DIFFERENTIAL GEOMETRY AND APPLICATIONS
Satellite Conference of ICM in Berlin, Aug. 10 – 14, 1998, Brno
Masaryk University in Brno (Czech Republic) 1999, 631–644

CAUSAL TRAJECTORIES BETWEEN SUBMANIFOLDS IN LORENTZIAN GEOMETRY

PAOLO PICCIONE

ABSTRACT. We set up an infinite dimensional differential structure for the set of causal curves joining two submanifolds P and Γ of a Lorentzian manifold $\mathcal{M}$, and we prove a variational principle that characterizes the timelike geodesics. Such principle, which is obtained using a sort of *arrival time* functional, is given by a general-relativistic version of the Fermat's principle for light rays in Classical Optics. The lightlike geodesics between P and Γ are obtained with a limit process, which is briefly discussed in the last section.

1. Introduction

The aim of this paper is to develop a variational theory for causal geodesics between submanifolds of a Lorentzian manifold.

I. Kovner [9] has suggested a general-relativistic variational principle for light rays and he has discussed its relevance in view of applications to gravitational lensing. This variational principle, which can be viewed as a general-relativistic *Fermat principle*, is formulated in the following way. In an arbitrary Lorentzian manifold $(\mathcal{M}, g)$, i.e., in an arbitrary spacetime according to general relativity, one fixes a point p and a timelike curve γ. The point p is to be interpreted as a pointlike light source at a particular instant of time and γ is to be interpreted as the worldline of a pointlike receiver. Now one considers all lightlike curves from p to γ, i.e., all possibilities to go from p to γ at the (vacuum) speed of light, and one assigns to each of those curves an "arrival time" with the help of a parameterization for γ. Kovner gave a plausibility argument that, indeed, the lightlike *geodesics* are characterized by stationary arrival time, i.e., that the actual (vacuum) light rays are local extrema or saddle-points of the arrival time. A rigorous mathematical proof of this claim was given in Ref. [14]. Kovner's variational principle was further investigated, both from a physical and from a mathematical point of view, e.g.,

---

1991 *Mathematics Subject Classification.* 53C22, 53C50, 58E30.

Key words and phrases. Lorentzian geometry, causal geodesics, Fermat's Principle.

The author gratefully acknowledges the financial support provided by FAPESP (São Paulo, Brazil), Processo n. 98/4729-3.

This paper is in final form and no version of it will be submitted for publication elsewhere.

by L. Bel and J. Martín [2] and in Refs. [16], [4] and [6]. For applications to gravitational lensing, we also refer to the book by P. Schneider, J. Ehlers, and E. Falco [20]. In Kovner's variational principle the light source and the receiver are assumed to be pointlike.

In [3], the Fermat's principle has been extended to the case of timelike geodesics between a pointlike source and receiver; moreover, in [5], under the assumption of stable causality for the Lorentzian manifold $\mathcal{M}$, using the timelike extension of the Fermat principle it has been developed an existence theory for timelike geodesics.

In [17], the authors have proven a more general variational principle for light rays where the light source and the receiver may have a spatial extension. In this paper we proceed analogously, and we prove a timelike extension of the Fermat principle between a source and a receiver that may have a spatial extension. To that end we replace the point p with a spacelike submanifold P and we replace the timelike curve γ with a timelike submanifold Γ. The manifold P could be a spacelike 2-sphere, to be interpreted as the surface of source at a particular instant of time, and Γ could be a timelike 3-manifold, to be interpreted as the history of a screen. For $\varepsilon > 0$ fixed, we consider the set $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ of (future pointing) timelike curves z joining P and Γ and such that $g(\dot{z},\dot{z}) = -\varepsilon^2$. Generalizing Kovner's idea, we assign an *arrival time* to each curve $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$ with the help of a time function on Γ, i.e., with the help of a function that slices Γ into spacelike submanifolds. The Fermat principle we are going to prove in Section 3 below says that, among all curves in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$, the curves of stationary arrival time are exactly the geodesics in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ which start perpendicularly to P and terminate perpendicularly to a time slice of Γ. We emphasize that the Fermat principle determines all curves of stationary arrival time and not only those of minimal arrival time; moreover, the stationary curves for the arrival time are affinely parameterized, unlike the lightlike case.

The main motivation for our result is that, in order to develop existence and multiplicity results for the lightlike geodesics, one encounters two the main difficulties in applying global analytical techniques. In first place, the set $\mathcal{L}_{P,\Gamma}$ of lightlike (i.e., $g(\dot{z},\dot{z}) = 0$) curves joining P and Γ does not have a reasonable Banach or Hilbert differentiable structure. Indeed, it is not difficult to prove that each curve whose tangent vector vanishes on a set of positive measure is a singular point for $\mathcal{L}_{P,\Gamma}$. Secondly, both the arrival time functional and the space $\mathcal{L}_{P,\Gamma}$ are invariant by reparameterization; in particular, the arrival time functional does *never* satisfy the Palais–Smale condition, moreover, since its critical points are not isolated, it is not a Morse functional.

What one can do to avoid these difficulties is to develop an existence theory for the critical points of the arrival time in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$, where the two problems mentioned above do not appear, and then pass to the limit as ε tends to zero to obtain the corresponding result for the lightlike geodesics. This approach has been used in the case of pointlike sources and receivers in Ref. [7], where the Morse relations for the space $\mathcal{L}_{P,\Gamma}$ have been obtained with a limit process. The aim of this paper is to set up the basis to develop a Ljusternik–Schnirelman theory and a Morse theory for timelike and lightlike geodesics between submanifolds.

In order to implement this project, the following three things are done in this article: it is established a Hilbert differentiable structure on the set $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ (Section 2); it is proven the Fermat Principle that characterizes the geodesics in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ (Section 3); finally, it is discussed the limit process as ε tends to 0 (Section 4). For all the definitions and basic properties of infinite dimensional differential geometry we refer to the book [10]; for the basic notions of Lorentzian geometry and their physical interpretation, we refer to classical textbooks, like [1, 8, 12, 18, 19].

2. The Infinite Dimensional Differential Structure

We consider an arbitrary Lorentzian manifold $(\mathcal{M}, g)$. More precisely, we assume that $\mathcal{M}$ is a finite-dimensional real C^∞ manifold of dimension greater than two whose topology satisfies the second countability axiom and the Hausdorff separation axiom, and we assume that g is a pseudo-Riemannian C^∞ metric of signature $(+, ..., +, -)$ on $\mathcal{M}$.

The symbol $\langle \cdot, \cdot \rangle$ will denote the bilinear form induced by g on the tangent spaces of $\mathcal{M}$; the usual nabla symbol ∇ will denote the covariant derivative relative to the Levi–Civita connection of g. At each point $p \in \mathcal{M}$, we denote the tangent space to $\mathcal{M}$ by $T_p\mathcal{M}$; we denote by exp the exponential map of $(\mathcal{M}, g)$, defined on an open subset A of $T\mathcal{M}$; for $p \in \mathcal{M}$, $\exp_p$ will denote the restriction of exp to $A \cap T_p\mathcal{M}$.

