PHYSICAL REVIEW D VOLUME 48, NUMBER 8 15 OCTOBER 1993

Relation between physical and gravitational geometry

Jacob D. Bekenstein*
Department of Physics, University of California at Santa Barbara, Santa Barbara, California 93106
and The Racah Institute of Physics, Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel†
(Received 19 November 1992)

The appearance of two geometries in a single gravitational theory is familiar. Usually, as in the Brans-Dicke theory or in string theory, these are conformally related Riemannian geometries. Is this the most general relation between the two geometries allowed by physics? We study this question by supposing that the physical geometry on which matter dynamics takes place could be Finslerian rather than just Riemannian. An appeal to the weak equivalence principle and causality then leads us to the conclusion that the Finsler geometry has to reduce to a Riemann geometry whose metric, the physical metric, is related to the gravitational metric by a generalization of the conformal transformation involving a scalar field.

PACS number(s): 04.50.+h

I. INTRODUCTION

The excellent description provided by special relativity of elementary particle phenomena is usually taken to imply that spacetime is described by a Riemannian geometry. This is because special relativity implies a Minkowski geometry for spacetime, but as shown by Schild [1], the experimental existence of the gravitational redshift makes it impossible for the Minkowski geometry to apply globally. An obvious way to mesh the Minkowski geometries at various points is to have a global Riemannian geometry which the Minkowski geometry of elementary particle physics is tangent to at each spacetime event. This is the situation in a one-geometry description of physics, e.g., general relativity (GR).

However, physics may not be that simple: Gravitation may naturally require two geometries for its description. Two geometries in a single theory made their debut in Nordström's 1913 gravitational theory [2] which preceded GR. As in Nordström's theory, and in theories such as Brans-Dicke theory [3], the variable mass theory [4], Dirac's theory of a variable gravitational constant [5], string theories [6], and many others, two conformally related geometries appear. Usually one of these describes gravitation while the other defines the geometry in which matter plays out its dynamics. The strong equivalence principle is violated by all these two-geometries theories, but they usually preserve weak equivalence. Theories of these sort have been of great value in clarifying the foundations of gravitation theory.

Thus, the two-geometries approach to the formulation of gravitational theory is an important paradigm. Whenever it becomes necessary to formulate a new theory of gravity, a conservative way to proceed in order to avoid immediate conflict with the tests of GR is to invoke a Riemannian metric $g_{\alpha\beta}$, build the Einstein-Hilbert action for the geometry's dynamics out of it, and effect the departure from standard GR by prescribing the relation between $g_{\alpha\beta}$ and the physical geometry on which matter propagates. Most known theories assume the relation is a simple conformal transformation.

However, the conformal transformation is but the simplest way to relate two geometries. Might the relation between gravitational and physical geometries be more complicated? In other words, within the two-geometries paradigm for gravitational theory, what are the most general theories that may be envisaged? To answer this question we consider physical geometry of the most general kind that might be of interest physically. For this purpose we have to discard conformal and affine geometries because they both lack the physically essential notion of distance. We are thus left with Finsler geometry as a possible substitute for Riemann geometry in describing the physical arena in which nongravitational physics takes place.

Finsler geometry, introduced by Riemann and first studied systematically by Finsler, is the most general geometry in which the squared line element is homogeneous of second degree in the coordinate increments [7]:

$$ds^2 = f(x^\alpha, dx^\beta), \qquad f(x^\alpha, \mu dx^\beta) = \mu^2 f(x^\alpha, dx^\beta). \quad (1)$$

Whenever $f(x^\alpha, dx^\beta)$ is a quadratic form in the dx^β, the geometry is Riemannian. Otherwise, we have a Finsler geometry as the playground of matter dynamics. We shall show presently that this state of affairs cannot be ruled out at the outset by the argument about the Minkowski geometry of elementary particle physics.

Although Finsler geometry is quite different from Riemann geometry, it is possible to introduce a metriclike tensor for it. If in the second of Eqs. (1) we replace $\mu \to 1+\epsilon$ where ϵ is an infinitesimal, expand in ϵ, and

*Electronic mail: bekenste@vms.huji.ac.il
†Permanent address.

focus on the terms of $O(\epsilon^2)$, we have (this is a facet of Euler's theorem)

$$\mathcal{G}_{\alpha\beta}\, dx^\alpha dx^\beta \equiv \frac{1}{2}\frac{\partial^2 f}{\partial dx^\alpha \partial dx^\beta}\, dx^\alpha dx^\beta = f. \tag{2}$$

Thus we may express the line element as [8]

$$ds^2 = \mathcal{G}_{\alpha\beta}\, dx^\alpha dx^\beta, \tag{3}$$

and think of $\mathcal{G}_{\alpha\beta}$ as a kind of metric. However, it must be remembered that $\mathcal{G}_{\alpha\beta}$ itself depends on dx^α. Because of this difference from the Riemann metric, we call $\mathcal{G}_{\alpha\beta}$ the quasimetric. Our arguments will make heavy use of it.

