A FINITE DIMENSIONAL APPROACH TO LIGHT RAYS IN GENERAL RELATIVITY

ROBERTO GIAMBÒ, FABIO GIANNONI, AND PAOLO PICCIONE

ABSTRACT. We propose a finite dimensional setup for the study of lightlike geodesics starting orthogonally to a spacelike $(n-2)$-submanifold and arriving orthogonally to the time-slices of an $(n-1)$-dimensional timelike submanifold of a n-dimensional spacetime. Under a transversality and a nonfocality assumption, we prove a finite dimensional reduction of a general relativistic Fermat principle, and we give a formula for the Morse index. We present some applications to bifurcation theory, and we conclude the paper with the discussion of some examples that illustrate our results.

1. INTRODUCTION

In a general relativistic space-time, light rays from an extended light source to an extended receiver (a screen) are modeled by lightlike geodesics that are orthogonal at their endpoints to two given spacelike submanifolds. Considering the worldline Γ of a receiver, and assuming that this set is a stably causal Lorentzian hypersurface of the space-time, i.e., a hypersurface that admits a (smooth) time function as a Lorentzian manifold of its own, then the light rays starting orthogonally to the initial submanifold and terminating orthogonally to the time slices of Γ are characterized by Fermat's principle as stationary points of the arrival time functional, see ref. [17].

This variational principle lacks regularity, in that the set of trial paths on which the arrival time is to be considered, which consists of all (future pointing, piecewise smooth) lightlike curves between the source and the receiver, does not admit a differentiable structure. This is an obstruction to the application of analytical techniques, such as Lusternik–Schnirelman theory, Morse theory, or bifurcation theory, whose setup requires a quite elaborate functional framework, and it makes unfeasible the use of singularity theory. In this paper we propose a finite dimensional (smooth) reduction of the Fermat principle, which is suited to give a local description of the orthogonal light rays near a degenerate one, and that in particular allows a direct application of bifurcation theory and singularity theory to study the caustics. This aims naturally at establishing multiplicity results for light rays between sources and observers, which models the so-called multiple image effect and the gravitational lensing phenomenon in General Relativity. The interested reader will find a very extensive literature on the subject, see for instance reference [13], or the living review [16] for a detailed account of the recent bibliography. Fermat's principle in general relativistic optics, and its applications to gravitational lensing are discussed thoroughly in the monograph [15]. Important aspects of the theory of gravitational lensing are presented in the survey [18]. Of course, a finite dimensional approach to light rays can be obtained using the normal exponential map. However, an essential point of the finite dimensional reduction which is presented here is the fact that it preserves the variational structure of the problem, and therefore also suited for developing Morse theoretical techniques, or to assess stability results.

Our model proposes to study orthogonal light rays using the arrival time functional restricted to the finite dimensional manifold of lightlike geodesics issuing orthogonally to the initial spacelike manifold $\mathcal{P}_0$ (the extended light source), and arriving transversally onto a timelike hypersurface Γ (the worldline of an extended receiver). The arrival time of lightlike geodesics in Lorentzian geometry plays the same role that the (squared) distance function plays in the study of focal properties of submanifolds in Riemannian geometry. Generically, the focal set is the bifurcation or catastrophe set for the family of distance functions from ambient points, see [21, 22]. We assume that, with the induced metric, Γ is a *stably causal* Lorentzian manifold in itself, i.e., it admits a smooth time function $T\colon \Gamma \to \mathbb{R}$; for $\tau \in \mathbb{R}$, Γ_τ will denote the time slice $T^{-1}(\tau)$. In this situation, the arrival time is a smooth function in the space of light rays issuing from $\mathcal{P}_0$ and arriving on Γ, and under a nonfocality assumption (Section 2.3), its critical points correspond to light rays that arrive orthogonally to the time slices of Γ (Theorem 3.1). Moreover, a second order variational principle also holds, in the following sense. First, nondegenerate critical points p of the arrival time functional correspond exactly to nondegenerate orthogonal light rays ℓ_p. Second, the Morse index of the critical point p is equal to the Morse index of the geodesic action functional at ℓ_p minus the *focal index* of ℓ_p. Such difference can be easily interpreted *geometrically*: it is the so-called *concavity index* that appears in the Morse index theorem for orthogonal geodesics (see Theorem 2.1), and it is given in terms of the second fundamental form of the target manifold, computed in a space associated to the $\mathcal{P}_0$-Jacobi fields. Our Morse index theorem provides a *physical* interpretation of the concavity index form along an orthogonal lightlike geodesic, which is now seen as the second variation of the arrival time functional.

As to the nonfocality assumption needed for our theory, a simple counterexample shows that it cannot be omitted, see Example 1. A discussion on this assumption is presented in Section 3.2, where we show that focal points correspond indeed to focusing points of families of lightlike geodesics issuing orthogonally from $\mathcal{P}_0$. We will also show here that the nonfocality assumption can be replaced by the assumption that τ, defined in (3.1), has only nondegenerate critical points, see Corollary 3.4.

Also, using the variational principle introduced in this paper, we give a notion of stability for light rays between an extended light source and an extended receiver. By the Morse index theorem, stability is equivalent to the positive-definiteness of the concavity index form (Corollary 4.8). In the case of lightlike geodesics between a pointwise source and a pointwise observer, the Morse index is always given by the number of conjugate instants along the ray, which happens to be independent on the orientation of the geodesic, i.e., a future-pointing is stable if and only if its backwards past-point reparameterization is stable. One of the interesting consequences of our theory is the fact that, when one considers extended source and receiver, the Morse index and the notion of stability *do* indeed depend on the time orientation of the lightlike geodesic. Thus, an observer may have different notions of stability for the image of an extended source on an extended receiver depending on whether he/she is located at the source or at the target. This is discussed in Section 4.3, where a formula relating the Morse indices of a lightlike geodesics and its backwards reparameterization is given (Proposition 4.6). Explicit examples of situations of orthogonal lightlike geodesics whose Morse index (and stability) changes according to its orientation are illustrated in Section 5. It is also interesting to observe that one can have stability for light rays that have focal points, and instability for light rays that do not contain focal points, see Example 3.

A remarkable interpretation of the result of this paper arises in a general relativistic context, in particular concerning the quasi–local analysis of spacetime singularities. Indeed, it is well known that the classical singularity theorems by Hawking and Penrose, requiring the existence of a global horizon, can be of cumbersome application in terms of the Cauchy problem for the Einstein field equations [20]. This problem can be overcome, in some sense, introducing the so–called marginally outer trapped surface (MOTS), and considering the (global) horizon as foliated by a family of these MOTS, each a 2–codimensional closed spacelike surface [1, 2]. In view of this, one can interpret the manifold $\mathcal{P}_0 \subset M$ as a MOTS lying on a initial data Cauchy hypersurface.

As for the submanifold Γ, it can play a prominent role in relativistic holography, in case (M, g) is the conformal compactification of an asymptotically anti–deSitter spacetime. In this case the conformal infinity is in fact a $(n-1)$–codimensional timelike hypersurface, and its intersection with the horizon is of interest in the so–called *topological censorship* context [5].

We emphasize that the finite dimensional reduction discussed in the present paper does not aim at the development of an existence theory for light rays, but rather to the study of the local geometry of nontransversal intersections of the images of normal exponential map along a lightlike geodesic. Given a future-pointing lightlike geodesic $\ell_0 : [0, 1] \to M$ which is orthogonal at the endpoints to two spacelike submanifolds $\mathcal{P}_0$ and Γ_0, then the finite dimensional setup can be applied to a neighborhood of $\ell_0'(0)$ in the normal lightlike bundle of $\mathcal{P}_0$. This yields interesting results when one considers the time evolutions $\mathcal{P}_\tau$ of the light source $\mathcal{P}_0$ and Γ_τ of Γ_0, $\tau \in [-\varepsilon, \varepsilon]$: when the concavity index form of ℓ_0 is degenerate, and it changes its index as t crosses 0, then bifurcation of orthogonal lightlike geodesics occurs (Proposition 4.5). In this situation, the type of bifurcation is determined by the singularities of a smooth real valued map on a finite dimensional manifold. This provides a simplified approach to the study of the singularities of the exponential map and the geometry of caustics, which constitutes a quite active research area both in the classical and in the modern literature (see for instance [3, 11, 23] and the references therein).

Using a Riemannian setup, several explicit examples are illustrated in Section 5 in the case of standard static spacetimes. As to bifurcation of orthogonal light rays, we obtain interesting examples which show the role of focal points. Unlike the fixed endpoint case, bifurcation may occur in absence of focal points, see Example 2, and may not occur in presence of focal points, see Example 4.

Finally, in Section 6 we will discuss an application of our theory in a general relativistic optical model. Here, the finite dimensional reduction is employed to obtain a noncompactness result for the visible region of a light emitting body in a null concave optical domain.

Acknowledgement. The authors gratefully acknowledge the support of Ivan Pontual Costa e Silva, who made several interesting remarks and suggested the application in Section 6. This work was carried out during a sabbatical semester of P.P. at the Department of Mathematics of the University of Notre Dame. P.P. would like to thank all staff and faculty at Notre Dame who provided this wonderful opportunity to him.

2. The setup

2.1. **The objects.** Let us consider a setup similar to [17]:

- (M, g) is an n-dimensional time-oriented Lorentzian manifold, $n \geqslant 4$, and exp is its exponential map;
- $\mathcal{P}_0 \subset M$ is a $(n-2)$-dimensional spacelike submanifold, to be interpreted as the surface of the light source;
- Γ is a $(n-1)$-dimensional timelike submanifold of M which is *stably causal* as a Lorentz manifold itself, i.e., it admits a smooth time function $T\colon \Gamma \to \mathbb{R}$. We can assume that T is future-pointing with respect to the time orientation on Γ induced by that of (M, g). Such a Γ represents the history of the surface of the receiver.

In order to avoid trivialities, we will assume that $\mathcal{P}_0 \cap \Gamma = \emptyset$; it will also be assumed later that Γ has nonempty intersection with the causal future of $\mathcal{P}_0$, see (HP1) below.

For $\tau \in \mathbb{R}$, let Γ_τ denote the *time-slice* $T^{-1}(\tau)$ of Γ, which is an $(n-2)$-dimensional spacelike submanifold of Γ (and of M). We will be interested in light rays issuing orthogonally from $\mathcal{P}_0$ and arriving orthogonally to a time-slice of Γ. We will first recall a few general facts on lightlike geodesics orthogonal to spacelike endmanifolds.

2.2. **Generalities on lightlike geodesics with endpoints orthogonal to two spacelike submanifolds.** Let us consider an affinely parameterized lightlike geodesic $\ell\colon [a, b] \to M$ joining orthogonally two spacelike surfaces Σ_1 and Σ_2, i.e., with $\ell(a) \in \Sigma_1$, $\ell(b) \in \Sigma_2$, $\ell'(a) \in T_{\ell(a)}\Sigma_1^\perp$ and $\ell'(b) \in T_{\ell(b)}\Sigma_2^\perp$.

2.2.1. *The index form.* Associated to this triple $(\Sigma_1, \Sigma_2, \ell)$ one has the *index form* $I^\ell_{\Sigma_1,\Sigma_2}$, which is a symmetric bilinear form on the space of vector fields V along ℓ, satisfying $V(a) \in T_{\ell(a)}\Sigma_1$, $V(b) \in T_{\ell(b)}\Sigma_2$, defined by:

$$I^{\ell}_{\Sigma_1,\Sigma_2}(V,W) = \int_a^b g\big(\tfrac{DV}{dt}, \tfrac{DW}{dt}\big) + g\big(R(\ell',V)\ell', W\big)\,dt + S^2_{\ell'(b)}\big(V(b), W(b)\big) - S^1_{\ell'(a)}\big(V(a), W(a)\big), \tag{2.1}$$

where $\frac{D}{dt}$ is the Levi–Civita covariant derivative operator along the curve ℓ, R is the curvature tensor of the Levi-Civita connection ∇ of g, chosen with the sign convention $R(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]}$, and $S^1_{\ell'(a)}$ (resp., $S^2_{\ell'(b)}$) is the second fundamental form[1] of the surface Σ_1 (resp., Σ_2) at the point $\ell(a)$ (resp., $\ell(b)$), in the orthogonal direction $\ell'(a)$ (resp., $\ell'(b)$). We will use the same symbol $S^1_{\ell'(a)}$ (reps. $S^2_{\ell'(b)}$) to denote the g-symmetric the linear endomorphism of $T_{\ell(a)}\Sigma_1$ (resp., $T_{\ell(b)}\Sigma_2$) that represents the second fundamental form $S^1_{\ell'(a)}$ (resp., $S^2_{\ell'(b)}$). Considered the restriction of $I^{\ell}_{\Sigma_1,\Sigma_2}$ to the (infinite dimensional) space of (piecewise smooth) vector fields V along γ satisfying $V(a) \in T_{\ell(a)}\Sigma_1$, $V(b) \in T_{\ell(b)}\Sigma_2$, and $g\big(\frac{DV}{dt}, \ell'\big) = 0$, its index is a nonnegative finite integer, that will be denoted by $i(\Sigma_1, \Sigma_2; \ell)$. This number is the Morse index of ℓ as a critical point of the geodesic action functional defined in the (infinite dimensional) space of all curves with free endpoints on Σ_1 and Σ_2.

