BRIEF REPORTS

Brief Reports are accounts of completed research which do not warrant regular articles or the priority handling given to Rapid Communications; however, the same standards of scientific quality apply. (Addenda are included in Brief Reports.) A Brief Report may be no longer than four printed pages and must be accompanied by an abstract.

Fermat's principle in general relativity

Rajaram Nityananda and Joseph Samuel
Raman Research Institute, Bangalore 560 080, India

(Received 20 November 1991)

Kovner has observed that Fermat's principle can be used to describe the motion of light rays in arbitrary gravitational fields, not just stationary ones. We give a simple demonstration of this fact.

PACS number(s): 04.20.Cv, 04.20.Fy

I. INTRODUCTION

Fermat's principle states that a light ray going from point A to B in space takes the path of least time. More precisely, the time taken to go from A to B is stationary with respect to small variations in the path. This principle provides an economical restatement of the laws of propagation, reflection, and refraction of light. As stated above however, Fermat's principle is firmly rooted in Newtonian physics, with its absolute notions of space and time. These notions do not survive the passage to general relativity (GR). In relativity, the time between two events depends on the reference frame used. It is not therefore clear how to formulate Fermat's principle in GR. The path followed by light rays is an intrinsic property of spacetime, whereas the time taken is an observer-dependent notion.

Can one describe the motion of light rays in curved spacetime using Fermat's principle? The literature in GR does address this question for static [1–3] and stationary [4–6] spacetimes. In these spacetimes, there is a timelike Killing vector field, which can be used to define preferred notions of "space" and "time." A point in space is an entire integral curve of the Killing vector field. Time is defined as the Killing parameter [7], or more plainly as the coordinate time in a stationary reference system. Given these notions of space and time, Fermat's principle applies [1–6] and describes the behavior of light rays, exactly as in Newtonian physics. However, a general spacetime need not be stationary. The question remains: Does Fermat's principle describe light rays in an arbitrary gravitational field?

Kovner [8] has recently answered this question by appropriately formulating Fermat's principle in GR. He was motivated by the application of Fermat's principle to the theory of gravitational lensing. Kovner's formulation follows.

Suppose that light is emitted at an event p in spacetime and received by an observer $\mathcal{O}$ at some event q on his world line $\mathcal{L}$. Denote by $\mathcal{N}$ the set of null curves starting at p and reaching $\mathcal{L}$. In general relativity, light rays follow null geodesics. Fermat's principle in GR states that these are characterized by the following property: Their arrival on $\mathcal{L}$ is stationary with respect to first order variations of the null curve within $\mathcal{N}$.

Several comments are in order at this point.

(1) Note that this formulation makes no reference to "space" or "time" and so applies in a general nonstationary spacetime.

(2) In this formulation, not all variations are permitted: The varied curve must also be null, start from p and reach $\mathcal{L}$. More concisely, the varied curve must also belong to $\mathcal{N}$. Henceforth, we will refer to such variations as "allowed" variations. Requiring that the varied curve be null is in complete analogy with the classical formulation of Fermat's principle: In computing the time taken for light to traverse a path in space, one does require that the rays travel at the local speed of light.

(3) The quantity which is stationary in this formulation is the arrival time of the light ray, as measured by the observer $\mathcal{O}$. This could be his proper time, or some parameter increasing along his world line. One can however take a more geometric stance and say that the *arrival* of the light ray, which is an event on the observer's world line, is stationary.

While we believe that Kovner's formulation is correct and interesting, there does seem to be scope for an improved treatment. In proving Fermat's principle, one encounters the technical problem of ensuring that the varied curves are in $\mathcal{N}$. Kovner solves this by explicitly constructing "zig zag" paths. He also occasionally regards null curves as the massless limit of timelike curves. This is a singular limit and may be dangerous. Such an elegant final result ought to be derivable directly for null curves using the standard apparatus of relativity. Our alternative presentation possibly brings out more clearly the essentially geometric nature of the result. In Sec. II we prove that the arrival of a null geodesic is stationary with respect to "allowed" variations. In Sec. III we prove the converse. Section IV is a brief concluding discussion.