We call a linear subspace $\mathcal{W}_p$ of $T_p\mathcal{M}$ *spacelike* if g is positive definite on $\mathcal{W}_p$, *lightlike* if g is positive semidefinite but not positive definite on $\mathcal{W}_p$, and *timelike* otherwise. This implies that $\mathcal{W}_p$ is spacelike if and only if the orthocomplement of $\mathcal{W}_p$ is timelike, and vice versa (cf., e.g., [19, p.20]). A vector $v \in T_p\mathcal{M}$ is called *spacelike*, *lightlike* or *timelike* if the linear subspace $\{sv | s \in \mathbb{R}\}$ has the respective property. Equivalently, v is *spacelike* if $\langle v, v \rangle > 0$ or $v = 0$, *lightlike* if $\langle v, v \rangle = 0$ but $v \neq 0$, and *timelike* if $\langle v, v \rangle < 0$. Finally, we call a submanifold Σ of $\mathcal{M}$ *spacelike*, *lightlike* or *timelike*, if at all points $p \in \Sigma$ the tangent space $T_p\Sigma$ has the respective property. Note that, according to this definition, a single point is a (zero-dimensional) spacelike submanifold.

Given a smooth function ϕ on $\mathcal{M}$, for $q \in \mathcal{M}$ we denote by $\nabla\phi(q)$ the gradient of ϕ at q with respect to g, which is the vector in $T_q\mathcal{M}$ defined by $\langle \nabla\phi(q), \cdot \rangle = \mathrm{d}\phi(q)[\,\cdot\,]$. We set $m = \dim(\mathcal{M})$; the physical interesting case is $m = 4$.

As customary, if $I \subseteq \mathbb{R}$ is any interval, we will denote by $H^2(I, \mathbb{R}^n)$ the Sobolev space of C^1 curves $z : I \longmapsto \mathbb{R}^n$ having absolutely continuous derivative, and such that the integral $\int_I |\ddot{z}|^2 \,\mathrm{d}t$ is finite, where the dots on the symbol of the curve z mean derivatives and $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$. We also denote by $a \cdot b$ the Euclidean inner product between vectors in $\mathbb{R}^n$.

As to the topology of H^2, observe that a sequence z_n is H^2-convergent if and only if it is uniformly convergent up to the first derivative, and the sequence of the second derivatives $\ddot{z}_n$ is L^2-convergent.

The space $H^2(I,\mathbb{R}^n)$ has the structure of an infinite dimensional separable Hilbert space, with inner product given by:

$$\langle z_1,z_2\rangle_2=\int_0^1 (z_1\cdot z_2+\dot z_1\cdot\dot z_2+\ddot z_1\cdot\ddot z_2)\ \mathrm{d}t\,.$$

Given *any* differentiable manifold N, the set $H^2([0,1],N)$ is defined as the set of all C^1-curves $z:[0,1]\longmapsto N$ such that, for every local chart (V,φ) on N, with $\varphi:U\longmapsto\mathbb{R}^n$ a diffeomorphism, and for every closed subinterval $I\subseteq[0,1]$ such that $z(I)\subset V$, it is $\varphi\circ z\in H^2(I,\mathbb{R}^n)$.

It is not difficult to see that this definition of $H^2([0,1],N)$ may be given equivalently in the following two ways:

- a curve $z:[0,1]\longmapsto N$ belongs to $H^2([0,1],N)$ if and only if there exists a finite sequence $I_1,\dots,I_k$ of closed subintervals of $[0,1]$ and a finite number of charts $\varphi_i:U_i\longmapsto\mathbb{R}^n$ on N, $i=1,\dots,k$, such that $\bigcup_{i=1}^k I_k=[0,1]$, $z(I_i)\subset U_i$, and $\varphi_i\circ z\in H^2(I_i,\mathbb{R}^n)$ for all $i=1,\dots,k$;
- a C^1-curve $z:[0,1]\longmapsto N$ is in $H^2([0,1],N)$ if and only if for one (hence for every) Riemannian metric g_{R} on N, the integral $\int_0^1 g_{\mathrm{R}}(\nabla^{(\mathrm{R})}_{\dot z}\dot z,\nabla^{(\mathrm{R})}_{\dot z}\dot z)\ \mathrm{d}t$ is finite, where $\nabla^{(\mathrm{R})}$ is the Levi–Civita connection of g_{R}.

A classical result of Global Analysis (see Ref. [13]) states that, for every differentiable manifold N, the Sobolev space $H^2([0,1],N)$ has the structure of an infinite dimensional manifold, modeled on the Hilbert space $H^2([0,1],\mathbb{R}^n)$.

Let $T\mathcal{M}$ denote the tangent bundle of $\mathcal{M}$, with $\pi:T\mathcal{M}\longmapsto\mathcal{M}$ the canonical projection. Given any curve $z:I\subseteq\mathbb{R}\longmapsto\mathcal{M}$, a *vector field along* z is a map $\zeta:I\longmapsto T\mathcal{M}$ such that $\pi\circ\zeta=z$.

We will consider from now on a timelike submanifold Γ and a spacelike submanifold P of our Lorentzian manifold $\mathcal{M}$; recall that this means that the restriction of g to $T\Gamma$ is again a Lorentzian (non degenerate) metric on Γ, and that the restriction of g to TP is positive definite. The reason for restricting the causal character of the submanifolds P and Γ will be clear in what follows (see Remark 3.3).

We will denote by $\Omega^{(2)}_{P,\Gamma}$ the set of curves in $H^2([0,1],\mathcal{M})$ that join P and Γ:

$$\Omega^{(2)}_{P,\Gamma}=\Big\{z\in H^2([0,1],\mathcal{M})\ \Big|\ z(0)\in P,\ z(1)\in\Gamma\Big\}\,. \tag{1}$$

It is well known (Ref. [13]) that also $\Omega^{(2)}_{P,\Gamma}$ has the structure of an infinite dimensional differentiable manifold, with the same regularity of the manifolds P and Γ. For all $z\in\Omega^{(2)}_{P,\Gamma}$, the tangent space $T_w\Omega^{(2)}_{P,\Gamma}$ is given by:

$$T_z\Omega^{(2)}_{P,\Gamma}=\Big\{\zeta\in H^2([0,1],T\mathcal{M})\ \Big|\ \zeta\ \text{ vector field along }\ z\Big\}\,. \tag{2}$$

If $\{E_1,E_2,\dots,E_m\}$ is any vector field basis for $T\mathcal{M}$ along z, for instance, if $\{E_1,E_2,\dots,E_m\}$ is a parallely transported basis along z, then one has a (non canonical) isomorphism $T_z\Omega^{(2)}_{P,\Gamma}\simeq H^2([0,1],\mathbb{R}^m)$ by setting

$$T_z\Omega^{(2)}_{P,\Gamma}\ni\zeta=\sum_{i=1}^m\lambda_i\,E_i\longmapsto(\lambda_1,\lambda_2,\dots,\lambda_m)\in H^2([0,1],\mathbb{R}^m)\,; \tag{3}$$

in particular, $T_z\Omega^{(2)}_{P,\Gamma}$ can be endowed with a Hilbert space structure by considering the pullback of the norm of $H^2([0,1],I\!R^m)$. It is easy to see that all the Hilbert norms on $T_z\Omega^{(2)}_{P,\Gamma}$ that arise from isomorphisms as in (3) are indeed equivalent. This fact will allow us to choose, depending on the situation, the most convenient Hilbert norm on $T_z\Omega^{(2)}_{P,\Gamma}$ to make explicit calculations.