2.2.2. *Jacobi fields.* The kernel of $I^{\ell}_{\Sigma_1,\Sigma_2}$ is the space of (Σ_1,Σ_2)-*Jacobi fields along* ℓ, which is the space of all Jacobi fields J along ℓ satisfying the boundary conditions:

$$J(a) \in T_{\ell(a)}\Sigma_1, \quad \tfrac{DJ}{dt}(a) + S^1_{\ell'(a)}\big(J(a)\big) \in T_{\ell(a)}\Sigma_1^{\perp}. \tag{2.2}$$

and

$$J(b) \in T_{\ell(b)}\Sigma_2, \quad \tfrac{DJ}{dt}(b) + S^2_{\ell'(b)}\big(J(b)\big) \in T_{\ell(b)}\Sigma_2.^{\perp} \tag{2.3}$$

The geodesic ℓ, or more precisely the triple $(\Sigma_1, \Sigma_2, \ell)$, is said to be *nondegenerate* if $I^{\ell}_{\Sigma_1,\Sigma_2}$ is a nondegenerate bilinear form, i.e., if $\mathrm{Ker}\big(I^{\ell}_{\Sigma_1,\Sigma_2}\big) = \{0\}$.

A Σ_1-*Jacobi field along* ℓ is a Jacobi field J satisfying the initial condition (2.2). Let us denote by $\mathbb{J}(\Sigma_1,\ell)$ the space of all Σ_1-Jacobi fields along ℓ; moreover, for $t \in\,]a,b]$, we will denote by $\mathbb{J}(\Sigma_1,\ell)[t] \subset T_{\ell(t)}M$ the space:

$$\mathbb{J}(\Sigma_1,\ell)[t] = \big\{J(t) : J \in \mathbb{J}(\Sigma_1,\ell)\big\}.$$

An instant $t_0 \in\,]a,b]$ is said to be Σ_1-*focal* if there exists $J \in \mathbb{J}(\Sigma_1,\ell) \setminus \{0\}$ with $J(t_0) = 0$. The *multiplicity* of a Σ_1-focal point t_0 is the dimension of the space of all Σ_1-Jacobi fields along ℓ vanishing at t_0. Equivalently, t_0 is Σ_1-focal along ℓ iff $\mathbb{J}(\Sigma_1,\ell)[t_0] \neq T_{\ell(t_0)}M$, and the multiplicity of t_0 is the codimension of $\mathbb{J}(\Sigma_1,\ell)[t_0]$ in $T_{\ell(t_0)}M$.

The Σ_1-focal instant along ℓ form a finite set; the sum of their multiplicity is called the Σ_1-*focal index of* ℓ, and it will be denoted by $i(\Sigma_1;\ell)$.

[1] $S^1_{\ell'(a)}(v_1,v_2) = g\big(\ell'(a), \nabla_{v_1}V_2\big)$ for all $v_1, v_2 \in T_{\ell(a)}\Sigma_1$, where V_2 is any extension of v_2 to a local field tangent to Σ_1

Alternatively, an instant $t_0 \in \,]a, b]$ is Σ_1-focal along ℓ if and only if the vector $(t_0 - a)\ell'(a)$ is a singular point of the normal exponential map $\exp\colon T\Sigma_1^\perp \to M$. In particular, if t_0 is not Σ_1-focal along ℓ, then exp gives a smooth diffeomorphism between a neighborhood of $(t_0 - a)\ell'(a)$ in the normal bundle $T\Sigma_1^\perp$ of Σ_1 and an open neighborhood of $\ell(t_0)$ in M.

2.2.3. *The Morse index theorem.* Consider now the vector space $\mathbb{J}(\Sigma_1;\ell)^{\Sigma_2}$ of all Σ_1-Jacobi fields J along ℓ satisfying $J(b) \in T_{\ell(b)}\Sigma_2$. Let us emphasize that the Jacobi fields in $\mathbb{J}(\Sigma_1;\ell)^{\Sigma_2}$ are *not* (Σ_1, Σ_2)-Jacobi fields along ℓ, in that they only satisfy the first of the two boundary conditions (2.3) at b.

Denote by $\mathcal{C}^\ell_{\Sigma_1,\Sigma_2}$ the restriction of the index form $I^\ell_{\Sigma_1,\Sigma_2}$ defined in (2.1) to the space $\mathbb{J}(\Sigma_1;\ell)^{\Sigma_2}$; this is called the *concavity index form* of $(\Sigma_1, \Sigma_2;\ell)$. Using partial integration, it is easily computed:

$$\mathcal{C}^\ell_{\Sigma_1,\Sigma_2}(J_1, J_2) = g\big(\tfrac{DJ_1}{dt}(b), J_2(b)\big) + S^2_{\ell'(b)}\big(J_1(b), J_2(b)\big), \tag{2.4}$$

for all $J_1, J_2 \in \mathbb{J}(\Sigma_1;\ell)^{\Sigma_2}$. Let us denote by $i_{\mathrm{conc}}(\Sigma_1, \Sigma_2;\ell)$ the index of $\mathcal{C}^\ell_{\Sigma_1,\Sigma_2}$ in $\mathbb{J}(\Sigma_1;\ell)^{\Sigma_2}$.

The reader must observe that we have associated to the triple $(\Sigma_1, \Sigma_2, \ell)$ three different integers: $i(\Sigma_1, \Sigma_2;\ell)$, $i(\Sigma_1, \ell)$ and $i_{\mathrm{conc}}(\Sigma_1, \Sigma_2;\ell)$. These three numbers are related by a Morse index theorem.

A Morse index theorem in this situation has been proved first by Ehrlich and Kim in [4], then by Perlick and Piccione in [17], while a different and more general statement for arbitrary geodesics with endpoints orthogonal to arbitrary submanifolds can be found in [19]. Let us recall here the index theorem in its formulation given in [19].

2.1. **Morse Index Theorem.** *Under the assumption:*

$$T_{\ell(b)}\Sigma_2 \subset \mathbb{J}(\Sigma_1;\ell)[b], \tag{2.5}$$

the following equality holds:

$$i(\Sigma_1, \Sigma_2;\ell) = i(\Sigma_1;\ell) + i_{conc}(\Sigma_1, \Sigma_2;\ell). \quad \square \tag{2.6}$$

Let us observe that assumption (2.5) holds in particular when b is not a Σ_1-focal instant along ℓ, in which case $\mathbb{J}(\Sigma_1;\ell)[b] = T_{\ell(b)}M$. See Section 3.2 for a further discussion on this.

2.3. **Transversality and nonfocality assumptions.** Let us now consider assumptions on the above objects, following the lines of [8]. The central assumption is that, for each point $p \in \mathcal{P}_0$, there exists an orthogonal (future-pointing) lightlike geodesic ℓ_p issuing from p orthogonally to $\mathcal{P}_0$, which eventually meets Γ transversally at some nonfocal instant. More precisely, let $\mathcal{L}^+_{\mathcal{P}_0}$ denote the future pointing lightlike normal bundle along $\mathcal{P}_0$; observe that, since $\dim(\mathcal{P}_0) = n - 2$, for all $p \in \mathcal{P}_0$, the fiber $\mathcal{L}^+_{\mathcal{P}_0}(p)$ consists of two directions. Let L be a smooth section of $\mathcal{L}^+_{\mathcal{P}_0}$, and for all $p \in \mathcal{P}_0$ let $\ell_p(t) = \exp_p\big(t \cdot L_p\big)$ be defined for all $t \in [0, 1]$.

Consider the following three assumptions:

(HP1) $\ell_p(t) \not\in \Gamma$ for all $t \in [0,1[$, and $\ell_p(1) \in \Gamma$;

(HP2) $\ell_p'(1) \not\in T_{\ell(1)}\Gamma$;

(HP3) $t = 1$ is not a $\mathcal{P}_0$-focal instant along ℓ_p.

Clearly, (HP1), (HP2) and (HP3) are open condition on $\mathcal{P}_0$, and therefore if they hold at some point p, they will hold in a neighborhood of p in $\mathcal{P}_0$. A general existence theory for light rays from $\mathcal{P}_0$ to Γ can be developed using global causality assumptions on the spacetime (M, g), as for instance [9, 12]. We stress however that no global assumptions are needed for the results of the present paper. A discussion on the nonfocality assumption (HP3) will be presented in Section 3.2. Loosely speaking, one can say that it holds *generically*, at least in stationary spacetimes, see [7].

Remark. Let us observe here that the above transversality and nonfocality assumptions only depend on $\mathcal{P}_0$ and Γ, and they could be stated without referring to particular choice of the section L of $\mathcal{L}^+_{\mathcal{P}_0}$. However, (HP1) makes L a somewhat *special* section of this bundle as far as focality is concerned: an instant $t \in \,]0,1]$ is focal along ℓ_p if and only if there exists a Jacobi field J along ℓ_p satisfying $J(0) \in T_p\mathcal{P}_0 \setminus \{0\}$, $\frac{DJ}{dt}(0) = \nabla_{J(0)}L$ and $J(t) = 0$, see Section 3.2.

3. A FINITE DIMENSIONAL REDUCTION FOR THE FERMAT PRINCIPLE

The Fermat principle for extended light sources and extended receivers proved in [17] characterize the (future-pointing) lightlike geodesics starting orthogonally from $\mathcal{P}_0$ and arriving spatially orthogonally to Γ as those lightlike piecewise-smooth curves from $\mathcal{P}_0$ to Γ that extremize the arrival time functional.

3.1. **The Fermat principle.**

Using the setup described above, we will now give a simplified finite dimensional reduction of this variational principle. We first observe that, under assumptions (HP1) and (HP2), the function $\tau\colon \mathcal{P}_0 \to \mathbb{R}^+$ defined by:

$$\tau(p) = T\big(\ell_p(1)\big) \tag{3.1}$$

is smooth. This follows easily from the implicit function theorem, using the transversality assumption (HP2). Our aimed finite dimensional reduction of the Fermat principle has the following statement:

Theorem 3.1. *In the above setup, assume that (HP1), (HP2) and (HP3) hold at every point* $p \in \mathcal{P}_0$.

(1) *A point* $p \in \mathcal{P}_0$ *is critical for* τ *if and only if* ℓ_p *is* spatially orthogonal *to* Γ *at* $t = 1$, *i.e., if* $\ell_p'(1) \in (T_{\ell_p(1)}\Gamma_{\tau(p)})^\perp$.

(2) *A critical point* $p \in \mathcal{P}_0$ *for* τ *is nondegenerate if and only if the lightlike geodesic* ℓ_p *is nondegenerate as an orthogonal geodesic in* M *between the spacelike submanifolds* $\mathcal{P}_0$ *and* $\Gamma_{\tau(p)}$.

(3) *If* $p \in \mathcal{P}_0$ *is a critical point of* τ*, then the Morse index* $i_{Morse}(\tau, p)$ *of* τ *at* p *is equal to the concavity index* $i_{conc}(\mathcal{P}_0, \Gamma_{\tau(p)}; \ell_p)$*, which is given by the difference:*

$$i(\mathcal{P}_0, \Gamma_{\tau(p)}; \ell_p) - i(\mathcal{P}_0; \ell_p). \tag{3.2}$$

Proof. For the proof of (1), observe that, for $p \in \mathcal{P}_0$ and $\nu \in T_p\mathcal{P}_0$, choosing a smooth curve $]-\varepsilon, \varepsilon[\ \ni s \mapsto p_s \in \mathcal{P}_0$ with $p_0 = p$ and $p_0' = \nu$, the differential $d\tau(p)\nu$ is given by:

$$d\tau(p)\nu = \frac{d}{ds}\Big|_{s=0} T(\ell_{p_s}) = dT(\ell_p(1))J_\nu(1),$$

where $J_\nu(1)$ is the $\mathcal{P}_0$-Jacobi field along ℓ_p satisfying:[2]

$$J_\nu(0) = \nu, \quad \text{and} \quad \frac{DJ_\nu}{dt}(0) = \nabla_\nu L. \tag{3.3}$$

Thus, p is a critical point of τ if and only if the vector space $\mathcal{E}_p \subset T_{\ell_p(1)}\Gamma$ defined by:

$$\mathcal{E}_p := \{J_\nu(1) : \nu \in T_p\mathcal{P}_0\} \tag{3.4}$$

is contained in $T_{\ell_p(1)}\Gamma_{\tau(p)}$. By Gauss Lemma, $\ell_p'(1)$ is orthogonal to $\mathcal{E}_p$; moreover, since $\ell_p'(1)$ is lightlike and Γ is timelike, then $\dim\big(\ell_p'(1)^\perp \cap T_{\ell_p(1)}\Gamma\big) = n-2$. Since 1 is not a focal instant[3] along ℓ, then also $\dim(\mathcal{E}_p) = \dim(\mathcal{P}_0) = n-2 = \dim(\Gamma_{\tau_p})$. Thus, $\ell_p'(1)^\perp \cap T_{\ell_p(1)}\Gamma = \mathcal{E}_p$, and p is critical if and only if

$$\mathcal{E}_p = T_{\ell_p(1)}\Gamma_{\tau(p)}. \tag{3.5}$$

In other words, p is critical if and only if $\ell_p'(1) \in (T_{\ell_p(1)}\Gamma_{\tau(p)})^\perp$.

