**, 1387 (1965), <https://doi.org/10.1063/1.1704788>, URL <https://doi.org/10.1063/1.1704788>.
- [40] A. N. Aleksandrov and K. A. Piragas, *Theoretical and Mathematical Physics* **38**, 48 (1979), ISSN 1573-9333, URL <https://doi.org/10.1007/BF01030257>.
- [41] I. Ciufolini and M. Demiański, *Phys. Rev. D* **34**, 1018 (1986), URL <https://link.aps.org/doi/10.1103/PhysRevD.34.1018>.
- [42] D. Puetzfeld and Y. N. Obukhov, *Phys. Rev.* **D93**, 044073 (2016), 1511.08465.
- [43] É. É. Flanagan, A. M. Grant, A. I. Harte, and D. A. Nichols (2018), 1901.00021.
- [44] M. Bartelmann, *Classical and Quantum Gravity* **27**, 233001 (2010), URL <https://doi.org/10.1088%2F0264-9381%2F27%2F233001>.
- [45] C. Clarkson, *Class. Quant. Grav.* **33**, 16LT01 (2016), 1603.04698.

- [46] C. Clarkson, *Class. Quant. Grav.* **33**, 245003 (2016), 1603.04652.
- [47] J. L. Synge, *Relativity: The general theory* (North-Holland Publishing Company, Amsterdam, 1960).
- [48] P. Fleury, *JCAP* **1406**, 054 (2014), 1402.3123.
- [49] B. S. DeWitt and R. W. Brehme, *Annals of Physics* **9**, 220 (1960), ISSN 0003-4916, URL <http://www.sciencedirect.com/science/article/pii/0003491660900300>.
- [50] W. G. Dixon, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **314**, 499 (1970).
- [51] V. Perlick, *Journal of Mathematical Physics* **31**, 1962 (1990), <http://dx.doi.org/10.1063/1.528645>, URL <http://dx.doi.org/10.1063/1.528645>.
- [52] R. J. Low, *Journal of Mathematical Physics* **34**, 315 (1993).
- [53] D.-E. Liebscher and P. Brosche, *Astronomische Nachrichten* **319**, 309 (1998), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/asna.2123190506>, URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/asna.2123190506>.
- [54] V. Perlick, *Living Reviews in Relativity* **7**, 9 (2004), ISSN 1433-8351, URL <https://doi.org/10.12942/lrr-2004-9>.
- [55] R. Sachs, *Proceedings of the Royal Society of London Series A* **264**, 309 (1961).
- [56] J. Ehlers, P. Jordan, and R. K. Sachs, *Beiträge zur Theorie der reinen Gravitationsstrahlung*, vol. 1 of *Abhandlungen der Mathematisch-Naturwissenschaftlichen Klasse* (Verlag der Akademie der Wissenschaften und der Literatur in Mainz, Wiesbaden, Germany, 1961).
- [57] W. H. McCrea, *Zeitschrift für Astrophysik* **9**, 290 (1935).
- [58] S. Weinberg, *The Astrophysical Journal* **161**, L233 (1970).
- [59] M. Kasai, *Prog. Theor. Phys.* **79**, 777 (1988), URL <http://dx.doi.org/10.1143/PTP.79.777>.
- [60] K. Rosquist, *Astrophys. J.* **331**, 648 (1988), URL <http://dx.doi.org/10.1086/166588>.
- [61] O. H. Marcori, C. Pitrou, J.-P. Uzan, and T. S. Pereira, *Phys. Rev. D* **98**, 023517 (2018), 1805.12121.
- [62] S. A. Klioner, *The Astronomical Journal* **125**, 1580 (2003).
- [63] L. Lindegren, U. Lammers, D. Hobbs, W. O'Mullane, U. Bastian, and J. Hernandez, *Astron. Astrophys.* **538**, A78 (2012), 1112.4139.
- [64] C. Hellaby and A. Walters, *JCAP* **1802**, 015 (2018), 1708.01031.
- [65] E. Fermi, *Rend. Lincei*, 1922, 31 (1), pp. 21-23, 51-52, 101-103 (in Italian) **31**, 21 (1922).
- [66] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Company, New York, 1973).
- [67] M. N. R. Wohlfarth and C. Pfeifer, *Phys. Rev. D* **87**, 024031 (2013), URL <https://link.aps.org/doi/10.1103/PhysRevD.87.024031>.
- [68] C. Bonvin, *Phys. Rev. D* **78**, 123530 (2008), 0810.0180.
- [69] I. Etherington, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **15**, 761 (1933), <http://dx.doi.org/10.1080/14786443309462220>, URL <http://dx.doi.org/10.1080/14786443309462220>.
- [70] I. M. H. Etherington, *General Relativity and Gravitation* **39**, 1055 (2007), ISSN 1572-9532, URL <https://doi.org/10.1007/s10714-007-0447-x>.
- [71] M. A. C. Perryman, K. S. de Boer, G. Gilmore, E. Høg, M. G. Lattanzi, L. Lindegren, X. Luri, F. Mignard, O. Pace, and P. T. de Zeeuw, *Astronomy and Astrophysics* **369**, 339 (2001), astro-ph/0101235.
- [72] A. G. Riess, S. Casertano, J. Anderson, J. Mackenty, and A. V. Filippenko, *Astrophys. J.* **785**, 161 (2014), 1401.0484.
- [73] W. O. Kermack, W. H. McCrea, and E. T. Whittaker, *Proceedings of the Royal Society of Edinburgh* **53**, 31-47 (1934).
- [74] X. Luri, A. G. A., Sarro, L. M., Arenou, F., Bailer-Jones, C. A. L., Castro-Ginard, A., Bruijne, J. de, Prusti, T., Babusiaux, C., and Delgado, H. E., *A&A* **616**, A9 (2018), URL <https://doi.org/10.1051/0004-6361/201832964>.
- [75] M. Grasso, M. Korzyński, and J. Serbenta, in preparation.
- [76] Note that unlike [48] we consider displacements not only in the two perpendicular, spatial directions, but in all 4 possible directions, including the time. The resulting system of ODE's has, therefore, a twice larger dimension. However, the Wronski matrix formalism extends to the fully 4-dimensional GDE without any problems. One can also check that the extended Wronski matrix is symplectic too, but this is beyond the scope of this article.
- [77] Another example for this nonlinearity is the second order vector differential equation $\frac{d^2 \mathbf{x}}{d\lambda^2} + A\mathbf{x} = 0$, where A is a constant diagonalisable matrix with positive eigenvalues. The general solution has the following form: $\mathbf{x} = \cos[\sqrt{A}(\lambda - \lambda_0)]\mathbf{x}_0 + \frac{1}{\sqrt{A}} \sin[\sqrt{A}(\lambda - \lambda_0)]\dot{\mathbf{x}}_0$. Even if A enters the differential equation linearly, the general solution contains a nonlinear dependence.