Using the Lorentzian exponential map in $\mathcal{M}$, associated to each $z\in\Omega^{(2)}_{P,\Gamma}$ and $\zeta\in T_z\Omega^{(2)}_{P,\Gamma}$ it is possible to define a smooth *variation* z_s of z in $\Omega^{(2)}_{P,\Gamma}$ with variational vector field ζ. This means that z_s, $s\in]-\delta,\delta[$ for some $\delta>0$, is a family of curves in $\Omega^{(2)}_{P,\Gamma}$ such that:

(a) $s\longmapsto z_s$ is a smooth map from $]-\delta,\delta[$ to $\Omega^{(2)}_{P,\Gamma}$;
(b) $z_0=z$;
(c) $\zeta(t)=\frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}z_s(t)$ for all $t\in[0,1]$.

Conversely, given any smooth variation z_s of z such that (a), (b) and (c) above hold true, then the corresponding variational vector field ζ is in $T_z\Omega^{(2)}_{P,\Gamma}$.

Given a variation $\eta(s,t)=z_s(t)$ of a curve $z\in\Omega^{(2)}_{P,\Gamma}$, we will denote by $\frac{\mathrm{D}}{\mathrm{d}s}$ and $\frac{\mathrm{D}}{\mathrm{d}t}$, the operations of covariant derivative of the Levi–Civita connection of g in the directions of $\frac{\partial\eta}{\partial s}$ and $\frac{\partial\eta}{\partial t}$ for vector fields along η; the usual symbols $\frac{\mathrm{d}}{\mathrm{d}s}$ and $\frac{\mathrm{d}}{\mathrm{d}t}$ will denote the differentials with respect to s and t of functions along η.

Since the Lie bracket $[\frac{\mathrm{D}}{\mathrm{d}s},\frac{\mathrm{D}}{\mathrm{d}t}]$ vanish and the Levi–Civita connection is torsion free, we have the following commutation relations involving the operators $\frac{\mathrm{D}}{\mathrm{d}s}$, $\frac{\mathrm{D}}{\mathrm{d}t}$, $\frac{\mathrm{d}}{\mathrm{d}s}$ and $\frac{\mathrm{d}}{\mathrm{d}t}$:

(4)
$$\frac{\mathrm{D}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}=\frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s},\quad \frac{\mathrm{D}}{\mathrm{d}s}\frac{\mathrm{D}}{\mathrm{d}t}=\frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{D}}{\mathrm{d}s}+R(\frac{\mathrm{D}}{\mathrm{d}s},\frac{\mathrm{D}}{\mathrm{d}t}),\quad \frac{\mathrm{D}}{\mathrm{d}t}\frac{\mathrm{D}}{\mathrm{d}s}=\frac{\mathrm{D}}{\mathrm{d}s}\frac{\mathrm{D}}{\mathrm{d}t}+R(\frac{\mathrm{D}}{\mathrm{d}t},\frac{\mathrm{D}}{\mathrm{d}s})\,;$$

where R is the *curvature tensor* of the Lorentzian metric g, defined with the following sign convention:

$$R(X,Y)=\nabla_X\nabla_Y-\nabla_Y\nabla_X-\nabla_{[X,Y]}.$$

for X,Y vector fields on $\mathcal{M}$.

In order to construct an *arrival time* functional for curves joining P and Γ, we need a function that *measures the time of events* of Γ. To this aim, let's assume now that $T:\Gamma\longmapsto I\!R$ is a smooth *time function* on Γ, i.e., the gradient ∇T of T is a timelike vector field on Γ. The existence of such a function T is the motivation for requiring that Γ be timelike, as mentioned above.

If Γ is one dimensional, i.e., if Γ is a timelike curve curve embedded in $\mathcal{M}$, then such a function T is just a parameterization of Γ.

If $\dim(\Gamma)>1$, then a time function $T:\Gamma\longmapsto I\!R$ exists if and only if Γ, viewed as a Lorentzian manifold in its own right, is *stably causal* (see Ref. [8, p. 198])

Let T be a fixed time a on Γ. Since ∇T is timelike, then, for all $c\in I\!R$, the inverse image $T^{-1}(c)$ is a possibly empty spacelike submanifold of Γ, that will be denoted by Γ_c.

We consider the smooth timelike vector field Y on Γ, given by a suitable normalization of ∇T:

$$Y = \frac{\nabla T}{\langle \nabla T, \nabla T \rangle}. \tag{5}$$

The following clearly holds:

$$\mathrm{d}T[Y] = \langle \nabla T, Y \rangle \equiv 1, \quad \text{and} \quad Y \perp \Gamma_c. \tag{6}$$

The existence of a time function on Γ allows to define an *arrival time functional* τ in the space of all curves with final endpoint in Γ: for all $z \in \Omega^{(2)}_{P,\Gamma}$, let $\tau(z)$ be defined by:

$$\tau(z) = T(z(1)). \tag{7}$$

Proposition 2.1. *The functional τ is smooth on $\Omega^{(2)}_{P,\Gamma}$. For all $z \in \Omega^{(2)}_{P,\Gamma}$ and all $\zeta \in T_z\Omega^{(2)}_{P,\Gamma}$, the Gateaux derivative $\tau'(z)[\zeta]$ is given by:*

$$\tau'(z)[\zeta] = \langle \nabla T(z(1)), \zeta(1) \rangle. \tag{8}$$

Proof. Let $z \in \Omega^{(2)}_{P,\Gamma}$ be fixed. Since (8) defines a smooth functional on $T_z\Omega^{(2)}_{P,\Gamma}$ (it is clearly linear and continuous in ζ), it suffices to show that τ is Frechet differentiable at z in the direction ζ, and that the derivative $\tau'(z)[\zeta]$ is given by $\langle \nabla T(z(1)), \zeta(1) \rangle$. To this aim, let ζ be fixed and let z_s be a smooth variation of z in $\Omega^{(2)}_{P,\Gamma}$ with variational vector field ζ. Then it is is easily computed:

$$\tau'(z)[\zeta] = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \tau(z_s) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} T(z_s(1)) = \langle \nabla T(z_s(1)), \frac{\mathrm{d}}{\mathrm{d}s} z_s(1) \rangle\big|_{s=0} =$$

$$= \langle \nabla T(z(1)), \zeta(1) \rangle,$$

which concludes the proof. □

Given any C^1-curve $z : [0,1] \longmapsto \mathcal{M}$ having endpoint $z(1)$ in Γ, we denote by U_z the C^1-vector field on z obtained by parallel transport of $Y(z(1))$ along z. This means that U_z is the unique vector field along z satisfying:

$$\nabla_{\dot{z}} U_z = 0, \quad U_z(1) = Y(z(1)).$$

Observe that, since the parallel transport is an isometry, then the quantity

$$\langle U_z, U_z \rangle = \langle Y(z(1)), Y(z(1)) \rangle$$

is constant (negative) along z.