For the proof of the remaining two statements, we need a second variational formula for the function τ at a given critical point $p \in \mathcal{P}_0$. To this aim, let us choose a differentiable curve $]-\varepsilon, \varepsilon[\ \ni s \mapsto p_s \in \mathcal{P}_0$, with $p_0 = p$, and $p_0' = \nu \in T_p\mathcal{P}_0$. Define the following two-parameter map in M:

$$]-\varepsilon, \varepsilon[\ \times [0,1] \ni (s,t) \longmapsto \eta(s,t) \in M,$$

by:

$$\eta(s,t) = \exp_{p_s}(t \cdot L_{p_s}) = \ell_{p_s}(t), \tag{3.6}$$

and denote by $\frac{\partial\eta}{\partial t}$, $\frac{\partial\eta}{\partial s}$ the corresponding vector fields along η. We want to compute

$$d^2\tau(p_0)[\nu,\nu] = \frac{d^2}{ds^2}\Big|_{s=0}\tau(p_s) = \frac{\partial^2}{\partial s^2}\Big|_{s=0} T\big(\eta(s,1)\big). \tag{3.7}$$

Since $t \mapsto \eta(s,t)$ is a geodesic, then:

$$g\big(\tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial t}\big) = 0,$$

[2] Let us observe that a Jacobi field along ℓ_p satisfying (3.3) is automatically a $\mathcal{P}_0$-Jacobi field.

[3] We will show in Section 3.2 that the condition $\dim(\mathcal{E}_p) = 2$ is equivalent to the nonfocality of the instant $t = 1$.

from which it follows:

$$0 = \frac{1}{2}\frac{\partial}{\partial s} g\big(\tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial t}\big) = g\Big(\nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial t}\Big). \tag{3.8}$$

From the symmetry of the Levi-Civita connection we also have:

$$\nabla_{\frac{\partial\eta}{\partial s}} \frac{\partial\eta}{\partial t} = \nabla_{\frac{\partial\eta}{\partial t}} \frac{\partial\eta}{\partial s}, \tag{3.9}$$

which substituted in (3.8) gives:

$$g\Big(\nabla_{\frac{\partial\eta}{\partial t}} \tfrac{\partial\eta}{\partial s}, \tfrac{\partial\eta}{\partial t}\Big) = 0. \tag{3.10}$$

For $s \in \left]-\varepsilon, \varepsilon\right[$, let $U_s \in T_{\eta(s,1)}\Gamma$ be the vector defined by:

$$U_s = -\nabla T\big(\eta(s,1)\big)/g\big(\nabla T\big(\eta(s,1)\big), \nabla T\big(\eta(s,1)\big)\big),$$

where ∇T is the vector field along Γ given by the gradient of the time function T. Set

$$C_s = -g\big(\tfrac{\partial\eta}{\partial t}(s,1), U_s\big),$$

this is a positive constant for all s, because $\frac{\partial\eta}{\partial t}(s,1)$ and U_s are future-pointing causal vectors. Let us consider a decomposition of the vector $\frac{\partial\eta}{\partial s}(s,1)$ as:

$$\frac{\partial\eta}{\partial s}(s,1) = a_s U_s + \frac{\partial\eta}{\partial s}(s,1)^{\perp}, \tag{3.11}$$

where $g\big(\frac{\partial\eta}{\partial s}(s,1)^{\perp}, U_s\big) = 0$. A direct computation gives:

$$a_s = g\big(\tfrac{\partial\eta}{\partial s}(s,1), \nabla T\big(\eta(s,1)\big)\big) = \frac{\partial}{\partial s} T\big(\eta(s,1)\big). \tag{3.12}$$

Using (3.10), we obtain:

$$\begin{aligned} 0 &= \int_0^1 g\big(\nabla_{\frac{\partial\eta}{\partial t}} \tfrac{\partial\eta}{\partial s}, \tfrac{\partial\eta}{\partial t}\big)\,dt = g\big(\tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial s}\big)\big|_{t=0}^{t=1} \\ &= g\big(\tfrac{\partial\eta}{\partial t}(s,1), \tfrac{\partial\eta}{\partial s}(s,1)\big) - g\big(\tfrac{\partial\eta}{\partial t}(s,0), \tfrac{\partial\eta}{\partial s}(s,0)\big) \\ &\overset{\text{by (3.11)}}{=} -C_s \cdot a_s + g\big(\tfrac{\partial\eta}{\partial t}(s,1), \tfrac{\partial\eta}{\partial s}(s,1)^{\perp}\big) - g\big(\tfrac{\partial\eta}{\partial t}(s,0), \tfrac{\partial\eta}{\partial s}(s,0)\big) \\ &\overset{\text{by (3.12)}}{=} -C_s \cdot \frac{\partial}{\partial s} T\big(\eta(s,1)\big) + g\big(\tfrac{\partial\eta}{\partial t}(s,1), \tfrac{\partial\eta}{\partial s}(s,1)^{\perp}\big) - g\big(\tfrac{\partial\eta}{\partial t}(s,0), \tfrac{\partial\eta}{\partial s}(s,0)\big), \end{aligned} \tag{3.13}$$

which gives:

$$C_s \cdot \frac{\partial}{\partial s} T\big(\eta(s,1)\big) = g\big(\tfrac{\partial\eta}{\partial t}(s,1), \tfrac{\partial\eta}{\partial s}(s,1)^{\perp}\big) - g\big(\tfrac{\partial\eta}{\partial t}(s,0), \tfrac{\partial\eta}{\partial s}(s,0)\big). \tag{3.14}$$

Keeping in mind that $\frac{\partial}{\partial s}\big|_{s=0} T\big(\eta(s,1)\big) = 0$, differentiating (3.14) gives:

$$C_0 \frac{\partial^2}{\partial s^2}\Big|_{s=0} T\big(\eta(s,1)\big) \tag{3.15}$$
$$= \Big[g\big(\nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial s}^{\perp}\big) + g\big(\tfrac{\partial\eta}{\partial t}, \nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial s}^{\perp}\big)\Big]\Big|_{s=0,\ t=1}$$
$$- \Big[g\big(\nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial t}, \tfrac{\partial\eta}{\partial s}\big) + g\big(\tfrac{\partial\eta}{\partial t}, \nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial s}\big)\Big]\Big|_{s=0,\ t=0}$$
$$\overset{\text{by (3.9)}}{=} \Big[g\big(\nabla_{\frac{\partial\eta}{\partial t}} \tfrac{\partial\eta}{\partial s}, \tfrac{\partial\eta}{\partial s}^{\perp}\big) + g\big(\tfrac{\partial\eta}{\partial t}, \nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial s}^{\perp}\big)\Big]\Big|_{s=0,\ t=1}$$
$$- \Big[g\big(\nabla_{\frac{\partial\eta}{\partial t}} \tfrac{\partial\eta}{\partial s}, \tfrac{\partial\eta}{\partial s}\big) + g\big(\tfrac{\partial\eta}{\partial t}, \nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial s}\big)\Big]\Big|_{s=0,\ t=0}.$$

From (3.6), we have:

$$\frac{\partial\eta}{\partial t}(0,0) = \ell_p'(0), \quad \frac{\partial\eta}{\partial s}(0,0) = v, \quad \frac{\partial\eta}{\partial t}(0,1) = \ell_p'(1),$$

and

$$\frac{\partial\eta}{\partial s}(0,1)^{\perp} \overset{\text{by (3.5)}}{=} \frac{\partial\eta}{\partial s}(0,1) = J_v(1), \tag{3.16}$$

where[4] J_v is the $\mathcal{P}_0$-Jacobi field along ℓ_p defined by the initial conditions (3.3). Substituting in (3.15) and using the definition of the second fundamental form, we obtain:

$$C_0 \cdot d^2\tau(p)[v,v] = g\big(\tfrac{DJ_v}{dt}(1), J_v(1)\big) + S^{\tau(p)}_{\ell_p'(1)}\big(J_v(1), J_v(1)\big), \tag{3.17}$$

where $S^{\tau(p)}_{\ell_p'(1)}$ is the second fundamental form of the spacelike submanifold $\Gamma_{\tau(p)}$ in the orthogonal direction $\ell_p'(1)$. Note that:

$$\Big[g\big(\nabla_{\frac{\partial\eta}{\partial t}} \tfrac{\partial\eta}{\partial s}, \tfrac{\partial\eta}{\partial s}\big) + g\big(\tfrac{\partial\eta}{\partial t}, \nabla_{\frac{\partial\eta}{\partial s}} \tfrac{\partial\eta}{\partial s}\big)\Big]\Big|_{s=0,\ t=0}$$
$$= g\big(\tfrac{DJ_v}{dt}(0), J_v(0)\big) + S^{\mathcal{P}_0}_{\ell_p'(0)}\big(J_v(0), J_v(0)\big) = 0$$

because J_v is a $\mathcal{P}_0$-Jacobi field.

Summarizing, formula (3.5) proves that the map $v \mapsto J_v$ gives an isomorphism between $T_p\mathcal{P}_0$ and the space $\mathbb{J}(\mathcal{P}_0;\ell_p)^{\Gamma_{\tau(p)}}$ consisting of $\mathcal{P}_0$-Jacobi fields J along ℓ_p such that $J(1) \in T_{\ell_p(1)}\Gamma_{\tau(p)}$ (see Section 2.2.3). Moreover, formula (3.17) shows that, using this isomorphism, the second variation of τ at a critical point p is identified with the concavity index form $\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_{\tau(p)}}$ of $(\mathcal{P}_0, \Gamma_{\tau(p)}; \ell_p)$, which was defined in (2.4). From these facts, we can draw the desired conclusions.

First, a vector $v \in T_p\mathcal{P}_0$ is in the kernel of $d^2\tau(p)$ if and only if the corresponding $\mathcal{P}_0$-Jacobi field J_v is in the kernel of the concavity index form $\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_{\tau(p)}}$, i.e.,

[4] Observe that equality (3.16) holds in fact also when assumption (HP3) is not satisfied, i.e., when (3.5) does not necessarily hold. Namely, for (3.16) it suffices to have the inclusion $\mathcal{E}_p \subset T_{\ell_p(1)}\Gamma_{\tau(p)}$, which holds for all critical points of τ, regardless of the nonfocality assumption.

if and only if J_ν is a $(\mathcal{P}_0, \Gamma_{\tau(p)})$-Jacobi field along ℓ_p. Thus, p is a nondegenerate critical point of τ if and only if ℓ_p is a nondegenerate orthogonal geodesic between $\mathcal{P}_0$ and $\Gamma_{\tau(p)}$, which proves part (2) of the statement.

Second, the Morse index of τ at p is equal to the index of the concavity index form $\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_{\tau(p)}}$. By the Morse index theorem, see formula (2.6), this index is given by the difference $i(\mathcal{P}_0, \Gamma_{\tau(p)}; \ell_p) - i(\mathcal{P}_0; \ell_p)$, proving (3.2). □

3.2. **Focusing of orthogonal light rays.** While assumptions (HP1) and (HP2) are certainly essential for the statement of Theorem 3.1, the reader may wonder whether assumption (HP3) is really necessary for the conclusion. Let us observe here that such assumption has been used in order to obtain that p is critical for τ if and only if equality (3.5) holds. More generally, without the assumption (HP3), one would deduce that p is critical if and only if $\mathcal{E}_p \subset T_{\ell_p(1)}\Gamma_{\tau(p)}$; recall that the space $\mathcal{E}_p$ was defined in (3.4).

Observe that Gauss Lemma only guarantees that $\ell_p'(1)$ is orthogonal to $\mathcal{E}_p$, so that orthogonality to the whole space $T_{\ell_p(1)}\Gamma_{\tau(p)}$ may fail when $\mathcal{E}_p \subsetneq T_{\ell_p(1)}\Gamma_{\tau(p)}$. Indeed, in Section 5 we will show an explicit example where (HP3) is not satisfied, and $\mathcal{E}_p \subsetneq T_{\ell_p(1)}\Gamma_{\tau(p)}$ at some critical point p, and for which the corresponding geodesic ℓ_p does *not* arrive orthogonally to $\Gamma_{\tau(p)}$.

Let us observe here that this phenomenon can only occur at *degenerate* critical points of the arrival time function.