II. FERMAT'S PRINCIPLE

We first prove that if γ is a null geodesic from p to the timelike world line $\mathcal{L}$, a first-order "allowed" variation in γ leads to a second- (or higher-)order variation in the arrival of the curve on $\mathcal{L}$. Consider a one-parameter family of null curves γ_s from p to $\mathcal{L}$ such that $\gamma_0=\gamma$. Choose a parameter λ along these curves so their points $\gamma_s(\lambda)$ are labeled by (s,λ) and $\gamma_s(0)=p$ and $\gamma_s(1)$ lies on $\mathcal{L}$. On the two-dimensional region $\gamma_s(\lambda)$, $0\le\lambda\le 1$, $-\epsilon\le s\le\epsilon$ there are defined vector fields.[1]

$$\boldsymbol{l}=\partial/\partial\lambda, \quad \mathbf{v}=\partial/\partial s \ .$$

$\boldsymbol{l}$ is the tangent vector[2] to the curve γ_s and $\mathbf{v}$ is a vector denoting the first-order variation of the null curve. Since the vector fields $\partial/\partial s$ and $\partial/\partial\lambda$ commute, their Lie brackets vanish:

$$[\boldsymbol{l},\mathbf{v}]=\nabla_l\mathbf{v}-\nabla_v\boldsymbol{l}=0 \ , \tag{2.1}$$

where $\nabla_v\boldsymbol{l}$ is $v^a\nabla_a l^b$, the covariant directional derivative of $\boldsymbol{l}$ along $\mathbf{v}$.

Since the varied curve must also be in $\mathcal{N}$, the vector field $\mathbf{v}$ which generates the variation is subject to constraints. The varied curve must start from p,

$$\mathbf{v}(0)=0 \ , \tag{2.2}$$

and arrive on $\mathcal{L}$:

$$\mathbf{v}(1)\propto\mathbf{t} \ , \tag{2.3}$$

where $\mathbf{t}$ is the (future-pointing) unit vector at q tangential to $\mathcal{L}$. The varied curve must also be a null curve: i.e.,

$$\boldsymbol{l}\cdot\boldsymbol{l}=0 \tag{2.4}$$

for all s. This condition is expressed to first order as

$$\frac{\partial}{\partial s}(\boldsymbol{l}\cdot\boldsymbol{l})=\nabla_v(\boldsymbol{l}\cdot\boldsymbol{l})=0=2\boldsymbol{l}\cdot\nabla_v\boldsymbol{l} \tag{2.5}$$

or using (2.1), as

$$\boldsymbol{l}\cdot\dot{\mathbf{v}}=0 \ , \tag{2.6}$$

where the overdot here and below denotes the directional covariant derivative ∇_l along $\boldsymbol{l}$.

It is assumed that $\gamma=\gamma_0$ is a geodesic. For $s=0$,

$$\dot{\boldsymbol{l}}=\rho(\lambda)\boldsymbol{l} \ , \tag{2.7}$$

for some function $\rho(\lambda)$. (Hereafter s is set to zero.) Differentiating (2.4) yields

$$\boldsymbol{l}\cdot\dot{\boldsymbol{l}}=0 \tag{2.8}$$

and

$$\boldsymbol{l}\cdot\ddot{\boldsymbol{l}}=-\dot{\boldsymbol{l}}\cdot\dot{\boldsymbol{l}} \ . \tag{2.9}$$