If z is a causal curve with final endpoint in Γ, we say that z is *future pointing* with respect to the time orientation of Γ induced by Y if, for all t, the tangent vector $\dot{z}(t)$ and the vector $U_z(t)$ belong to the same light half-cone, i.e., if $\langle \dot{z}(t), U_z(t) \rangle < 0$ everywhere. A similar definition is given for *past pointing* curves.

For our variational principle, we want to consider the set of timelike curves in $\Omega^{(2)}_{P,\Gamma}$ parameterized by a constant multiple of the proper time and that are future pointing with respect to the time orientation of Γ induced by the timelike vector field Y. To this aim, for all $\varepsilon > 0$ we introduce the set:

$$\mathcal{R}^+_{P,\Gamma,\varepsilon} = \Big\{ z \in \Omega^{(2)}_{P,\Gamma} : \big\langle \dot z, \dot z \big\rangle \equiv -\varepsilon^2 \text{ and } \big\langle \dot z, U_z \big\rangle < 0 \text{ on } [0,1] \Big\}. \tag{9}$$

Proposition 2.2. *For all $\varepsilon > 0$, the set $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ is a smooth submanifold of $\Omega^{(2)}_{P,\Gamma}$, and, in particular, the restriction of the functional τ on $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ is smooth. For all $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$, the tangent space $T_z\mathcal{R}^+_{P,\Gamma,\varepsilon}$ can be identified with the Hilbert subspace of $T_z\Omega^{(2)}_{P,\Gamma}$ given by:*

$$T_z\mathcal{R}^+_{P,\Gamma,\varepsilon} = \Big\{ \zeta \in T_z\Omega^{(2)}_{P,\Gamma} \;\Big|\; \big\langle \nabla_{\dot z}\zeta, \dot z \big\rangle = 0 \textit{ on } [0,1] \Big\}. \tag{10}$$

Proof. We define a map $G : \Omega^{(2)}_{P,\Gamma} \longmapsto H^1([0,1], \mathbb{R})$ by:

$$z \longmapsto \big\langle \dot z, \dot z \big\rangle. \tag{11}$$

It is easy to see that G is a smooth map. Indeed, G is clearly continuous, moreover, given a variational vector field $\zeta \in T_z\Omega^{(2)}_{P,\Gamma}$ with corresponding variation z_s, recalling (4), the Frechet derivative $G'(z)[\zeta]$ is computed as:

$$\begin{aligned} G'(z)[\zeta] &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \Big\langle \frac{\mathrm{d}}{\mathrm{d}t} z_s, \frac{\mathrm{d}}{\mathrm{d}t} z_s \Big\rangle = 2\Big\langle \frac{\mathrm{D}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t} z_s, \frac{\mathrm{d}}{\mathrm{d}t} z_s \Big\rangle = \\ &= 2\Big\langle \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{D}}{\mathrm{d}s} z_s, \frac{\mathrm{d}}{\mathrm{d}t} z_s \Big\rangle = 2\big\langle \nabla_{\dot z}\zeta, \dot z \big\rangle. \end{aligned} \tag{12}$$

Formula (12) gives a smooth map in z and linear continuous in ζ, hence G is smooth.

Given $\varepsilon > 0$, let f_ε denote the constant function equal to $-\varepsilon^2$ and let G_ε be the subset of $\Omega^{(2)}_{P,\Gamma}$ given by the inverse image $G^{-1}(f_\varepsilon)$. We use the Implicit Function Theorem (see [10]) to prove that G_ε is a smooth submanifold of $\Omega^{(2)}_{P,\Gamma}$. To this aim, we need to show that the differential $G'(z) : T_z\Omega^{(2)}_{P,\Gamma} \longmapsto H^1([0,1], \mathbb{R})$ is surjective for all $z \in G_\varepsilon$, i.e., from (12), given any $h \in H^1([0,1], \mathbb{R})$, there exists $\zeta \in T_z\Omega^{(2)}_{P,\Gamma}$ such that:

$$\big\langle \nabla_{\dot z}\zeta, \dot z \big\rangle = h. \tag{13}$$

To this aim, given such a z and h, we define $\zeta = \mu \cdot U_z$, with $\mu \in H^2([0,1], \mathbb{R})$ to be determined in such a way that (13) holds. Observe that, since $U_z(1) = Y(z(1)) \in T_{z(1)}\Gamma$, the vector field $\mu \cdot U_z$ belongs to $T_z\Omega^{(2)}_{P,\Gamma}$ provided that $\mu(0) = 0$. Moreover, since U_z is parallel, then:

$$\nabla_{\dot z}(\mu \cdot U_z) = \mu' \cdot U_z.$$

So, we need a solution in $H^2([0,1], \mathbb{R})$ for the following Cauchy problem:

$$\mu' = \frac{h}{\big\langle U_z, \dot z \big\rangle}, \quad \mu(0) = 0. \tag{14}$$

Observe that, by the *wrong-way Schwartz's inequality*, for all $z \in G_\varepsilon$ it is:

$$\langle U_z, \dot z\rangle \geq \langle U_z, U_z\rangle \cdot \langle \dot z, \dot z\rangle = -\varepsilon^2 \cdot \langle Y(z(1)), Y(z(1))\rangle > 0,$$

hence $\frac{h}{\langle U_z, \dot z\rangle} \in H^1([0,1], I\!R)$. The desired function μ is therefore given by:

$$\mu(t) = \int_0^t \frac{h}{\langle U_z, \dot z\rangle}\,\mathrm{d}r,$$

which proves that G_ε is a smooth submanifold of $\Omega^{(2)}_{P,\Gamma}$.

The condition $\langle \dot z, U_z\rangle < 0$ is *open* in $\Omega^{(2)}_{P,\Gamma}$, so $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ is an open submanifold of G_ε, which proves the first part of the thesis.

For the second part, by the Implicit Function Theorem, the tangent space $T_z\mathcal{R}^+_{P,\Gamma,\varepsilon}$ is given by the kernel of $G'(z)$. Hence, (10) follows at once from (12), which concludes the proof. □

3. The Fermat Principle

We need the following representation of the differential τ' on the space $\mathcal{R}^+_{P,\Gamma,\varepsilon}$.