Lemma 3.2. *Assume that (HP1) and (HP2) hold at p, and that p is a nondegenerate critical point of τ. Then, $\mathcal{E}_p = T_{\ell_p(1)}\Gamma_{\tau(p)}$, and therefore ℓ_p arrives orthogonally onto $\Gamma_{\tau(p)}$.*

Proof. The crucial observation is that the second variational formula (3.17) holds for all critical points of the function τ, regardless of the nonfocality assumption. This has been observed in footnote 4 on page 10. The equality $\mathcal{E}_p = T_{\ell_p(1)}\Gamma_{\tau(p)}$ follows once we show that $J_\nu(1) \neq 0$ for all $\nu \in T_p\mathcal{P}_0 \setminus \{0\}$. But this follows readily from the nondegeneracy assumption and the observation that, by formula (3.17), $\mathrm{Ker}\big(d^2\tau(p)\big) = \big\{\nu \in T_p\mathcal{P}_0 : J_\nu(1) = 0\big\}$. □

One could then wonder whether some assumption weaker than (HP3) implies the equality $\mathcal{E}_p = T_{\ell_p(1)}\Gamma_{\tau(p)}$ at every critical point p. Note that $\mathcal{E}_p \neq T_{\ell_p(1)}\Gamma_{\tau(p)}$ when there exists $\nu \in T_p\mathcal{P}_0 \setminus \{0\}$ such that $J_\nu(1) = 0$, and this occurs if and only if $\ell_p(1)$ is a *focusing point* for a family of lightlike geodesics starting orthogonally from $\mathcal{P}_0$. More precisely, $J_\nu(1) = 0$ if and only if there exists a smooth curve $]-\varepsilon, \varepsilon[\ni s \mapsto p_s \in \mathcal{P}_0$, with $p_0 = p$ and $p_0' = \nu$, such that:

$$\frac{d}{ds}\Big|_{s=0} \ell_{p_s}(1) = 0,$$

recalling that ℓ_{p_s} is the geodesic defined in (3.6).

However, the usual focality property is equivalent to the focusing of geodesics starting orthogonally from $\mathcal{P}_0$ having *arbitrary* causal character. Thus, it is not clear in principle whether one could have a situation where $\ell_p(1)$ is a $\mathcal{P}_0$-focal point along ℓ_p, and yet $\mathcal{E}_p = T_{\ell_p(1)}\Gamma_{\tau(p)}$. Next Lemma establishes that this is never the case.

For $p \in \mathcal{P}_0$, set:

$$\mathbb{J}_p^{\mathcal{P}_0} = \big\{\mathcal{P}_0\text{-Jacobi fields } J \text{ along } \ell_p \text{ with } J(1) = 0\big\}; \tag{3.18}$$

clearly, $\ell_p(1)$ is $\mathcal{P}_0$-focal iff $\mathbb{J}_p^{\mathcal{P}_0} \neq \{0\}$.

Proposition 3.3. *Assume that (HP1) and (HP2) hold at* $p \in \mathcal{P}_0$. *Then, the following conditions are equivalent:*

(a) $\ell_p(1)$ *is a* $\mathcal{P}_0$*-focal point along* ℓ_p*;*
(b) $\dim(\mathcal{E}_p) < n - 2$*;*

Proof. We have seen in the proof of Theorem 3.1 that (b) implies (a).

Assume now that (a) holds, and let $J \in \mathbb{J}_p^{\mathcal{P}_0}$ be a nontrivial $\mathcal{P}_0$-Jacobi field along ℓ_p satisfying $J(1) = 0$; let us prove that

$$\frac{DJ}{dt}(0) = \nabla_{J(0)}L + \alpha \cdot L_p \tag{3.19}$$

for some $\alpha \in \mathbb{R}$. By definition of $\mathcal{P}_0$-Jacobi field, one has:

$$\tfrac{DJ}{dt}(0) + S_{L_p}^{\mathcal{P}_0}\big(J(0)\big) = \tfrac{DJ}{dt}(0) - \nabla_{J(0)}L \in (T_p\mathcal{P}_0)^{\perp}. \tag{3.20}$$

Since both $\frac{DJ}{dt}(0)$ and $\nabla_{J(0)}$ are orthogonal to L_p, from (3.20) we get:

$$\tfrac{DJ}{dt}(0) + S_{L_p}^{\mathcal{P}_0}\big(J(0)\big) = \tfrac{DJ}{dt}(0) - \nabla_{J(0)}L \in (T_p\mathcal{P}_0)^{\perp} \cap L_p^{\perp}. \tag{3.21}$$

Since $\dim\big((T_p\mathcal{P}_0)^{\perp}\big) = 2$, then $(T_p\mathcal{P}_0)^{\perp} \cap L_p^{\perp} = \mathbb{R} \cdot L_p$, which proves (3.19).

Set $v = J(0)$; we now claim that $v \neq 0$. For, otherwise it would be $\alpha \neq 0$, and J would be the variational vector field associated to the variation of ℓ_p by orthogonal geodesics $(\ell_s)_{s \in]-\varepsilon,\varepsilon[}$:

$$\ell_s(t) = \ell_p\big((1+\alpha s)t\big), \quad t \in [0,1].$$

Observe indeed that $\ell_0 = \ell_p$, $\frac{d}{ds}\big|_{s=0}\ell_s(0) = 0$, and:

$$\frac{D}{ds}\Big|_{s=0}\ell_s'(0) = \frac{d}{ds}\Big|_{s=0}(1+\alpha s)\ell_p'(0) = \alpha \cdot \ell_p'(0) = \alpha \cdot L_p.$$

Hence, $J = \frac{d}{ds}\big|_{s=0}\ell_s$. However, for such a field J, it must be:

$$J(1) = \frac{d}{ds}\Big|_{s=0}\ell_s(1) = \frac{d}{ds}\Big|_{s=0}\ell_p(1+\alpha s) = \alpha \cdot \ell_p'(1) \neq 0,$$

i.e., J cannot belong to $\mathbb{J}_p^{\mathcal{P}_0}$. This proves the claim that $v \neq 0$.

Now, we claim that since $J(1) = 0$, then it must be $\alpha = 0$, i.e., J is equal to the $\mathcal{P}_0$-Jacobi field J_v defined by the initial conditions (3.3). Observe indeed that the field $J_\alpha(t) = \alpha \cdot \ell_p'(t)$ is a $\mathcal{P}_0$-Jacobi field along ℓ_p, and that $J = J_v + J_\alpha$. However, if $\alpha \neq 0$, then $J_\alpha(1) = \alpha\ell_p'(1) \neq 0$. On the other hand, $J_v(1) \in \mathcal{E}_p$, which is a spacelike subspace of $T_{\ell_p(1)}\Gamma$ and it does not contain nontrivial multiples of $\ell_p'(1)$. It follows that if $\alpha \neq 0$, then $J(1)$ cannot vanish, proving our claim.

The conclusion is that $J_v(1) = 0$, and therefore (b) holds. This concludes the proof. □

Using Lemma 3.2 and Proposition 3.3, we obtain readily the following:

Corollary 3.4. *Under assumption (HP1) and (HP2), if* p *is a nondegenerate critical point of* τ*, then* $\ell_p(1)$ *is not* $\mathcal{P}_0$*-focal along* ℓ_p*.* □

4. Application: bifurcation of light rays and the gravitational lensing effect

As an application of our finite dimensional reduction for the Fermat principle, we will discuss here a bifurcation result for light rays between an extended light source and an extended receiver. Having the path paved by the finite dimensional Fermat principle, the main results of this section will require no formal proof, as they follow immediately from standard variational bifurcation theory and the implicit function theorem. In view to a physical interpretation, the interesting case is when $n = 4$; nevertheless, this assumption will not be necessary in our theory. Bifurcation of light rays has been studied in reference [6]. We discuss here a different type of approach, which is more directly related to the multiple image effect and gravitational lensing in General Relativity.

4.1. **Worldline of the light source.** Let us consider the setup in Section 2.1, with the additional requirement that the spacelike submanifold $\mathcal{P}_0$ is an element of a continuous 1-parameter family of spacelike $(n-2)$-dimensional submanifolds:

$$[-\varepsilon, \varepsilon] \ni r \longmapsto \mathcal{P}_r, \tag{4.1}$$

for some $\varepsilon > 0$. More precisely, we will assume the existence of a 1-parameter group of local diffeomorphisms $\{\phi_r\}_r$ of M, which is continuous with respect to the C^2-topology, such that $\mathcal{P}_r = \phi_r(\mathcal{P}_0)$. The interesting case is when ϕ_r is the local flow of some future-pointing timelike vector field defined in a neighborhood of $\mathcal{P}_0$; in this situation, (4.1) is interpreted as the worldline of the extended light source $\mathcal{P}_0$. Note that we use a different symbol for the time parameter r, which in principle is distinct from the time τ measured on the target manifold Γ.

Up to considering a sufficiently small ε, we can assume that the line bundle $\mathcal{L}^+_{\mathcal{P}_0}$ can be continuously extended to an orthogonal lightlike line bundle $\mathcal{L}^+_P$ on

$$P = \bigcup_{r\in[-\varepsilon,\varepsilon]} \big(\mathcal{P}_r \times \{r\}\big).$$

Thus, for all $(p, r) \in \mathcal{P}_0 \times [-\varepsilon, \varepsilon]$, we have a future-pointing lightlike direction at the point $\phi_r(p)$ which is orthogonal to the spacelike submanifold $\mathcal{P}_r$, and which depends continuously on p and r.

Let us now assume that assumptions (HP1) and (HP2) hold on $\mathcal{P}_0$, and let us observe that these conditions are stable by sufficiently C^1-small perturbations of the manifold $\mathcal{P}_0$. This means that we have the existence of a continuous section L of $\mathcal{L}^+_P$, such that $L_{p,r}$ is smooth in p for all fixed r, and such that, setting:

$$\ell_{p,r}(t) = \exp_{\phi_r(p)}(L_{p,r} \cdot t)$$

the following hold, $\forall r \in [-\epsilon, \epsilon]$:

$(\mathrm{HP1})_P$ $\ell_{p,r}(t) \not\in \Gamma$ for all $t \in [0, 1[$, and $\ell_{p,r}(1) \in \Gamma$;

$(\mathrm{HP2})_P$ $\ell'_{p,r}(1) \not\in T_{\ell(1)}\Gamma$.

We therefore have a map:

$$\tau \colon P \longrightarrow \mathbb{R}^+$$

defined by $\tau(p, r) = T\big(\ell_{p,r}(1)\big)$, which is smooth in the variable $p \in P$ and continuous in r. For $r \in [-\varepsilon, \varepsilon]$, we also set $\tau_r : \mathcal{P}_0 \to \mathbb{R}^+$:

$$\tau_r = \tau(\cdot, r).$$

A similar observation can be made on assumption (HP3): if it holds on $\mathcal{P}_0$, then for ε small enough:

$(\mathrm{HP3})_P$ $t = 1$ is not a (lightlike) $\mathcal{P}_r$-focal instant along $\ell_{p,r}$, for all $r \in [-\varepsilon, \varepsilon]$.

The above discussion can be summarized as follows. Denote by $\mathcal{O}_{P,\Gamma}$ the set of initial points in P of future pointing lightlike geodesics starting orthogonally to P and arriving spatially orthogonally to Γ, i.e.:

$$\mathcal{O}_{P,\Gamma} = \Big\{(p, r) \in P : \ell'_{p,r}(1) \in \big(T_{\ell_{p,r}(1)}\Gamma_{\tau_r(p)}\big)^{\perp}\Big\}. \tag{4.2}$$

Proposition 4.1. *Under assumptions* $(\mathrm{HP1})_P$, $(\mathrm{HP2})_P$ *and* $(\mathrm{HP3})_P$*:*

$$\mathcal{O}_{P,\Gamma} = \Big\{(p, r) \in P : \tfrac{\partial \tau}{\partial p}(p, r) = 0\Big\}.$$

Moreover, given $(p, r) \in \mathcal{O}_{P,\Gamma}$*, then* $\ell_{p,r}$ *is nondegenerate as an orthogonal geodesic between* $\mathcal{P}_r$ *and* $\Gamma_{\tau_r(p)}$ *if and only if the second derivative* $\frac{\partial^2\tau}{\partial p^2}(p.r)$ *is a nondegenerate bilinear form on* $T_p\mathcal{P}_0$. □

4.2. **The bifurcation setup.** We will now assume that the 1-parameter family of local diffeomorphisms $\{\phi_r\}_r$ depends smoothly on r, which makes the set P a smooth manifold, and $\mathcal{L}^+_P$ a smooth line bundle over P. Suppose that we are given a continuous curve $[-\varepsilon, \varepsilon] \ni r \mapsto p_r \in \mathcal{P}_0$ such that the lightlike geodesic

$$\ell_r \colon = \ell_{p_r, r}$$

arrives spatially orthogonally to Γ, i.e., such that $\ell'_r(1)$ is orthogonal to $\Gamma_{\tau_r(p_r)}$ for all r. This is interpreted as the fact that an image of the extended source is seen by the (extended) receiver along the time. An immediate application of the implicit function theorem to the equation $\frac{\partial\tau}{\partial p} = 0$ gives the following:

Proposition 4.2. *Given* $r_* \in]-\varepsilon, \varepsilon[$*, if* ℓ_{r_*} *is a nondegenerate orthogonal geodesic between* $\mathcal{P}_{r_*}$ *and* $\Gamma_{\tau_{r^*}(p_{r^*})}$*, then the map* $r \mapsto p_r$ *is smooth near* $r = r_*$*, and a sufficiently small neighborhood of* (p_{r_*}, r_*) *in* $\mathcal{O}_{P,\Gamma}$ *consists of points of the form* (p_r, r)*, with* r *near* r_*. □

As to the geometry of $\mathcal{O}_{P,\Gamma}$ near a degenerate geodesic, this can be studied by looking at the singular zeros of the map $(p, r) \mapsto \frac{\partial\tau}{\partial p}(p, r)$. In this situation, a basic question is about the existence of bifurcating branches of orthogonal lightlike geodesics that converge to a degenerate one. Let us give the following definition.