II. SPIRIT OF COVARIANCE

Because in the theory being discussed there is already a symmetric tensor $g_{\alpha\beta}$, we may rewrite Eq. (1) in general as

$$ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta\, \mathcal{F}(x^\alpha, dx^1/dx^0, dx^2/dx^0, dx^3/dx^0). \tag{4}$$

This is because the expressions such as dx^1/dx^0 are the only independent combinations of the coordinate increments which are homogeneous of degree zero in dx^α.

As it stands, Eq. (4) is still the most general Finsler geometry. However, there is something about it which does violence to the spirit of the principle of covariance. Suppose we make a general coordinate transformation. We know that the form $g_{\alpha\beta}\, dx^\alpha dx^\beta$ is invariant. This is achieved by the components of $g_{\alpha\beta}$ changing in an appropriate way so that at a fixed spacetime point the metric is described by ten numbers of which four may be chosen arbitrarily. In other words, covariance requires that four numbers be free at any spacetime point. Since we also want invariance of ds^2, not just of $g_{\alpha\beta}\, dx^\alpha dx^\beta$, it is plain that the form of the function $\mathcal{F}$ cannot be invariant. It will vary with coordinate systems in such a way as to compensate for the transformations of the ratios dx^i/dx^0 into rational functions of themselves.

Not only is this ugly, but it also means that the freedom inherent in coordinate transformations is, at a fixed point in spacetime, not just that in four numbers, but rather that in a function of three variables. Although the letter of the principle of covariance is still obeyed by having this free function $\mathcal{F}$, it would seem that the spirit of the principle is violated. The vast freedom engendered by coordinate transformations would seem to empty the principle of any physical content.

One may recover the situation where only a few quantities are free at a point by confining attention to a function F of coordinate invariants alone. However, out of the invariant $g_{\alpha\beta}\, dx^\alpha dx^\beta$ alone one can form only one homogeneous function of second order in the dx^α: the invariant itself. This would bring us back to the situation where the physical metric is identical to the gravitational one. To go beyond triviality one needs more invariants. In general this requires the introduction of new fields. First suppose there exists a dimensionless scalar field ψ (such as the dilaton in many contemporary field theories) and a length scale L, e.g., the Planck length. Then a nontrivial Finsler line element may be written

$$ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta\, F(I, H, \psi), \tag{5}$$

$$I \equiv L^2 g^{\alpha\beta}\psi_{,\alpha}\,\psi_{,\beta}\,, \tag{6}$$

$$H \equiv \frac{L^2\,(\psi_{,\alpha}\, dx^\alpha)^2}{-g_{\alpha\beta}\, dx^\alpha dx^\beta}. \tag{7}$$

Note that both F and its arguments I, H, and ψ are dimensionless. We now note that covariance of Eq. (5) is to be had at the same price as that for ordinary Riemannian geometry: four free metric components at every point in spacetime. The function F is fixed, one and the same for all coordinate systems.

Of course, we could have added further arguments to F constructed out of second and higher derivatives of ψ. We refrain from this in order to preclude higher derivative terms from entering in the matter equations of motion (after all the matter action will be built on the line element ds^2).

The introduction of additional scalar fields and invariants constructed from them as arguments of F is likewise not logically excluded. However, given that ψ is a field with a special status (building block of the physical geometry), simplicity requires that we abstain from multiplying such entities. One may also consider introduction of invariants built with the help of additional vector and tensor fields v_α and $t_{\alpha\beta}$, e.g., $(v_\alpha dx^\alpha)^2/g_{\alpha\beta}dx^\alpha dx^\beta$. Again, such a move can be criticized on the grounds of economy: A vector field would add four more fields to the theory, and a tensor field at least ten more. Furthermore, once the inclusion in the theory of $t_{\alpha\beta}$ is permitted, it is simplest from the point of view of the two geometries paradigm to just identify the symmetric part of $t_{\alpha\beta}$ as the Riemann metric of the physical geometry and forego the appeal to Finsler geometry. However, gravitational theories based on two independent metrics have never fared well in the confrontation with experiment [9]. In view of all this, we conclude that the line element in Eq. (5) is the most general that may be constructed with a minimum of elements: gravitational metric and a scalar field.