Proposition 3.1. *Let $\varepsilon > 0$ be fixed. For all $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$ and all $\zeta \in T_z\mathcal{R}^+_{P,\Gamma,\varepsilon}$, the derivative $\tau'(z)[\zeta]$ has the following integral representation:*

$$\tau'(z)[\zeta] = \frac{\langle \dot z(0), \zeta(0)\rangle}{\langle \dot z(0), U_z(0)\rangle} - \frac{\langle \dot z(1), \zeta(1)^\perp\rangle}{\langle \dot z(1), U_z(1)\rangle} + \int_0^1 \langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, \zeta\rangle\,\mathrm{d}t\,, \tag{15}$$

where $\zeta(1)^\perp$ is the component of $\zeta(1)$ in the orthocomplement of $Y(z(1))$.

Proof. Let $z \in \Omega^{(2)}_{P,\Gamma}$ and $\zeta \in T_z\Omega^{(2)}_{P,\Gamma}$ be fixed. Using (10) and Integration by parts of the quantity $\langle \nabla_{\dot z}\zeta, \dot z\rangle\langle U_z, \dot z\rangle^{-1}$ gives:

$$0 = \int_0^1 \frac{\langle \nabla_{\dot z}\zeta, \dot z\rangle}{\langle U_z, \dot z\rangle}\,\mathrm{d}r = \frac{\langle \zeta, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|_0^1 - \int_0^1 \langle \zeta, \frac{\dot z}{\langle U_z, \dot z\rangle}\rangle\,\mathrm{d}r\,. \tag{16}$$

Given any tangent vector v along z, we denote by $v^\perp$ and $v^{||}$ respectively the components of v orthogonal and parallel to U_z; i.e.,

$$v = v^\perp + v^{||}, \quad v^{||} = \frac{\langle v, U_z\rangle}{\langle U_z, U_z\rangle}U_z, \quad v^\perp = v - v^{||}.$$

Applying the above decomposition to the vector $\zeta(1)$ and recalling that $U_z(1) = Y(1) = \nabla T\langle \nabla T, \nabla T\rangle^{-1}$ and formula (8), we have:

(17)

$$\begin{aligned}\frac{\langle \zeta, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|^1 &= \frac{\langle U_z, \zeta\rangle}{\langle U_z, U_z\rangle}\frac{\langle U_z, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|^1 + \frac{\langle \zeta^\perp, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|^1 = \langle \nabla T(z(1)), \zeta(1)\rangle + \frac{\langle \zeta^\perp, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|^1 = \\ &= \tau'(z)[\zeta] + \frac{\langle \zeta^\perp, \dot z\rangle}{\langle U_z, \dot z\rangle}\Big|^1.\end{aligned}$$

Formula (15) now follows easily by substituting (17) into (16). □

We are ready to state and prove our variational principle for timelike geodesics between P and Γ.

Theorem 3.2 (Fermat Principle for Timelike Geodesics)**.** *Let $\varepsilon > 0$ be fixed and z a curve in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$. Then, z is a stationary point for τ in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ if and only if z is a geodesic such that $\dot z(0)$ is perpendicular to P and $\dot z(1)$ is perpendicular to the time slice $\Gamma_{\tau(z)}$ of Γ.*

Proof. Let $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$ be fixed and let V be any H^2-vector field along z vanishing at the endpoints $z(0)$ and $z(1)$. Let $f_V \in H^2([0,1], I\!R)$ be the function defined by:

$$f_V(r) = -\int_0^r \frac{\langle \nabla_{\dot z} V, \dot z\rangle}{\langle U_z, \dot z\rangle}\, \mathrm{d}t\,. \tag{18}$$

With this choice, we have that the vector field $\zeta = V + f_V \cdot U_z$ belongs to $T_z\mathcal{R}^+_{P,\Gamma,\varepsilon}$; namely, $\zeta(0) = 0$ and $\zeta(1) = f(1)\cdot U_z(1) \in T_{z(1)}\Gamma$. Moreover, since U_z is parallel, it is $\nabla_{\dot z}\zeta = \nabla_{\dot z} V + f'\cdot U_z$, hence:

$$\langle \nabla_{\dot z}\zeta, \dot z\rangle = \langle \nabla_{\dot z} V, \dot z\rangle + f'\cdot\langle U_z, \dot z\rangle = 0\,.$$

We now insert such a ζ into formula (15); since $V(0) = f(0) = 0$, then the first term of (15) vanishes. Since $V(1) = 0$, then $\zeta(1)^\perp = 0$ and also the second term of (15) is zero. Finally, we compute:

$$\begin{aligned}\tau'(z)[\zeta] &= \int_0^1 \Big\langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, V\Big\rangle\,\mathrm{d}t + \int_0^1 f\Big\langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, U_z\Big\rangle\,\mathrm{d}t = \\ &= \int_0^1 \Big\langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, V\Big\rangle\,\mathrm{d}t + f\cdot\frac{\langle \dot z, U_z\rangle}{\langle \dot z, U_z\rangle}\Big|_0^1 - \int_0^1 f'\cdot\frac{\langle \dot z, U_z\rangle}{\langle \dot z, U_z\rangle}\,\mathrm{d}t = \\ &= \int_0^1 \Big\langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, V\Big\rangle\,\mathrm{d}t\,.\end{aligned} \tag{19}$$

In particular, if z is a stationary point for τ, then:

$$\int_0^1 \Big\langle \nabla_{\dot z}\frac{\dot z}{\langle \dot z, U_z\rangle}, V\Big\rangle\,\mathrm{d}t = 0 \tag{20}$$

for all H^2-vector field V along z vanishing at the endpoints. Hence, by the fundamental Lemma of Calculus of Variations, if $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$ is stationary for τ then, setting $\lambda = \langle \dot z, U_z\rangle^{-1}$, we have::

$$\nabla_{\dot z}\big[\lambda\cdot\dot z\big] = \lambda'\cdot\dot z + \lambda\cdot\nabla_{\dot z}\dot z = 0 \quad \text{almost everywhere on } [0,1]\,. \tag{21}$$

Multiplying (21) by $\dot z$ and keeping in mind that $\langle \nabla_{\dot z}\dot z, \dot z\rangle = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\langle \dot z, \dot z\rangle = 0$, we have:

$$-\varepsilon^2\cdot\lambda' = 0,$$

which implies that $\langle \dot z, U_z\rangle$ is constant. Then equation (21) tells us that $\nabla_{\dot z}\dot z = 0$, and thus z is a geodesic.

To prove the boundary conditions, we specialize to vector fields ζ with arbitrary initial values $\zeta(0)$ tangent to P and $\zeta(1) = 0$. Then, setting formula (15) equal to 0 gives $\langle z'(0), \zeta(0) \rangle = 0$. This shows that $z'(0)$ must be orthogonal to $T_{z(0)}P$. Similarly, we can specialize to vector fields V with arbitrary end values $\zeta(1)$ tangent to $\Gamma_{\tau(z)}$ and $\zeta(0) = 0$. Then, equating (15) to 0 reduces to $\langle \dot z(1), \zeta(1)^{\perp} \rangle = 0$. This shows that $z'(1)$ must be orthogonal to $T_{z(1)}\Gamma_{\tau(z)}$.