Definition 4.3. We say that $r_* \in [-\varepsilon, \varepsilon]$ is an *instant of bifurcation of images of* P *on* Γ if there exists a sequence $(p_n, r_n)_{n\in\mathbb{N}}$ in P, with $p_n \neq p_{r_n}$ for all n, such that:

(a) $\lim_{n\to\infty}(p_n, r_n) = (p_{r_*}, r_*)$;
(b) $(p_n, r_n) \in \mathcal{O}_{P,\Gamma}$, for all $n \in \mathbb{N}$.

As an immediate corollary of Proposition 4.2 and Definition 4.3, we have:

Corollary 4.4. *If* $r_* \in [-\varepsilon, \varepsilon]$ *is an instant of bifurcation of images of* P *on* Γ*, then* p_{r_*} *is a degenerate critical point of* τ_{r_*} *in* P. □

It is well known that degeneracy is a necessary condition for bifurcation, which is not sufficient in general. Sufficient conditions for bifurcation are given in terms of jump of topological invariants associated to critical points, such as the Morse index. Given $r \in [-\varepsilon, \varepsilon]$ and a critical point p of τ_r, we will denote by $i_{\mathrm{Morse}}(\tau_r; p)$ the Morse index of τ_r at p.

Proposition 4.5. *Assume that:*

(1) $p_{\pm\varepsilon}$ *is a nondegenerate critical point of* $\tau_{\pm\varepsilon}$ *on* $\mathcal{P}_0$*;*
(2) $i_{Morse}(\tau_{-\varepsilon}; p_{-\varepsilon}) \neq i_{Morse}(\tau_\varepsilon; p_\varepsilon)$.

Then, there exists an instant $r_* \in\,]-\varepsilon, \varepsilon[$ *of bifurcation of images of* P *on* Γ. □

We emphasize that, in the case of an extended source and an extended receiver, bifurcation of light rays is *not* related to focal points, see Example 2 in Section 5.

4.3. **Interchanging the role of P and Γ.** Let us now assume that also the worldline of the light source can be described as a stably causal Lorentzian hypersurface of M, where the sets $\mathcal{P}_r$ are the time slices of the time function defined on the disjoint union $\mathcal{P} = \bigcup_r \mathcal{P}_r$. In this situation, the role of $\mathcal{P}$ and Γ can be reversed, and one can consider past-pointing lightlike geodesics starting orthogonally from a time slice of Γ, and ending on $\mathcal{P}$. In this situation, let us assume that ℓ_0 is a lightlike geodesic with endpoints orthogonal to a time slice $\mathcal{P}_0$ of $\mathcal{P}$ at the initial point, and to a time slice Γ_0 of Γ at the final instant. Then, ℓ_0 is obtained from a critical point p_0 on $\mathcal{P}$ off the arrival time function τ^Γ on Γ, and also from a critical point γ_0 on Γ of the departure time $\tau^{\mathcal{P}}$. Let us also assume that the endpoint of ℓ_0 on Γ_0 is not $\mathcal{P}_0$-focal, and that the endpoint of ℓ_0 on $\mathcal{P}_0$ is not Γ_0-focal.

Proposition 4.6. *In the above situation,* p_0 *is a nondegenerate critical point of* τ^Γ *if and only if* ℓ_0 *is a nondegenerate critical point of* $\tau^{\mathcal{P}}$*. The difference of the Morse indices* $i_{Morse}(\tau^{\mathcal{P}}; \ell_0) - i_{Morse}(\tau^\Gamma, p_0)$ *is equal to the number of* $\mathcal{P}_0$*-focal instants minus the number of* Γ_0*-focal instants along* ℓ_0*:*

$$i_{Morse}(\tau^{\mathcal{P}}; \ell_0) - i_{Morse}(\tau^\Gamma; p_0) = i(\mathcal{P}_0; \ell_0) - i(\Gamma_0; \ell_0^-), \tag{4.3}$$

where ℓ_0^- *denotes the backwards reparametrization of* ℓ_0.

Proof. The first statement follows easily from the fact that the nondegeneracy of a critical point for both the arrival and the departure time, from Theorem 3.1, part (2) is equivalent to the nondegeneracy of ℓ_0 as an orthogonal geodesic between the two end submanifolds, which does not depend on the orientation chosen for ℓ_0. Formula (4.3) follows readily from Theorem 3.1, part (3), and formula (2.6), as $i(\mathcal{P}_0, \Gamma_0; \ell_0) = i(\Gamma_0, \mathcal{P}_0; \ell_0^-)$. □

Thus, as a critical point of the arrival time functional, the Morse index of a lightlike geodesic ℓ_0 and the Morse index of its backwards reparameterization ℓ_0^- will change if the number of $\mathcal{P}_0$-focal points along ℓ_0 is different from the number of Γ_0-focal points along ℓ_0^-. An explicit example of this situation will be given in Section 5.

4.4. **On a notion of stability for orthogonal light rays.** Using the variational characterization of Theorem 3.1, we can introduce the following notion of stability for light rays.

Definition 4.7. A (future-pointing) light ray $\ell\colon [0, 1] \to M$ which is orthogonal to the initial spacelike surface $\mathcal{P}_0$ and spatially orthogonal to the final timelike submanifold Γ is said to be *stable* if the point $\ell(0)$ is a local minimum of the arrival time function $\tau^{\mathcal{P}_0}$.

Equivalently, ℓ is stable if the Morse index of $\tau^{\mathcal{P}_0}$ at $p_0 = \ell(0)$ is equal to 0. Recalling part (3) of Theorem 3.1, we obtain immediately:

Corollary 4.8. *ℓ is stable if and only if the concavity index form $\mathcal{C}^{\ell}_{\mathcal{P}_0, \Gamma_{\tau(p_0)}}$ is positive semidefinite.* □

As explained above, when one interchanges the role of the initial and the final manifold, the stability of ℓ and its backwards reparameterization ℓ^- may change, depending on the number of focal points, see Figure 4.

5. Examples

In this section we will exhibit explicit examples to illustrate our results. For most of our constructions, we will use standard static Lorentz manifolds, in which case lightlike geodesics project onto Riemannian geodesics in the spatial part. More precisely, let us assume that $(\widetilde{M}, \widetilde{g})$ is a Riemannian manifold, with $\dim(\widetilde{M}) = n - 1 \geqslant 2$, and let us consider the manifold $M = \widetilde{M} \times \mathbb{R}$, endowed with the Lorentz metric $g = \widetilde{g} - d\tau^2$, where τ denotes the coordinate in $\mathbb{R}$, and time-oriented by the timelike Killing vector field $\frac{\partial}{\partial \tau}$. A curve $\ell = (x, \tau) : [a, b] \to M$ is a geodesic in (M, g) if and only if $x : [a, b] \to \widetilde{M}$ is a geodesic in $(\widetilde{M}, \widetilde{g})$ and $\tau : [a, b] \to \mathbb{R}$ is an affine map; in this situation, ℓ is lightlike (and future-pointing) when $\widetilde{g}(\dot{x}, \dot{x}) = \dot{\tau}^2$ (and $\dot{\tau} > 0$).

Assume that $\widetilde{\mathcal{P}_0} \subset \widetilde{M}$ is a hypersurface, that $\widetilde{\Gamma}_0$ is an $(n-2)$-dimensional manifold, and that $\phi_\tau : \widetilde{\Gamma}_0 \to \widetilde{M}$ is a smooth 1-parameter map of embeddings, with $\tau \in [-\varepsilon, \varepsilon]$, $\varepsilon > 0$; we identify $\widetilde{\Gamma}_0$ with $\phi_0(\widetilde{\Gamma}_0)$, and we set $\widetilde{\Gamma}_\tau = \phi_\tau(\widetilde{\Gamma}_0)$.

Let $x_0 : [0, L_0] \to \widetilde{M}$ be a unit speed geodesic in $(\widetilde{M}, \widetilde{g})$ with

- $x_0(0) \in \widetilde{\mathcal{P}_0}$,
- $\dot{x}(0) \in (T_{x(0)}\widetilde{\mathcal{P}_0})^{\perp}$,

- $x_0\big([0,L_0[\big)\cap\widetilde{\Gamma}_0=\emptyset$,
- $x_0(L_0)\in\widetilde{\Gamma}_0$, and $\dot{x}_0(L_0)\notin T_{x(L_0)}\widetilde{\Gamma}_0$.

Set $\Gamma=\bigcup_{\tau\in[-\varepsilon,\varepsilon]}\big(\widetilde{\Gamma}_\tau\times\{\tau+L_0\}\big)\subset M$, and $\mathcal{P}_0=\widetilde{\mathcal{P}_0}\times\{0\}\subset M$; clearly, $\mathcal{P}_0$ is a spacelike $(n-2)$-dimensional submanifold of M, and Γ is a timelike hypersurface of (M,g). The map $T:\widetilde{M}\times\mathbb{R}\to\mathbb{R}$ defined by:

$$T(x,\tau)=\tau+L_0$$

gives an everywhere defined smooth future pointing timelike function on (M,g), and the time slices of Γ are precisely the spacelike submanifolds $\Gamma_\tau=\widetilde{\Gamma}_\tau\times\{\tau+L_0\}$.

The curve $\ell_0:[0,1]\to M$, defined by $\ell_0(t)=\big(x_0(tL_0),tL_0\big)$ is a future pointing lightlike geodesic starting orthogonally from $\mathcal{P}_0$ and arriving transversally on Γ. By replacing $\mathcal{P}_0$ with a suitable neighborhood of $p_0=\big(x_0(0),0\big)$ in $\mathcal{P}_0$, we are in the setup described in Section 2.1, and assumptions (HP1) and (HP2) of Section 2.3 are satisfied. Moreover, it is easy to prove the following:

- the instant L_0 is $\widetilde{\mathcal{P}_0}$-focal along the geodesic x_0 in $(\widetilde{M},\widetilde{g})$ if and only if $\ell_p(1)$ is a lightlike $\mathcal{P}_0$-focal point along ℓ_0;
- ℓ_0 is spatially orthogonal to Γ (i.e., orthogonal to Γ_0) if and only if x_0 arrives orthogonally to $\widetilde{\Gamma}_0$.

It is also clear that, in the above situation, the *arrival time function* τ defined in (3.1) corresponds to the Riemannian $\widetilde{g}$-length of geodesics starting orthogonally from $\widetilde{\mathcal{P}_0}$ and arriving (transversally) onto $\widetilde{\Gamma}_\tau$. There is a one-to-one correspondence between $\mathcal{P}_0$-Jacobi field along ℓ_0 that are tangent at Γ_0 at $\ell_0(1)$ and $\widetilde{\mathcal{P}_0}$-Jacobi fields along x_0 that are tangent to $\widetilde{\Gamma}_0$ at $x_0(1)$. Using this correspondence, when ℓ_0 is spatially orthogonal to Γ, the concavity index form $\mathcal{C}_{\ell_0}^{\mathcal{P}_0,\Gamma_0}$ determined by ℓ_0 corresponds to the concavity index form $\mathcal{C}_{x_0}^{\widetilde{\mathcal{P}_0},\widetilde{\Gamma}_0}$ of the g_0-geodesic x_0. Thus, degeneracy and Morse index for a $(\mathcal{P}_0,\Gamma_0)$-orthogonal lightlike geodesic ℓ_0 in (M,g) are the same as degeneracy and Morse index for the $(\widetilde{\mathcal{P}_0},\widetilde{\Gamma}_0)$-orthogonal geodesic x_0 in $(\widetilde{M},\widetilde{g})$.

For the bifurcation setup, we will consider $\mathcal{P}_r=\mathcal{P}_0\times\{r\}$, $r\in[-\varepsilon,\varepsilon]$, so that $P=\bigcup_r\big(\mathcal{P}_0\times\{r\}\big)$; in this situation, lightlike geodesics that are spatially orthogonal to P and Γ correspond to geodesics in $(\widetilde{M},\widetilde{g})$ that are orthogonal to $\widetilde{\mathcal{P}_0}$ and to $\widetilde{\Gamma}_t$ at their endpoints for some t.

Keeping in mind this Riemannian setup, we will now describe our examples.

Example 1. This example shows that assumption (HP3) cannot be omitted in Theorem 3.1. Consider the case where $(\widetilde{M},\widetilde{g})$ is the Euclidean plane $\mathbb{R}^2$. The hypersurface $\widetilde{\mathcal{P}_0}$ is a neighborhood of the point $(0,1)$ in the circle centered at $(0,0)$ having radius 1, and $\widetilde{\Gamma}_0$ is a neighborhood of $(0,0)$ of the line $y=x$. The origin $(0,0)$ belongs to $\widetilde{\Gamma}_0$, and it is a $\widetilde{\mathcal{P}_0}$-focal point. All the segments starting orthogonally from $\widetilde{\mathcal{P}_0}$ arrive on $\widetilde{\Gamma}_0$ at $(0,0)$, so that the length function is constant on $\widetilde{\mathcal{P}_0}$, i.e.,

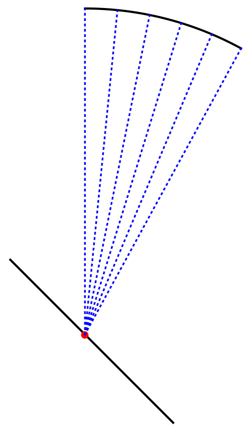

FIGURE 1. Critical points of the length functional do not correspond to orthogonal geodesics when (HP3) is not satisfied. In the picture, the segments starting orthogonally to the circular curve are rays that meet at the center of the circle. In this case, the length function is constant.

all the points of $\widetilde{\mathcal{P}_0}$ are critical for the distance function. However, none of these critical points correspond to a segment that arrive orthogonally to $\widetilde{\Gamma}_0$. See Figure 1.