We wish to show that $\mathbf{v}(1)$ (which generates the first-order change in the arrival of γ on $\mathcal{L}$) vanishes. It is simpler to deal instead with the scalar quantity

$$\psi(\lambda):=\boldsymbol{l}\cdot\mathbf{v}$$

(evaluated at $s=0$, of course). $\psi(0)=0$, since $\mathbf{v}$ vanishes (2.2) at p. And

$$\dot{\psi}(\lambda)=\dot{\boldsymbol{l}}\cdot\mathbf{v}+\boldsymbol{l}\cdot\dot{\mathbf{v}}=\rho(\lambda)\psi$$

because γ_0 is a geodesic (2.7) and $\mathbf{v}$ is an "allowed" variation (2.6). It follows from the differential equation for $\psi(\lambda)$ and the initial condition that

$$\psi(1)=\boldsymbol{l}(1)\cdot\mathbf{v}(1)=0 \ .$$

Since $\mathbf{v}(1)$ is timelike (2.3) and $\boldsymbol{l}$ is a nonzero null vector, $\mathbf{v}(1)$ must vanish:

$$\mathbf{v}(1)=0 \ . \tag{2.10}$$

This proves that the arrival of null geodesics on $\mathcal{L}$ is stationary with respect to "allowed" variations.

III. CONVERSE

We now prove the converse result: any null curve from p to $\mathcal{L}$ whose arrival at $\mathcal{L}$ is stationary with respect to arbitrary "allowed" variations is a geodesic. Let us parallel transport $\mathbf{t}$ in (2.3) along γ and so define a vector field $\mathbf{t}(\lambda)$ all along γ satisfying

$$\dot{\mathbf{t}}=0 \ .$$

Consider the following choice for $\mathbf{v}$:

$$\mathbf{v}=a(\lambda)\dot{\boldsymbol{l}}+\beta(\lambda)\mathbf{t} \ . \tag{3.1}$$

In order that $\mathbf{v}$ generate an "allowed" variation, it must satisfy (2.2), (2.3). These conditions are met by choosing a,β such that

$$a(0)=a(1)=\beta(0)=0 \ . \tag{3.2}$$

Condition (2.6) is satisfied by choosing β as

$$\beta(\lambda)=-\int_0^\lambda\frac{a(\lambda')(\boldsymbol{l}\cdot\ddot{\boldsymbol{l}})}{(\boldsymbol{l}\cdot\mathbf{t})}d\lambda'=\int_0^\lambda\frac{a(\lambda')(\dot{\boldsymbol{l}}\cdot\dot{\boldsymbol{l}})}{(\boldsymbol{l}\cdot\mathbf{t})}d\lambda' \ ,$$

which is well defined because $\boldsymbol{l}\cdot\mathbf{t}$ is strictly positive.

We are given that for *all* $\mathbf{v}$ satisfying (2.2), (2.3), (2.6), $\mathbf{v}(1)$ or equivalently, $\beta(1)$ vanishes. It follows that, for all a satisfying (3.2),

$$\beta(1)=\int_0^1\frac{a(\lambda)\dot{\boldsymbol{l}}\cdot\dot{\boldsymbol{l}}}{(\boldsymbol{l}\cdot\mathbf{t})}d\lambda=0 \ .$$

This implies $\dot{\boldsymbol{l}}\cdot\dot{\boldsymbol{l}}=0$, i.e., $\dot{\boldsymbol{l}}$ is a null vector. But $\dot{\boldsymbol{l}}$ is orthogonal to $\boldsymbol{l}$(2.8). Therefore $\dot{\boldsymbol{l}}$ must be proportional to $\boldsymbol{l}$, which proves that γ is a geodesic.

IV. CONCLUSION

We have given an elementary proof of Fermat's principle in general relativity. While Fermat's principle for stationary metrics has been known for a long time, the general version has only recently been formulated by Kovner. It is perhaps not widely known, even in the rela-

[1]For those used to more explicit display of indices, we briefly clarify the notation used below, which follows [7]: $\mathbf{v}\cdot\boldsymbol{l}=v^a l_a, \nabla_l\mathbf{v}=l^a v^b{}_{;a}$. The tangent vector to the curve $x^a(\lambda)$ with components $\partial x^a/\partial\lambda$ is denoted $\partial/\partial\lambda$.

[2]We require the parametrization to be such that $\boldsymbol{l}$ is nonvanishing.

tivity community, that such a formulation exists.