Conversely, if $z \in \mathcal{R}^+_{P,\Gamma,\varepsilon}$ is a geodesic such that $\dot z(0)$ is perpendicular to P and $\dot z(1)$ is perpendicular to $\Gamma_{\tau(z)}$, then it is immediate to check that the three terms appearing on the right side of formula (15) vanish, and z is a stationary point for τ in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$.

This concludes the proof. $\square$

Remark 3.3. Observe that the orthogonal vector to a causal vector, i.e., lightlike or timelike, is always spacelike. For this reason, since the solution to our variational problem are timelike geodesics orthogonal to P, we need to require that P be spacelike.

4. Lightlike Geodesics between P and Γ

In this section we will discuss briefly a method of approximating the lightlike geodesics between P and Γ with timelike geodesics, for which the results of the previous sections apply. The idea of such approximation is not new in General Relativity (see T. Levi–Civita [11]).

We recall that, for any *nondegenerate* submanifold Σ of our Lorentzian manifold $\mathcal{M}$, i.e., such that the restriction of the Lorentzian metric g to the tangent bundle $T\Sigma$ is non degenerate, the *second fundamental form* S^Σ (also known as the *shape tensor* of Σ) can be defined in analogy to the Riemannian case. Namely, for each $p \in \Sigma$ and each vector $n \in T_p\mathcal{M}$ which is perpendicular to $T_p\Sigma$ the second fundamental form of Σ in the direction of n is the bilinear form $S^\Sigma_n : T_p\Sigma \times T_p\Sigma \longrightarrow \mathbb{R}$ defined by

$$S^\Sigma_n(v_1, v_2) = \langle n, \nabla_{v_1} V_2 \rangle \tag{22}$$

where V_2 is any C^∞ vector field on Σ which takes the value v_2 at p. As in Riemannian geometry, one can show that S^Σ_n is, indeed, well defined (i.e., independent of the extension V_2 of v_2) and symmetric. For details we refer the reader to J. Beem, P. Ehrlich and K. Easly [1] and to B. O'Neill [12]. Observe that every spacelike submanifold of $\mathcal{M}$ is nondegenerate.

We give the following definition of P-focal point along a lightlike geodesic:

Definition 4.1. Let $z : [0,1] \longmapsto \mathcal{M}$ be a lightlike geodesic with $z(0) \in P$ and $\dot z(0) \perp T_{z(0)}P$. A *$P$-Jacobi field* along z is a smooth vector field J along z satisfying:

1. $\nabla^2_{\dot z} J - R(\dot z, J)\,\dot z = 0$;
2. $J(0) \in T_{z(0)}P$;
3. $\langle \nabla_{\dot z(0)} J, v \rangle + S^P_{\dot z(0)}(J(0), v) = 0$ for all $v \in T_{z(0)}P$.

A point $z(t_0)$ of z is said to be a *P-focal point* if there exists a non zero P-Jacobi field along z such that $J(t_0) = 0$. We denote by $\mathcal{J}_z^P$ the vector space of all P-Jacobi fields along z.

Equation 1 is the Jacobi equation, its geometric meaning is that J is the variational field corresponding to a (smooth) variation z_s of z consisting of geodesics; condition 2 means that, in first approximation, each z_s starts at some point of P; condition 3 means that $\dot z_s(0)$ is orthogonal to P for all s.

A vector field that satisfies the differential equation 1 is said to be a *Jacobi field* along z; a Jacobi field J is uniquely determined by the initial conditions $J(0)$ and $\nabla_{\dot z(0)} J$. There is a linear isomorphism between the vector space $\mathcal{J}_z$ of Jacobi fields along z and the tangent space $T_{\dot z(0)}(T\mathcal{M})$ to the tangent bundle $T\mathcal{M}$ at the point $\dot z(0)$. To describe such an isomorphism, we make the following consideration.

Given any point $p \in \mathcal{M}$ and any vector $v \in T_p\mathcal{M}$, we can identify the tangent space $T_v(T\mathcal{M})$ at v of the tangent bundle $T\mathcal{M}$ with the direct sum $T_p\mathcal{M} \oplus T_p\mathcal{M}$. The first summand $T_p\mathcal{M}$ is the *horizontal* space of the Levi–Civita connection of g, and the second summand $T_p\mathcal{M}$ is the *vertical* subspace, which is tangent to the fiber $T_p\mathcal{M}$. For $\lambda \in T_v(T\mathcal{M})$, we will write $\lambda = (\lambda_h, \lambda_v)$, with $\lambda_v, \lambda_h \in T_p\mathcal{M}$ being the components of λ with respect to this decomposition; this identification will be made throughout the section.

The isomorphism between $T_{\dot z(0)}(T\mathcal{M}) \simeq T_{z(0)}\mathcal{M} \oplus T_{z(0)}\mathcal{M}$ and $\mathcal{J}_z$ is obtained by assigning to each $\lambda = (\lambda_h, \lambda_v) \in T_{\dot z(0)}T\mathcal{M}$ the unique Jacobi field $J \in \mathcal{J}_z$ satisfying $J(0) = \lambda_h$ and $\nabla_{\dot z(0)} J = \lambda_v$. An inverse for this map is obtained as follows. Given $J \in \mathcal{J}_z$, let z_s be a smooth variation of z consisting of geodesics and with variational vector field J. Then, the element $\lambda \in T_{\dot z(0)}T\mathcal{M}$ is given by $\lambda = \frac{\mathrm{D}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t} z_s(t)|_{t=s=0}$. Observing that $z_s(1) = \exp(\frac{\mathrm{d}}{\mathrm{d}t} z_s(0)$, we have:

$$J(1) = \mathrm{d}\exp_{z(0)}(\dot z(0))[\lambda]\,. \tag{23}$$

Let's denote by $TP^\perp$ the normal bundle of P in $\mathcal{M}$, which is a submanifold of $T\mathcal{M}$; the P-Jacobi fields along z correspond to vectors $\lambda \in T_{\dot z(0)}(TP^\perp)$ tangent to the normal bundle of P.