A perturbation of this situation produces an example of an isolated (degenerate) critical point of the distance function on $\widetilde{\mathcal{P}_0}$ which does not correspond to a segment that arrives orthogonally to $\widetilde{\Gamma}_0$, see Figure 2 .

Example 2. This example shows that one can have bifurcation of spatially orthogonal light rays also when there are no focal points. Consider again the case where $(\widetilde{M}, \widetilde{g})$ is the Euclidean plane $\mathbb{R}^2$. The hypersurface $\widetilde{\mathcal{P}_0}$ is the circle centered at $(-2, 0)$ having radius 2; given $\tau \in [1, 3]$, the hypersurface $\widetilde{\Gamma}_\tau$ is a neighborhood of the point $(1, 0)$ in the circle with center in $(-\tau, 0)$ and radius equal to $\tau + 1$. For all τ, the segment $x_0(t) = (t, 0)$, $t \in [0, 1]$, is an orthogonal geodesic from $\widetilde{\mathcal{P}_0}$ to $\widetilde{\Gamma}_\tau$. For $\tau \in [1, 2[$, the curve x_0 realizes the minimum distance between $\widetilde{\mathcal{P}_0}$ and $\widetilde{\Gamma}_\tau$. For $\tau \in]2, 3]$, the curve x_0 gives a local maximum of the distance function, so that a jump of Morse index occurs at $\tau = 2$. As expected, bifurcation of orthogonal geodesics occurs at $\tau = 2$, when $\widetilde{\mathcal{P}_0}$ and $\widetilde{\Gamma}_\tau$ have the same center: all the segments starting orthogonally from $\widetilde{\mathcal{P}_0}$ arrive orthogonally to $\widetilde{\Gamma}_2$. Observe that x_0 does not contain neither $\widetilde{\mathcal{P}_0}$-focal instants nor $\widetilde{\Gamma}_\tau$-focal instants. See Figure 3. This is an example of supercritical pitchfork bifurcation.

Example 3. In Figure 4 there is an example where, interchanging the role of the initial and the final submanifold, the Morse index of the orthogonal geodesic changes. This is due to the presence of focal points, as explained in Section 4.3. Observe that on the left we have a stable critical point of the arrival time which is a geodesic

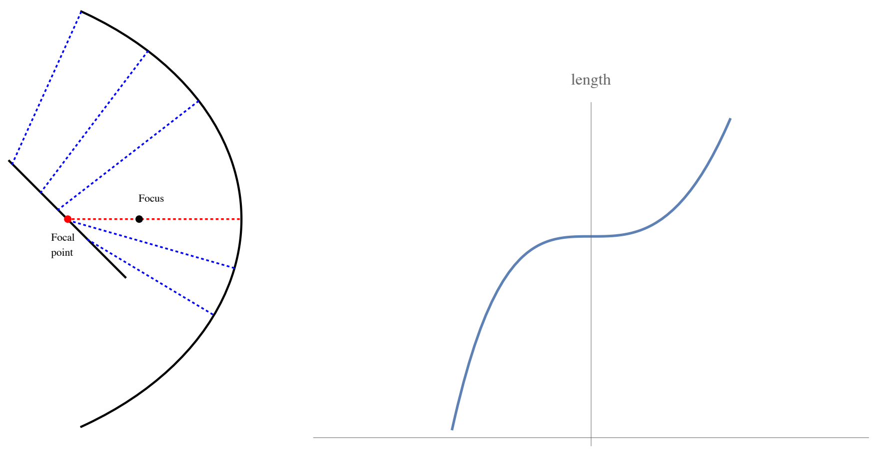

FIGURE 2. On the left, $\widetilde{\mathcal{P}_0}$ is the arc of an ellipse containing a vertex along the greater axis (horizontal), and $\widetilde{\Gamma}_0$ is a straight non-vertical segment through the closest focal point of the ellipse (this is not the focus of the ellipse!). Thus, assumption (HP3) is not satisfied. Segments that depart orthogonally from $\widetilde{\mathcal{P}_0}$ arrive transversally, but never orthogonally, to $\widetilde{\Gamma}_0$. In red, the segment along the axis, which gives a critical point of the length function, but it does not arrive orthogonally to $\widetilde{\Gamma}_0$. On the right, a graph of the distance function, with a degenerate critical point at $t = 0$ corresponding to the vertex of the ellipse. Note that, as in the example of Figure 1, we have here a degenerate critical point of length function, see Lemma 3.2.

that contains a focal point. On the right, an unstable critical point with no focal points.

Example 4. In the Euclidean plane, the hypersurface $\widetilde{\mathcal{P}_0}$ is (a neighborhood of the point $(-1,0)$ of) the circle centered at $(0,0)$ having radius 1. For $\tau \in [-\varepsilon, \varepsilon]$, the hypersurface $\widetilde{\Gamma}_\tau$ is the vertical line of equation $y = \tau$. For all τ, the horizontal segment $x_\tau(t) = \big(-1 + (\tau+1)t, 0\big)$, $t \in [0,1]$, is the unique geodesic starting orthogonally on $\mathcal{P}_0$ and ending orthogonally to $\widetilde{\Gamma}_\tau$. For $\tau = 0$, the final endpoint $x_0(1) = (0,0)$ is $\widetilde{\mathcal{P}_0}$-focal.

Let us observe that for $\tau < 0$, x_τ gives a maximum of the length function in the set of segments that start orthogonally from $\widetilde{\mathcal{P}_0}$ and terminate on $\widetilde{\Gamma}_\tau$. Thus, the Morse index of the corresponding arrival time function is equal to 1; this is also easily computed as the index of the concavity form, which in this case is defined on a 1-dimensional space and it is negative definite. On the other hand, for $\tau > 0$, x_τ realizes the minimum of the length function, and it gives Morse index equal to 0. Also in this case, a direct elementary computation shows that the concavity index form is positive definite. Thus, we have jump of the Morse index, but no bifurcation

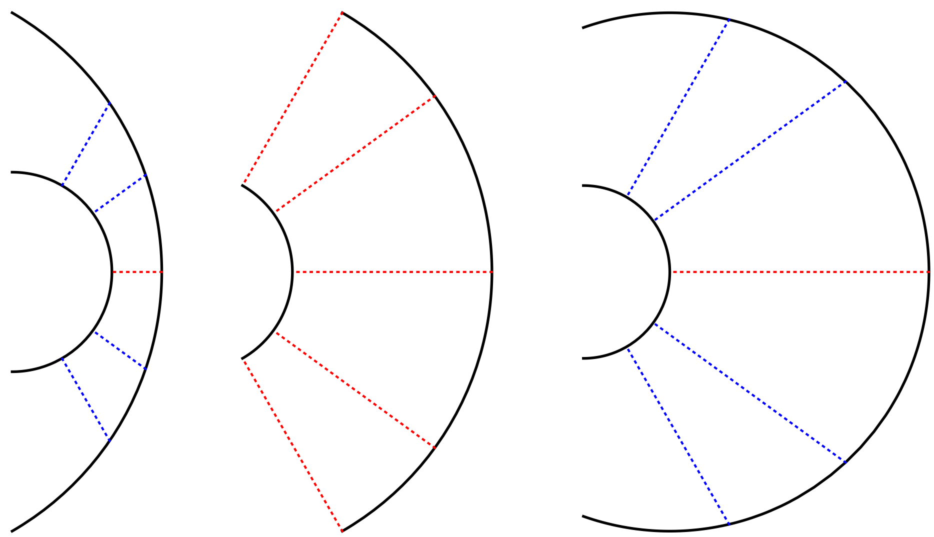

FIGURE 3. In black two circular curves, with segments starting orthogonally from the circle on the left. Bifurcation occurs when the center of the circles coincide, in which case there are infinitely many segments that arrive orthogonally to the right circle. This is an example of supercritical pitchfork bifurcation: on the left, the orthogonal geodesic is stable, and it becomes unstable after the bifurcation instant. The bifurcating branch consists of stable orthogonal geodesics.

of orthogonal geodesics. This is due to the fact that at $\tau = 0$ the assumption (HP3) is not satisfied. More precisely, one has bifurcation of critical points for the arrival time function, but they do not correspond to orthogonal geodesics (they are rays of the circle $\widetilde{\mathcal{P}}_0$ terminating at the center, see Figure 5).

We should also note that the Morse index of the geodesic x_τ as a critical point of the geodesic action functional in the space of all paths from $\widetilde{\mathcal{P}}_0$ to $\widetilde{\Gamma}_\tau$ has Morse index equal to 1 for all $\tau \neq 0$. Namely, by Theorem 2.1, this index is given by the sum of the index of the concavity index form and the number of $\widetilde{\mathcal{P}}_0$-focal points. Thus, we have an elementary geometric example that shows that degeneracy without jump of Morse index does not produce bifurcation in general.

Example 5. Let us consider the setup of Section 4.1. Besides bifurcation theory, other techniques from Singularity Theory can be employed to study the geometry of the set

$$\mathcal{O}_{P,\Gamma} = \left\{(p, r) \in \mathcal{P}_0 \times [-\varepsilon, \varepsilon] : \ell'_{p,r}(1) \in \left(T_{\ell_{p,r}(1)}\Gamma_{\tau_r(p)}\right)^{\perp}\right\}$$

near a degenerate orthogonal light ray. Recalling Proposition 4.1:

$$\mathcal{O}_{P,\Gamma} = \left\{(p, r) \in \mathcal{P}_0 \times [-\varepsilon, \varepsilon] : \tfrac{\partial \tau}{\partial p}(p, r) = 0\right\},$$

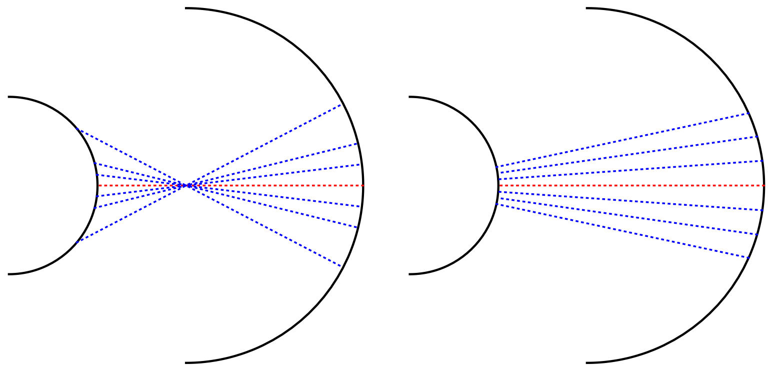

FIGURE 4. Interchanging the role of the initial and the final manifold, the same orthogonal geodesic has different Morse indices. On the left, the orthogonal segment in red corresponds to a minimum of the length functional on the space of segments starting orthogonally from the larger circle. On the right, the same geodesic corresponds to a maximum of the length functional on the space of segments that start orthogonally from the smaller circle.

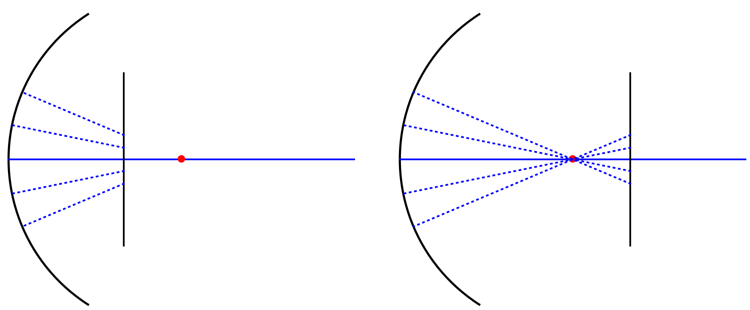

FIGURE 5. Jump of the Morse index that does not produce bifurcation. On the left, the orthogonal geodesic corresponds to a maximum of the length function. On the right, to a minimum.

where $\tau : \mathcal{P}_0 \times [-\varepsilon, \varepsilon] \to \mathbb{R}$ is the smooth function that represents the arrival time functional defined in the set of (future pointing) lightlike geodesics issuing orthogonally from P and arriving on Γ.

Let us assume that we are given a smooth curve $[-\varepsilon, \varepsilon] \ni r \mapsto p_r \in \mathcal{P}_0$ such that:

(a) $(p_r, r) \in \mathcal{O}_{P,\Gamma}$ for all r;
(b) the Hessian $\dfrac{\partial^2 \tau}{\partial p^2}(p_r, r)$ is degenerate only at $r = 0$.

In order to study the geometry of the set $\mathcal{O}_{P,\Gamma}$ near $(p_0, 0)$, one can use local coordinates on $\mathcal{P}_0$ around p_0, and consider the situation of a smooth one-parameter family of real valued maps $\tau_r = \tau(\cdot, r)$ defined on a neighborhood of $\mathbf{0}$ in $\mathbb{R}^{n-2}$.