It is consistent with all our previous discussion to postulate that the classical trajectories of free particles are those which extremize the action

$$S = \frac{1}{2}\int g_{\alpha\beta}\, \dot{x}^\alpha \dot{x}^\beta\, F(I, H, \psi)\, d\lambda. \tag{8}$$

Here by H we mean the expression in Eq. (7) with $dx^\alpha \to \dot{x}^\alpha \equiv dx^\alpha/d\lambda$ (λ is a parameter along the trajectory). Equation (8) is the straightforward generalization to Finsler geometry of the action used in GR for classical particles (in a form not invariant under changes of parameter λ). It is easy to see that a trajectory with

$ds = 0$ all along it automatically extremizes S [which is simply the integral of $(ds/d\lambda)^2$]. Thus in this theory, just as in GR, null curves are automatically trajectories of free particles. We do not require that the trajectories which extremize S in the Finsler geometry coincide with the geodesics of $g_{\alpha\beta}$. There is no physical basis for such an assumption in our context: The metric $g_{\alpha\beta}$ is for gravitational phenomena, whereas the Finsler geometry is for matter dynamics.

Before passing on let us rebut the argument that infers a global Riemannian geometry from the Minkowski geometry of elementary particle physics. Consider Eq. (5). For general F the line element certainly does not look like one that could locally be brought to Minkowski form by a coordinate transformation. However, suppose $F(I, H, \psi)$ is regular and nonvanishing in the limit $I \to 0$ and $H \to 0$. Then in a region where the ψ field varies slowly (presumably the solar system is like that), the line element is seen to be of the form $ds^2 \approx F(0,0,\psi)\, g_{\alpha\beta}\, dx^\alpha dx^\beta$ which corresponds to a Riemann geometry. Therefore, under everyday circumstances some Finslerian geometries can masquerade as Riemannian ones, and can thus be consistent with the evidence from elementary particle physics.

III. FORM OF THE FINSLER FUNCTION F

As long as all we have to deal with is classical particle motion, Finsler geometry seems perfectly adequate as the physical geometry. However, classical physics also involves field equations, i.e., Maxwell's, and once we enter into quantum physics even particles must be discussed in terms of wave (field) equations. But it is unclear how to formulate the familiar field equations on a general Finsler geometry. The problem is that to formulate the typical field equation (think of the scalar equation for concreteness), one requires a contravariant "metric" to raise indices of derivatives and so form divergences. To put it another way, to form a scalar action out of derivatives of fields, one requires an object capable of raising indices. In Riemannian geometry $g^{\alpha\beta}$ serves this purpose. If we try to use the inverse of the Finsler quasimetric, $\mathcal{G}^{\alpha\beta}$, we will find in general that it depends on the increments dx^α and not just on the spacetime point. Clearly, field equations constructed with $\mathcal{G}^{\alpha\beta}$ would be meaningless in general.

One way out of this problem is to confine attention to geometries for which $\mathcal{G}_{\alpha\beta}$ is independent of dx^α; this guarantees that $\mathcal{G}^{\alpha\beta}$ will be too. Let us use definition (2) to write the quasimetric for the Finsler geometry of Eq. (5). The result is

$$\mathcal{G}_{\alpha\beta} = (F - HF')g_{\alpha\beta} - L^2\,(F' + 2HF'')\psi_{,\alpha}\,\psi_{,\beta} -2H^2F''\Big[\frac{\psi_{,(\alpha}\, g_{\beta)\mu}dx^\mu}{\psi_{,\mu}\, dx^\mu} - \frac{g_{\alpha\mu}dx^\mu\, g_{\beta\nu}dx^\nu}{g_{\mu\nu}dx^\mu dx^\nu}\Big], \quad (9)$$

where a prime denotes a partial derivative with respect to H and parentheses around subscripts denote symmetrization. This $\mathcal{G}_{\alpha\beta}$ will be independent of dx^α only if F'' vanishes, i.e., if

$$F = A(I,\psi) - B(I,\psi)H, \quad (10)$$

with A and B two dimensionless functions of the shown arguments. With an F like this it is possible to construct the familiar field equations.

When we substitute Eq. (10) into Eq. (5) we find that the quasimetric reduces to a Riemann metric $\tilde{g}_{\alpha\beta}$:

$$ds^2 = \tilde{g}_{\alpha\beta}\, dx^\alpha dx^\beta \equiv (Ag_{\alpha\beta} + L^2\, B\, \psi_{,\alpha}\, \psi_{,\beta}\,)\, dx^\alpha dx^\beta. \quad (11)$$

But here the relation between the gravitational metric $g_{\alpha\beta}$ and the physical metric $\tilde{g}_{\alpha\beta}$ is more complex than via a conformal transformation.