It is now easy to prove the following:

Lemma 4.2. *Let $z : [0,1] \longmapsto \mathcal{M}$ be a (lightlike) geodesic with $z(0) \in P$ and $\dot z(0) \in TP^\perp$. If $z(1)$ is not a P-focal point along z, then the exponential map* exp *of $\mathcal{M}$ gives a local diffeomorphism between a neighborhood of $\dot z(0)$ in $TP^\perp$ and a neighborhood of $z(1)$ in $\mathcal{M}$.*

Proof. If $z(1)$ is not P-focal, by (23), we have that for all $\lambda \in T_{\dot z(0)}TP^\perp$ it is $\mathrm{d}\exp_{z(0)}(\dot z(0))[\lambda] = 0$ if and only if $\lambda = 0$, i.e., $\mathrm{d}\exp_{z(0)}(\dot z(0))$ is injective on $T_{\dot z(0)}(TP^\perp)$. Since $\dim(TP^\perp) = \dim(\mathcal{M})$, then $\mathrm{d}\exp_{z(0)}(\dot z(0))$ is an isomorphism between $T_{\dot z(0)}(TP^\perp)$ and $T_{z(1)}\mathcal{M}$, hence exp is a local diffeomorphism between a neighborhood of $\dot z(0)$ in $TP^\perp$ and a neighborhood of $z(1)$ in $\mathcal{M}$. ∎

Let A denote the open subset of $T\mathcal{M}$ where the exponential map is defined. For $p \in \mathcal{M}$, $v \in A \cap T_p\mathcal{M}$ and $w \in T_p\mathcal{M}$ we have:

$$\mathrm{d}\exp_p(v)[w] = \mathrm{d}\exp(v)[(0, w)].$$

Moreover, if $z(t) = \exp(tv)$ is the geodesic such that $z(0) = p$ and $\dot z(0) = v$, then $\mathrm{d}\exp_p(v)[v] = \frac{\mathrm{d}}{\mathrm{d}t}\exp(tv)|_{t=1} = \dot z(1)$.

We can now prove the following generalization of the classical *Gauss Lemma* in semi-Riemannian geometry (see [1]). We recall that the Gauss Lemma states that the differential $\mathrm{d}\exp_p$ of the exponential map at a *fixed* point p is a *radial isometry.* In our case, we need to prove a similar property for the exponential map when the basepoint is not necessarily fixed.

Proposition 4.3. *Let* $p \in \mathcal{M}$, $v \in A \cap T_p\mathcal{M}$ *and* $\lambda = (\lambda_h, \lambda_v) \in T_v(T\mathcal{M})$ *be given. Then, it is:*

$$\big\langle \mathrm{d}\exp(v)[(0, v)], \mathrm{d}\exp(v)[\lambda]\big\rangle = \big\langle v, \lambda_h + \lambda_v\big\rangle\,. \tag{24}$$

Proof. Since both sides of (24) are linear in λ, we can consider separately the two cases $\lambda = (w, 0)$ and $\lambda = (0, w)$, with $w \in T_p\mathcal{M}$.

The case of a vertical vector $\lambda = (0, w)$ reduces to the classical Gauss Lemma, and its proof can be found, for instance, in Reference [1, Theorem 10.18].

Let $v(s)$ be a curve in $T\mathcal{M}$ with $v(0) = v$ and $\dot v(0) = \lambda$, i.e., $v(s)$ is a vector field along some curve $p(s)$ in $\mathcal{M}$, with $\dot p(0) = \lambda_h$ and covariant derivative $\frac{\mathrm{D}}{\mathrm{d}s}v(0) = \lambda_v$.

Let's define $H(s, t) = \exp(t \cdot v(s))$. Then, H is defined in a neighborhood of $[0, 1] \times \{0\}$ in $\mathbb{R}^2$. We have:

$$\begin{aligned} &\frac{\partial H}{\partial t}(0,0) = v\,, \qquad &&\frac{\partial H}{\partial t}(1,0) = \mathrm{d}\exp(v)[(0,v)]\,,\\ &\frac{\partial H}{\partial s}(0,0) = \lambda_h\,, &&\frac{\partial H}{\partial s}(1,0) = \mathrm{d}\exp(v)[\lambda]\,. \end{aligned} \tag{25}$$

Suppose now that $\lambda = (w, 0)$ is a given horizontal vector in $T_v(T\mathcal{M})$. Then, $\lambda_v = 0$, and, by (25), we have the thesis once we prove the equality:

$$\big\langle \frac{\partial H}{\partial t}(0,0), \frac{\partial H}{\partial s}(0,0)\big\rangle = \big\langle \frac{\partial H}{\partial t}(1,0), \frac{\partial H}{\partial s}(1,0)\big\rangle\,. \tag{26}$$

Let's prove that $\big\langle \frac{\partial H}{\partial t}, \frac{\partial H}{\partial s}\big\rangle$ does not depend on t: Since $t \longmapsto H(t, s)$ is a geodesic, it is $\frac{\mathrm{D}}{\mathrm{d}t}\,\frac{\partial H}{\partial t} = 0$, and we have:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\big\langle \frac{\partial H}{\partial t}, \frac{\partial H}{\partial s}\big\rangle &= \big\langle \frac{\mathrm{D}}{\mathrm{d}t}\,\frac{\partial H}{\partial t}, \frac{\partial H}{\partial s}\big\rangle + \big\langle \frac{\partial H}{\partial t}, \frac{\mathrm{D}}{\mathrm{d}t}\,\frac{\partial H}{\partial s}\big\rangle = \big\langle \frac{\partial H}{\partial t}, \frac{\mathrm{D}}{\mathrm{d}s}\,\frac{\partial H}{\partial t}\big\rangle = \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\big\langle \frac{\partial H}{\partial t}, \frac{\partial H}{\partial t}\big\rangle\,. \end{aligned}$$

Now, the *length* of the tangent vector to a geodesic is constant, hence $\big\langle \frac{\partial H}{\partial t}, \frac{\partial H}{\partial t}\big\rangle = \big\langle v(s), v(s)\big\rangle$. Since $\lambda = (w, 0)$, we can choose any curve $p(s)$ with $\dot p(0) = w$, and $v(s)$ to be a parallel field along p. It follows that the quantity $\big\langle v(s), v(s)\big\rangle$ is constant, which implies (26) and concludes the proof. □

We are finally ready to prove that the lightlike geodesics between P and Γ can be obtained as limits of timelike geodesics between P and Γ. We will state and prove the result only for the case when Γ is a unidimensional submanifold of $\mathcal{M}$, i.e., when Γ is a timelike curve.

Proposition 4.4. *Let $\Gamma : I\!R \longmapsto \mathcal{M}$ be a (smooth) timelike curve of $\mathcal{M}$, P be any spacelike submanifold of $\mathcal{M}$ and $z : [0,1] \longmapsto \mathcal{M}$ be a lightlike geodesic with $z(0) \in P$, $\dot z(0) \in TP^\perp$ and $z(1) = \Gamma(s_0)$ for some $s_0 \in I\!R$. Assume that $z(1)$ is not a P-focal point along z. Then, there exists a neighborhood V of $z([0,1])$ in $\mathcal{M}$ and a positive number ε_0 such that, for all $\varepsilon \in]-\varepsilon_0, \varepsilon_0[$ there exists a unique timelike geodesic $z_\varepsilon : [0,1] \longmapsto \mathcal{M}$ joining P and Γ, with $\dot z_\varepsilon(0) \in TP^\perp$. Moreover, z_ε tends to z in $C^2([0,1], \mathcal{M})$ as ε tends to 0. If z is future-pointing with respect to Γ, i.e., if $\langle \dot z(1), \dot\Gamma(s_0)\rangle < 0$, then also z_ε is future pointing with respect to Γ for all ε.*

Proof. By Lemma 4.2, since $z(1)$ is not a P-focal point along z, then the exponential map gives a diffeomorphism between a neighborhood of $\dot z(0)$ in $TP^\perp$ and a neighborhood of $z(1)$ in $\mathcal{M}$. Hence, there exists a smooth curve $v :]-\sigma, \sigma[\longmapsto TP^\perp$, $\sigma > 0$, such that $v(0) = \dot z(0)$ and

$$\exp(v(s)) = \Gamma(s_0 + s)\,, \qquad \text{for all } s \in]-\sigma, \sigma[\,. \tag{27}$$

Let $\pi : TP^\perp \longmapsto P$ be the canonical projection and let $p(s)$ be the smooth curve in P given by $p(s) = \pi(v(s))$, so that $p(0) = z(0)$. If we set $v'(0) = \lambda = (\lambda_h, \lambda_v)$, we have $\lambda_h = p'(0)$ and $\lambda_v = \frac{\mathrm{D}}{\mathrm{d}s} v(0)$.