Let us consider for simplicity the case $n = 4$, so that the maps τ_r are defined on an open neighborhood of $(0, 0)$ in $\mathbb{R}^2$, and the gradient $G_r = \nabla\tau_r$ is seen as a map from $\mathbb{R}^2$ to itself. The set $\mathcal{O}_{P,\Gamma}$ corresponds to the inverse image $G_r^{-1}(0, 0)$. Condition (b) above tells us that $(0, 0)$ is a singularity of G_0. Let us assume that the differential $dG_0(0, 0)$ (or equivalently, the Hessian $\frac{\partial^2\tau}{\partial p^2}(p_0, 0)$) is not identically 0, i.e., its kernel has dimension equal to 1. By a classical result of Whitney, see [24], in a generic situation the map G_0 is given, in suitable coordinates, in one of the following two forms:

- $\mathbb{R}^n \ni (\mathfrak{x}, \mathfrak{y}) \mapsto (\mathfrak{x}^2, \mathfrak{y}) \in \mathbb{R}^2$ (fold singularity);
- $\mathbb{R}^n \ni (\mathfrak{x}, \mathfrak{y}) \mapsto (\mathfrak{x}^3 - \mathfrak{x}\mathfrak{y}, \mathfrak{y}) \in \mathbb{R}^2$ (cusp singularity).

These local forms are stable by C^3-small perturbations, hence we can assume that, for $|r|$ small, also the maps G_r have the same form after a change of coordinates that depends on r. Such a change of coordinates carries $(0, 0)$ to some other point (a_r, b_r) in a neighborhood of $(0, 0)$; the two functions $r \mapsto a_r$ and $r \mapsto b_r$ are smooth, and they satisfy $a_0 = b_0 = 0$. Again, by assumption (b) above, for all $r \neq 0$, (a_r, b_r) does not belong to the singular set of the map G_r, which is:

- the single point $(0, 0)$ in the fold case;
- the parabola $\mathfrak{y} = 3\mathfrak{x}^2$ in the cusp case.

Assumption (a) tells us also that (a_r, b_r) belongs to the image of G_r for all r.

Set $H_{\text{fold}}(\mathfrak{x}, \mathfrak{y}) = (\mathfrak{x}^2, \mathfrak{y})$ and $H_{\text{cusp}}(\mathfrak{x}, \mathfrak{y}) = (\mathfrak{x}^3 - \mathfrak{x}\mathfrak{y}, \mathfrak{y})$. In the fold case, $G_r^{-1}(a_r, b_r) = H_{\text{fold}}^{-1}(a_r, b_r)$ consists of exactly two points $(\pm\sqrt{a_r}, b_r)$ when $r \neq 0$. One of them corresponds to the point p_r (the curve $r \mapsto p_r$ is the "trivial branch" of solutions of the equation $d\tau_r(p) = 0$). The other solution belongs to the bifurcation branch. This is a continuous path $r \mapsto q_r$, with $q_0 = p_0$, which is smooth for $r < 0$ and for $r > 0$, by the implicit function theorem, using assumption (b). Since $-\sqrt{a_r} < a_r < \sqrt{a_r}$ for $a_r > 0$ small, then the two smooth portions of the bifurcating branch lie on different sides of the trivial branch. A typical bifurcation picture in the fold case is illustrated in Figure 6.

In the cusp case, it is easy to see that the equation $H_{\text{cusp}}(\mathfrak{x}, \mathfrak{y}) = (a_r, b_r)$, with (a_r, b_r) near $(0, 0)$ admits one, two or three solutions near $(0, 0)$. More precisely, it admits three solutions when (a_r, b_r) is in the region $\mathcal{R}$ given by $\mathfrak{x}^2 < \frac{4}{27}\mathfrak{y}^3$ (shaded region in Figure 7), two solutions when $a_r^2 = \frac{4}{27}b_r^3$, and one solution when (a_r, b_r) is in the region $\mathfrak{x}^2 > \frac{4}{27}\mathfrak{y}^3$. Thus, in this situation bifurcation may or may not occur, with possibly continuous or discrete bifurcating branch, depending on the intersections of the curve $r \mapsto (a_r, b_r)$ with the region $\mathcal{R}$. See Figure 7.

A similar analysis can be carried out also when $n > 4$, using canonical forms of singularities for maps on manifolds of arbitrary dimensions, see [10, 14].

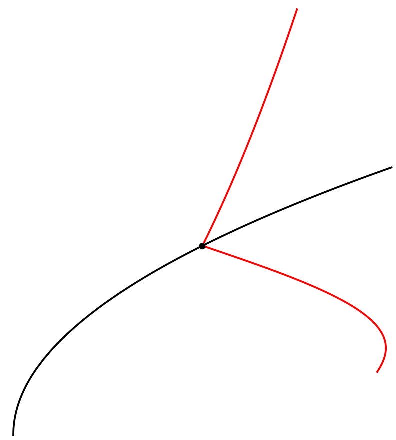

FIGURE 6. In black the trivial branch of solutions for the equation $d\tau_r(p) = 0$ near the bifurcation instant. In red, the two smooth portions of the bifurcating branch.

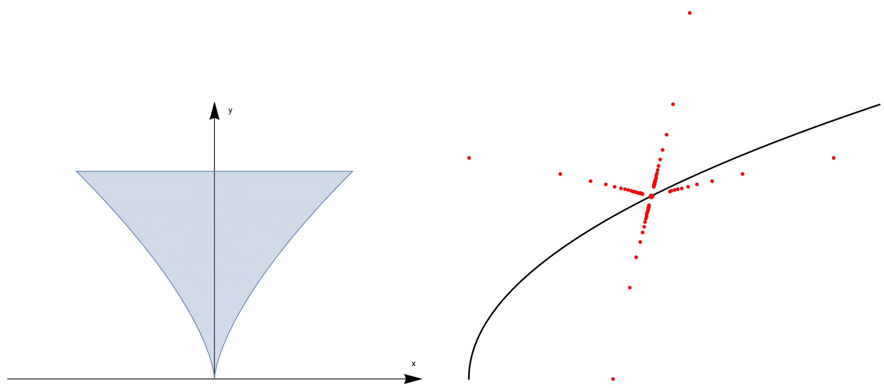

FIGURE 7. On the left, shaded, the region $x^2 < \frac{4}{27}y^3$; when (a_r, b_r) is in this region, the equation $(x^3 - xy, y) = (a_r, b_r)$ has three solutions. When (a_r, b_r) is on the boundary of this region, there are two solutions. Elsewhere, there is a unique solution. Correspondingly, bifurcation may or may not occur, and the bifurcation branch can be either discrete or continuous. On the right, a possible picture: in red, a discrete bifurcation branch consisting of two or four sequences of solutions converging to the bifurcation point on the trivial branch.

6. AN APPLICATION: NULL CONCAVE OPTICAL DOMAINS

6.1. **The physical setting.** We will now discuss an example with a sensible physical interpretation as further illustration of our results. Consider a celestial body B, say a star going supernova, whose surface suddenly gives off an outburst of light or gravitational radiation at an instant $t = 0$ as measured on Earth. The pulse travels into our solar system, and a portion of it is captured by a network of satellite detectors orbiting around the Earth. For simplicity of the description we may assume that the detectors are so densely distributed that we may effectively imagine they span the surface of a round sphere D concentric with Earth, but with some larger radius.

6.2. The mathematical model.

Definition 6.1. An *optical domain* in a spacetime (M^{n+1}, g) $(n \geqslant 3)$ is a pair (T, Ω) such that

- (a) $\Omega \subset M$ is a connected open set whose boundary $\partial\Omega$ is a smooth timelike hypersurface in M, with exactly two connected components which we denote by Σ and Γ;
- (b) $T : \partial\Omega \to \mathbb{R}$ is a smooth time function with past pointing timelike gradient, that can be assumed to be surjective.

The set $\Sigma_0 := \Sigma \cap T^{-1}\{0\}$ is the *emitting body* for (T, Ω).

Fix an optical domain (T, Ω) in (M, g). We can view $(\overline{\Omega}, g|_{\overline{\Omega}})$ as a spacetime with smooth boundary and effectively "forget" the "outer" region $M \setminus \overline{\Omega}$. Here, the two boundary components Σ and Γ purport to describe the spacetime history of B and of D, respectively.

It might well happen that only part of the emitted pulse reaches D, and we can consider an open submanifold $\mathcal{P}_0$ of Σ_0 describing the "visible" region of B for the detector, which may or may not be equal to Σ_0. If the wavelength of the pulse radiation is very small compared with the distance traveled to the detector, we may adopt a geometric optical approximation, and describe the pulse by a family $\mathcal{F}$ of normal future pointing affinely parametrized null geodesics emanating from $\mathcal{P}_0$ and arriving at Γ.

More precisely, we adopt the following definition.

Definition 6.2. Let (T, Ω) be an optical domain for the spacetime (M, g). Let N be the unit spacelike inward pointing normal vector field on $\partial\Omega$. We define a section L_0 of the future pointing lightlike normal bundle $\mathcal{L}^+_{\Sigma_0}$ of the emitting body Σ_0 by

$$L_0 := -\frac{\nabla^\Sigma T}{|\nabla^\Sigma T|}\Big|_{\Sigma_0} + N\big|_{\Sigma_0}, \tag{6.1}$$

where ∇^Σ denotes the gradient with respect to the induced metric on Σ, and $|\nabla^\Sigma T| = \big[-g(\nabla^\Sigma T, \nabla^\Sigma T)\big]^{\frac{1}{2}}$.

Let $\Sigma_0 \ni p \mapsto L_p \in [0, +\infty[$ be a smooth function. The collection $\mathcal{F}$ of future pointing null geodesics $\ell_p : [0, 1] \to M$ given by $\ell_p(t) = \exp_p(t \cdot L_p)$ for all $t \in [0, 1]$, and all $p \in \Sigma_0$, is called a *pulse* for the optical domain. The set

$$\mathcal{P}_0 := \big\{p \in \Sigma_0 \;\big|\; \ell_p \text{ satisfies (HP1)—(HP3) (Section 2.3)}\big\}$$

is called *the visible region* of the emitting body Σ_0. Clearly, $\mathcal{P}_0$ is a (possibly empty) open subset of Σ_0.

Definition (6.1) is meant to describe the fact that the pulse moves *outwardly* from the body B. This translates into certain *null convexity* assumptions on Σ_0. The convergence/divergence of the family $\mathcal{F}$ of null geodesics is described by the (trace of the) second fundamental form S^{Σ_0} of Σ_0 in the orthogonal direction L: the fact that the family is *diverging* is indicated by the condition:

$$\mathrm{tr}(S^{\Sigma_0}) \leqslant 0. \tag{6.2}$$

Analogously, for each level set Γ_t in Γ we define a section L_t of the future pointing lightlike normal bundle $\mathcal{L}^+_{\Gamma_t}$ of Γ_t pointing to the outside of Ω by

$$L_t := -\frac{\nabla^\Gamma T_\Gamma}{|\nabla^\Gamma T_\Gamma}\big|_{\Gamma_t} - N\big|_{\Gamma_t}. \tag{6.3}$$

Future pointing lightlike inbound geodesics normal to Γ_t arrive parallel to L_t. We assume they converge as they arrive, and hence we assume that the trace of the second fundamental form S^{Γ_t} of Γ_t with respect to L_t is positive:

$$\mathrm{tr}(S^{\Gamma_t}) > 0. \tag{6.4}$$

We embody these concepts into an appropriate definition.

Definition 6.3. An optical domain (T, Ω) is said to be *null concave* if

(i) all the null geodesics in the pulse $\mathcal{F}$ starting at the visible region $\mathcal{P}_0$ of the emitting body Σ_0 are entirely contained in $\overline{\Omega}$;
(ii) (6.2) holds on the emitting body Σ_0 and (6.4) holds on Γ_t for each $t \in \mathbb{R}$;
(iii) for every lightlike vector $v \in T\overline{\Omega}$ and every $w \in T\overline{\Omega}$ orthogonal to v

$$g\big(R(v, w)v, w\big) \geqslant 0. \tag{6.5}$$

Remark. Inequality (6.5) may be interpreted as a non-convergence condition on null geodesics. Geometrically, this will help avoiding premature focal points, but how realistic is that from a physical perspective? It obviously holds in a flat spacetime. Thus, in a region of our Universe away from large curvature regions such as the vicinity black holes, such an assumption would not be too restrictive. Globally, our Universe is usually described by a Robertson–Walker spacetime $(I \times S, -dt^2 \oplus f^2 \cdot h)$ with $I \subseteq \mathbb{R}$ an open interval, (S, h) a Riemannian manifold of constant curvature k and $f : I \to \,]0, +\infty[$ is a smooth positive function. We can normalize k to be 0, 1 or -1. In such a spacetime, any future-directed null vector is a multiple of

$$v = \partial_t + \frac{z}{f},$$

where $z \in TS$ is a unit vector with respect to h. For any $w \in TS$ orthogonal to v, we then have

$$g\big(R(v, w)v, w\big) = -\frac{f''}{f} + \left(\frac{f'}{f}\right)^2 + \frac{k}{f^2}.$$

According to current observations, $k = 0$ and the universe is expanding with positive acceleration, so $f' > 0$ and $f'' > 0$. However, the acceleration is very small, so the second term in the previous equation dominates, and condition (6.5) is satisfied.