While our primary interest is in the description of light rays in a gravitational field, there is also [8] a version of Fermat's principle for massive particles. In that case (as the reader can infer from a careful reading of [8], only those variations are "allowed" that do not change the total arc length of γ. In the limit that γ is null, i.e., has zero length, this coincides with Fermat's principle for light. The proof given above for the massless case also goes through in the massive case with appropriate modification. Since the arc lengths of all the timelike curves γ_s are assumed to be the same, one can choose λ to be the fraction of arc length traversed along the curve and achieve $l \cdot l = 1$, from which follows (2.8) and finally (2.10). Conversely, the same arguments given above show that $\dot{l}$ is a null vector orthogonal (2.8) to the timelike vector l and therefore vanishes. This proves that γ is an (affinely parametrized) geodesic.

The geodesic equation for massive particles is usually derived from the stationarity of the arc length, keeping the end points fixed. It is appealing that in the massive case one can either extremize the arc length, keeping the end points fixed, or extremize one end point (the arrival), keeping the arc length fixed. For the massless case only the second option is available because the first variational principle breaks down. A general variation of a null curve will have timelike and spacelike segments. And the integral of the square root of $l \cdot l$ along the path is not differentiable with respect to such variations.

Fermat's principle for light appears closely related to causality. The boundary of the causal future[3] of p is ruled by null geodesics. If the arrival of these null geodesics on a timelike world line $\mathcal{L}$ were not stationary, one could so choose the variation as to make the varied null curve arrive earlier than the null geodesic. Since the null geodesic belonged to the boundary of the causal future of p, the varied null curve arrives outside the causal future of p. This is a clear violation of causality. While causality and Fermat's principle are related, they are also distinct. For, Fermat's principle says nothing about the sign of the second- (and higher-)order variations in the arrival time. The principle of causality (applied to null geodesics ruling the boundary of a causal future) implies not only that the first-order variation in arrival of null curves must vanish, but also that the higher-order variations must keep the varied curve within the causal future of p. Further, not all null geodesics emanating from p rule the boundary of its causal future. Null geodesics which have focused enter into the interior of the causal future. The principle of causality imposes no restriction on the variation of arrival of these geodesics. But Fermat's principle still applies and requires the first variation to vanish.

[3]More precisely, the difference between the causal and chronological futures [7] of p.

As is well known, null geodesics are sensitive only to the conformal structure of the spacetime. However, the null geodesic equation is not manifestly invariant under conformal transformations. An appealing feature of Fermat's principle for light in relativity is that it too depends only on the conformal structure of the spacetime. The requirements that γ_s be null and $\mathcal{L}$ be timelike and the arrival be stationary are all invariant under conformal transformations. From this it is manifestly clear that null geodesics are conformally invariant. It is perhaps slightly unsatisfactory that our proof of Fermat's principle (like the null geodesic equation) does use more than the conformal structure of spacetime. But it remains true that Fermat's principle gives us a conformally invariant principle to describe the motion of light rays in curved spacetime.

Note added. After this paper was submitted for publication, we learned of Ref. [9], which also deals with the same subject.

[1] W. Pauli, in *Theory of Relativity* (Pergamon, Oxford, (1958).
[2] H. Weyl, Ann. Phys. **54**, 117 (1917).
[3] C. W. Misner, K. S. Thorne, and J. A. Wheeler, in *Gravitation* (Freeman, San Francisco, 1973).
[4] Pham Mau Quan, Les Theories Relativistes de la Gravitation, Royaumont Conference Proceedings, Paris, 1962 (unpublished), p. 165.
[5] D. Brill, in *Relativity, Astrophysics and Cosmology,* edited by W. Israel (Reidel, Dordrecht, 1973), p. 134.
[6] L. D. Landau and E. M. Lifshitz, *Classical Theory of Fields* (Pergamon, Oxford, 1975).
[7] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).
[8] I. Kovner, Astrophys. J. **351**, 114 (1990).
[9] V. Perlick, Class. Quantum Grav. **7**, 1319 (1990).