If we differentiate (27) with respect to s and evaluate at $s = 0$, we obtain:

$$\mathrm{d}\exp(\dot z(0))[\lambda] = \dot\Gamma(s_0)\,. \tag{28}$$

We have:

$$\frac{\mathrm{d}}{\mathrm{d}s}\langle v(s), v(s)\rangle\big|_{s=0} = 2\langle \frac{\mathrm{D}}{\mathrm{d}s} v(0), v(0)\rangle = 2\langle \lambda_v, \dot z(0)\rangle.$$

Since $\lambda_v = \dot p(0) \in TP$ and $\dot z(0) \in TP^\perp$, by Proposition 4.3 we have:

$$\begin{aligned} \langle \lambda_v, \dot z(0)\rangle &= \langle \lambda_h + \lambda_v, \dot z(0)\rangle = \langle \mathrm{d}\exp(\dot z(0))[\lambda], \mathrm{d}\exp(\dot z(0))[(0, \dot z(0))]\rangle = \\ &= \langle \dot\Gamma(s_0), \dot z(1)\rangle\,. \end{aligned} \tag{29}$$

Since $\dot z(1)$ is lightlike and $\dot\Gamma(s_0)$ is timelike, then $\langle \dot\Gamma(s_0), \dot z(1)\rangle \neq 0$, from which we get $\frac{\mathrm{d}}{\mathrm{d}s}\langle v(s), v(s)\rangle|_{s=0} \neq 0$. Since, $\langle v(0), v(0)\rangle = \langle \dot z(0), \dot z(0)\rangle = 0$, it follows $\langle v(s), v(s)\rangle < 0$ for either $s > 0$ or $s < 0$ small enough. More precisely, $\langle v(s), v(s)\rangle < 0$ for $s > 0$ if and only if z is future pointing with respect to Γ.

Define $\varepsilon(s) = \sqrt{-\langle v(s), v(s)\rangle}$ and $z_\varepsilon(t) = \exp(t \cdot v(s))$. Clearly, z_ε is a timelike geodesic joining P and Γ, with $\dot z_\varepsilon(0) \in TP^\perp$. Moreover, by the continuity of the function $s \longmapsto \langle \dot z_{\varepsilon(s)}, \dot\Gamma(s_0 + s)\rangle$, the curve z_ε is future pointing with respect to Γ if and only if z is.

When ε tends to 0, $\dot z_\varepsilon(0)$ tends to $\dot z(0)$ in $TP^\perp$, which implies the C^2-convergence of z_ε to z.

This concludes the proof. □

The conclusion that we get from Proposition 4.4 is that the lightlike geodesics between P and Γ that start orthogonally at P can be obtained as limits of critical points for the arrival time functional in $\mathcal{R}^+_{P,\Gamma,\varepsilon}$ for $\varepsilon \to 0$.

This approach to the study of light rays between extended light sources and receivers is the subject of a followup paper.

References

[1] Beem, J. K., Ehrlich, P. E., Easly, K. L., *Global Lorentzian Geometry*, Marcel Dekker, Inc., New York and Basel, 1996.

[2] Bel, L., Martin, J., *Fermat's Principle in General Relativity*, Gen. Rel. Grav. **26** (1994), 567–585.

[3] Giannoni, F., *A Timelike Extension of Fermat's Principle in General Relativity. Comparisons with the Lightlike Case*, Proceedings of the WCNA96, Athens, Nonlinear Analysis 30, vol. 2 (1997), 759–766.

[4] Giannoni, F., Masiello, A., Piccione, P., *A Variational Theory for Light Rays in Stably Causal Lorentzian Manifolds. Existence and Regularity Results*, Commun. Math. Phys. **187**, 375.

[5] Giannoni, F., Masiello, A., Piccione, P., *A Timelike Extension of Fermat's Principle in General Relativity and Applications*, Calculus of Variations and PDE **6**, vol. 3 (1998), 263–283.

[6] Giannoni, F., Masiello, A., Piccione, P., *A Morse Theory for Light Rays in Stably Causal Lorentzian Manifolds*, Ann. Inst. H. Poincaré (Physique Theorique) 69, vol. 4 (1998), 359–412.

[7] Giannoni, F., Masiello, A., Piccione, P., *A Morse Theory for Massive Particles and Photons in General Relativity*, preprint 1998.

[8] Hawking, S. W., Ellis, G. F., *The Large Scale Structure of Space–Time*, Cambridge Univ. Press, London, New York, 1973.

[9] Kovner, I., *Fermat Principle in Arbitrary Gravitational Fields*, Astrophysical Journal **351** (1990), 114–120.

[10] Lang, S., *Differential Manifolds*, Springer-Verlag, Berlin, 1985.

[11] Levi–Civita, T., *Fondamenti di Meccanica Relativistica*, Zanichelli, Bologna, 1928 (new version 1986).

[12] O'Neill, B., *Semi–Riemannian Geometry with Applications to Relativity*, Academic Press, New York, 1983.

[13] Palais, R., *Foundations of Global Nonlinear Analysis*, W. A. Benjamin, 1968.

[14] Perlick, V., *On Fermat's Principle in General Relativity: I. The General Case*, Class. Quantum Grav. **7** (1990), 1319–1331.

[15] Perlick, V., *On Fermat's Principle in General Relativity: I. The Conformally Stationary Case*, Class. Quantum Grav., **7** (1990), 1849.

[16] Perlick, V., *Infinite-dimensional Morse Theory and Fermat's Principle in General Relativity*, J. Math. Phys. **36** (1995), 6915–6928.

[17] Perlick, V., Piccione, P., *A General–Relativistic Fermat's Principle for Extended Light Sources and Extended Receivers*, Gen. Rel. Grav. 30, vol. 10 (1998), 1461–1476.

[18] Rindler, W., *Essential Relativity*, Springer, New York, 1977.

[19] R. Sachs, R., Wu, H., *General Relativity for Mathematicians*, Springer, New York, 1977.

[20] Schneider, P., Ehlers, J., Falco, E., *Gravitational Lenses*, Springer, New York 1992.

Departamento de Matemática
Universidade de São Paulo, BRAZIL
E-mail: piccione@ime.usp.br