We are ready to state our results.

Proposition 6.4. *Let* (T, Ω) *be a null concave optical domain in a spacetime, with emitting body* Σ_0 *and visible region* $\mathcal{P}_0 \subset \Sigma_0$*. Let* $p \in \mathcal{P}_0$ *and assume that* ℓ_p *arrives orthogonally to some* Γ_t*. Then, the concavity index form* $\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_t}$ *(see (2.4)) is* not *negative semidefinite. In particular, none of the null geodesics of the pulse* $\mathcal{F}$ *emanating from* $\mathcal{P}_0$ *maximize the arrival time.*

Proof. Let $\{e_1, \ldots, e_{n-1}\}$ be an orthonormal basis of $T_p\Sigma_0$, and consider:

$$\text{(6.6)}\quad \mathrm{tr}\big(\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_t}\big) = \sum_{s=1}^{n-1} \mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_t}(e_s, e_s) \overset{\text{by (2.4)}}{=} \sum_{s=1}^{n-1} g\big(J'_{e_s}(1), J_{e_s}(1)\big) + \mathrm{tr}(S^{\Gamma_t}) \overset{\text{by (6.4)}}{>} \sum_{s=1}^{n-1} g\big(J'_{e_s}(1), J_{e_s}(1)\big).$$

Note that, since J_{e_s} is a Σ_0-Jacobi field, then $g\big(J'_{e_s}(0), w\big) + S^{\Sigma_0}\big(J_{e_s}(0), w\big) = 0$ for all $w \in T_{p_0}\Sigma_0$, and in particular:

$$\text{(6.7)}\qquad g\big(J'_{e_s}(0), J_{e_s}(0)\big) + S^{\Sigma_0}\big(J_{e_s}(0), J_{e_s}(0)\big) = 0, \quad \forall\, s = 1, \ldots, n-1.$$

Now, set:

$$h(t) = \sum_{s=1}^{n-1} g\big(J'_{e_s}(t), J_{e_s}(t)\big);$$

the derivative $h'(t)$ is computed as follows:

$$\text{(6.8)}\quad h'(t) = \sum_{s=1}^{n-1} g\big(J'_{e_s}(t), J'_{e_s}(t)\big) + \sum_{s=1}^{n-1} g\big(J''_{e_s}(t), J_{e_s}(t)\big) = \sum_{s=1}^{n-1} g\big(J'_{e_s}(t), J'_{e_s}(t)\big) + \sum_{s=1}^{n-1} g\big(R(t)J_{e_s}(t), J_{e_s}(t)\big) \overset{\text{by (6.5)}}{\geqslant} 0.$$

By $R(t)$ we mean the endomorphism of $\ell'_p(t)^\perp$ given by:

$$R(t)w = g\big(\ell'_p(t), w\big)\ell'_p(t).$$

For the last inequality in (6.8) above observe that $J_{e_s}(t)$ is orthogonal to the lightlike vector $\ell'_p(t)$, and in particular $g\big(R(t)J_{e_s}(t), J_{e_s}(t)\big) \geqslant 0$ for all s.

Thus, h is nondecreasing, and therefore:

$$\text{(6.9)}\quad \sum_{s=1}^{n-1} g\big(J'_{e_s}(1), J_{e_s}(1)\big) \geqslant \sum_{s=1}^{n-1} g\big(J'_{e_s}(0), J_{e_s}(0)\big) \overset{\text{by (6.7)}}{=} -\mathrm{tr}(S^{\Sigma_0}) \overset{\text{by (6.2)}}{\geqslant} 0.$$

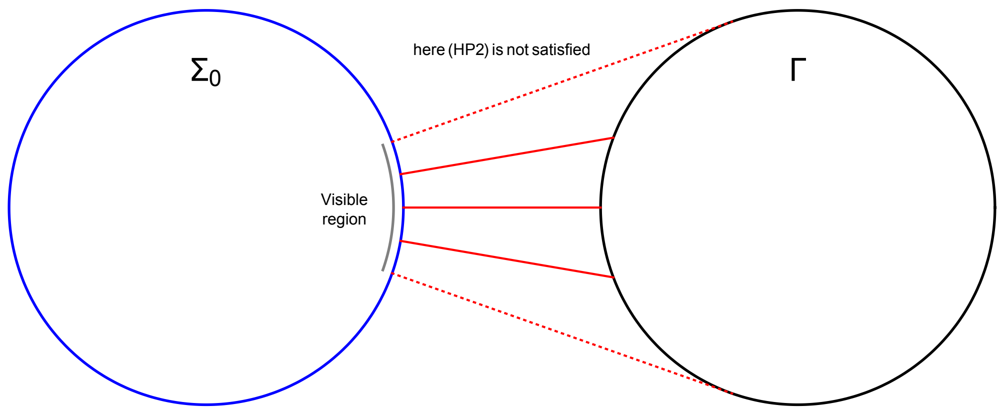

FIGURE 8. A noncompact visible region

From (6.6) and (6.9) we get $\mathrm{tr}(\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_t}) > 0$, which proves the first statement. Now, recall from Theorem 3.1, part (3), that $\mathcal{C}^{\ell_p}_{\mathcal{P}_0,\Gamma_t}$ is the second variation of the arrival time. The conclusion is that p is not a local maximum of the arrival time. □

An immediate consequence is the following.

Corollary 6.5. *With the notation and assumptions of Proposition 6.4, the visible region* $\mathcal{P}_0 \subseteq \Sigma_0$ *cannot be compact.* □

Note that the fact that although the visible region is never compact for a null convex optical domain, the emitting body Σ_0 can of course be compact, and there might well exist a null geodesic in the pulse $\mathcal{F}$ which maximizes the arrival time. However in this situation conditions (HP1–HP3) must be violated. This is illustrated by the following situation. Take $M = \mathbb{R}^{n+1}$ with $n \geqslant 3$ with the flat metric

$$g = -dx_0^2 + \sum_{i=1}^{n} dx_i^2,$$

and time-orientation such that the timelike vector field ∂_{x_0} is future pointing. Adopt the time function $T : (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \mapsto x_0 \in \mathbb{R}$, and fix any number $c \geqslant 3$. Let

$$\Omega_c = \Big\{(x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; \sum_{i=1}^{n} x_i^2 > 1 \text{ and } (x_1 - c)^2 + \sum_{i=2}^{n} x_i^2 > 1\Big\}.$$

It is easy to see that (T, Ω_c) is a null concave optical domain. The boundaries are the cylinders

$$\Sigma = \Big\{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; \sum_{i=1}^{n} x_i^2 = 1\Big\}$$

and

$$\Gamma = \Big\{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; (x_1 - c)^2 + \sum_{i=2}^{n} x_i^2 = 1\Big\},$$

so that the emitting body is the $(n-1)$ sphere

$$\Sigma_0 = \Big\{(0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; \sum_{i=1}^{n} x_i^2 = 1\Big\}.$$

Let $\mathbf{x}_\alpha : t \in [0, +\infty[\mapsto p_\alpha t \in \mathbb{R}^n$ be the one-parameter family of half-lines emanating from the center of the sphere S^{n-1}

$$\sum_{i=1}^{n} x_i^2 = 1$$

which are tangent to the sphere

$$(x_1 - c)^2 + \sum_{i=2}^{n} x_i^2 = 1,$$

where the parametrization has been chosen such that $\mathbf{x}_\alpha(1) = p_\alpha$ are the tangency points. This family of curves comprise a cone whose interior intersects S^{n-1} in the open set C, see Figure 8.

It is not difficult to check that the curves $\ell_\alpha : t \in [0, 1] \mapsto (\|\mathbf{x}_\pm(t)\|, \mathbf{x}_\pm(t)) \in \mathbb{R}^{n+1}$ are lightlike geodesics maximizing the time of arrival from Σ_0 to Γ among all those null geodesics from the origin of $\mathbb{R}^{n+1}$ which do reach Γ. The visible region is

$$\mathcal{P}_0 = \Big\{(0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} \;\Big|\; (x_1, \ldots, x_n) \in C\Big\},$$

which is non-compact. The existence of "slowest" null geodesics does not contradict Proposition 6.4 because the geodesics ℓ_α are tangent to Γ, and thus violate (HP2). They do not emanate from $\mathcal{P}_0$, but from its (compact) closure.

References

[1] L. Andersson, M. Mars, W. Simon, *Local Existence of Dynamical and Trapping Horizons*, Phys. Rev. Lett. **95**, 111102 (2005)

[2] L. Andersson, M. Mars, J. Metzger, W. Simon, *The time evolution of marginally trapped surfaces*, Class. Quantum Grav. **26** (2009) 085018

[3] J. Ehlers, E. Newman, *The theory of caustics and wave front singularities with physical applications*, J. Math. Phys. **41** (2000), no. 6, 3344–3378.

[4] P. Ehrlich, S.-B. Kim, *A focal index theorem for null geodesics*, J. Geom. Phys. **6** (1989), no. 4, 657–670.

[5] M. Eichmair, G. J. Galloway, D. Pollack, *Topological censorship from the initial data point of view*, J. Diff. Geom. **95** (2013) 389–405

[6] R. Giambò, F. Giannoni, P. Piccione, *Gravitational lensing in general relativity via bifurcation theory*, Nonlinearity **17** (2004), no. 1, 117–132.

[7] R. Giambò, F. Giannoni, P. Piccione, *Genericity of nondegeneracy for light rays in stationary spacetimes*, Comm. Math. Phys. **287** (2009), no. 3, 903–923.

[8] R. Giambò, F. Giannoni, P. Piccione, *Functions on the sphere with critical points in pairs and orthogonal geodesic chords*, J. Differential Equations **260** (2016), no. 11, 8261–8275.

[9] F. GIANNONI, A. MASIELLO, P. PICCIONE, *A variational theory for light rays in stably causal Lorentzian manifolds: regularity and multiplicity results*, Comm. Math. Phys. **187** (1997), no. 2, 375–415.

[10] M. GOLUBITSKY, V. GUILLEMIN, *Stable mappings and their singularities*, Graduate Texts in Mathematics, Vol. 14. Springer–Verlag, New York-Heidelberg, 1973.

[11] W. HASSE, M. KRIELE, V. PERLICK, *Caustics of wavefronts in general relativity* Classical Quantum Gravity **13** (1996), no. 5, 1161–1182.

[12] W. HASSE, V. PERLICK, *A Morse-theoretical analysis of gravitational lensing by a Kerr-Newman black hole*, J. Math. Phys. **47** (2006), no. 4.

[13] P. JETZER, Y. MELLIER, V. PERLICK, *Editorial on special issue on gravitational lensing*, Gen. Relativity Gravitation **42** (2010), no. 9, 2009–2010.

[14] B. MORIN, *Formes canoniques des singularités d'une application différentiable*,(C. R. Acad. Sci. Paris **260** (1965) 5662–5665, 6503–6506.

[15] V. PERLICK, *Ray optics, Fermat's principle, and applications to general relativity*. Lecture Notes in Physics. Monographs, 61. Springer–Verlag, Berlin, 2000.

[16] V. PERLICK, *Gravitational Lensing from a Spacetime Perspective*, Living Rev. Relativity **7**, (2004), 9. http://www.livingreviews.org/lrr-2004-9.

[17] V. PERLICK, P. PICCIONE, *A general-relativistic Fermat principle for extended light sources and extended receivers*, Gen. Relativity Gravitation **30** (1998), no. 10, 1461–1476.

[18] A. O. PETTERS, *Gravity's action on light*, Notices Amer. Math. Soc. **57** (2010), no. 11, 1392–1409.

[19] P. PICCIONE, D. V. TAUSK, *A note on the Morse index theorem for geodesics between submanifolds in semi-Riemannian geometry*, J. Math. Phys. **40** (1999), no. 12, 6682–6688.

[20] J.M.M. SENOVILLA, D. GARFINKLE, *The 1965 Penrose singularity theorem*, Class. Quantum Grav. **32** (Focus issue: Milestones of general relativity) 124008

[21] R. THOM, *Structural stability and morphogenesis. An outline of a general theory of models*. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1989.

[22] C.T.C. WALL, *Geometric properties of generic differentiable manifolds*. Geometry and topology (Proc. III Latin Amer. School of Math., Inst. Mat. Pura Aplicada CNPq, Rio de Janeiro, 1976), pp. 707–774.

[23] G. WASSERMANN, *Stability of caustics*, Math. Ann. **216** (1975), 43–50.

[24] H. WHITNEY, *On singularities of mappings of Euclidean spaces. I. Mappings of the plane into the plane*, Ann. of Math. (2) **62** (1955), 374–410.

SCUOLA DI SCIENZE E TECNOLOGIE, SEZIONE DI MATEMATICA
UNIVERSITÀ DI CAMERINO, ITALY
E-mail address: roberto.giambo@unicam.it, fabio.giannoni@unicam.it

DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE DE SÃO PAULO
RUA DO MATÃO, 1010
05508-090 SÃO PAULO, SP, BRAZIL
URL: http://www.ime.usp.br/~piccione
E-mail address: piccione.p@gmail.com