Might not a more general relation between physical and gravitational geometry be possible than that we have exhibited? Perhaps reasonable matter field equations can be formulated on the basis of some structure other than the quasimetric. For example, consider a Finsler geometry determined by a symmetric fourth-rank tensor [10]:

$$ds^4 = \mathcal{E}_{\alpha\beta\gamma\delta}\, dx^\alpha dx^\beta dx^\gamma dx^\delta. \quad (12)$$

We assume $\mathcal{E}_{\alpha\beta\gamma\delta}$ is not degenerate; i.e., it cannot be written as $q_{\alpha\beta}\, q_{\gamma\delta}$. Suppose we use the inverse tensor $\mathcal{E}^{\alpha\beta\gamma\delta}$ in lieu of a metric to construct invariant field Lagrangians. For a scalar field Φ the simplest choice for the Lagrangian $\mathcal{L} = \mathcal{E}^{\alpha\beta\gamma\delta}\, \Phi_{,\alpha}\, \Phi_{,\beta}\, \Phi_{,\gamma}\, \Phi_{,\delta}$ is quartic in the field. It cannot lead to a linear field equation. And if we take the square root of the above invariant as the Lagrangian, linearity is still out of reach. The full symmetry of $\mathcal{E}^{\alpha\beta\gamma\delta}$ also prevents us from forming a Lagrangian for an antisymmetric Maxwell field $F_{\alpha\beta}$ which is quadratic in the field. A fourth-order invariant can be built, but even if we use its square root as the Lagrangian, we cannot obtain linear equations. Needless to say, we cannot do without linear equations in physics. And in the more general case when ds^n is given by an nth-order form in dx^α with $n \geq 3$, the problems mentioned will persist. This discussion underscores the implausibility of a viable Finsler (but non-Riemannian) physical geometry for matter dynamics. So far we have only been able to make do with the special linear Finsler function of Eq. (10) which is equivalent to a Riemannian metric.

Of course it is possible that a Finsler geometry picks out a certain Riemann geometry more complex than $\tilde{g}_{\alpha\beta}$ as special. If so we might contemplate using that metric to construct field equations for matter. In that case one would have to check whether the physics is consistent with the Finsler geometry being the arena for matter dynamics. We shall now explore this program.

Our principal tool will be considerations of causality. In order to be clear, let us introduce the terms "graviton" and "photon" with very specific meaning. We have agreed that $g_{\alpha\beta}$ is the metric which the Einstein-Hilbert action is written with. This means that the characteristics of the Einstein-like equations which govern gravitational dynamics in the envisaged theory must lie on the null surfaces of the metric $g_{\alpha\beta}$. Short wavelength perturbations of $g_{\alpha\beta}$ will thus propagate on these null surfaces. We call these gravitons (no quantum connotation implied).

By photons we mean short wavelength excitations of matter fields such as the scalar field, the Maxwell field, or the Weyl neutrino field. In GR these travel on the null cone of the metric. We shall adopt this wisdom as an axiom of the present theory and take it to mean that if viewed as a point particle, a photon follows a trajectory in the Finsler geometry with $ds^2 = 0$. We have seen in Sec. II that such trajectories correspond to free particles in the theory. The totality of such trajectories passing through a point in spacetime defines the physical light cone at that event. Note that we do not assume that the null surfaces of $g_{\alpha\beta}$ coincide with the physical light cones as would be the case in theories where the geometries are necessarily conformally related. Thus we do not assume that $F > 0$ everywhere as is often done in studies of Finsler geometry [8]. We do assume, and this is the content of the causality principle, that all physical particles travel on trajectories with $ds^2 \leq 0$. It is still true here that nothing travels faster than light.

Note that the point $H = +\infty$ of the geometry is to be identified with $H = -\infty$. This is because the passage from one to the other corresponds to $g_{\alpha\beta}\, dx^\alpha dx^\beta$ passing through zero from negative to positive values. Therefore, the line element ds^2 should be continuous as H jumps from $+\infty$ to $-\infty$, so that as $H \to \pm\infty$, F must either be bounded or blow up no faster than linearly with H. If F blows up linearly, the coefficient of H must be identical in both limits to preclude a discontinuity in ds^2. We shall discount the possibility that F can blow up slower than H because this would entail nonanalytic behavior, e.g., $F \sim H^{1/3}$.

A. Finsler function F with no zeros

Suppose F is of one sign throughout with no zeros in the finite H axis. We take its sign positive by convention and refer to this case as case A. Now the light cone is delineated by coordinate increments which make $g_{\alpha\beta}\, dx^\alpha dx^\beta = 0$. But in order for ds^2 to actually vanish for such increments, F is not allowed to blow up (even just linearly with H) as $H \to \pm\infty$ for in that case the zero of $g_{\alpha\beta}\, dx^\alpha dx^\beta$ would be canceled out.

Now massive particles (not photons) follow trajectories with $ds^2 < 0$. So since $F > 0$, Eqs. (5) and (7) tell us that $g_{\alpha\beta}\, dx^\alpha dx^\beta < 0$ and $H > 0$ for the corresponding trajectories. Thus physical trajectories fill the whole positive H axis, with photon trajectories lying at $H = \pm\infty$ at which point F must be bounded.

Suppose F does not tend to zero as $H \to +\infty$. Clearly we must require $F(I, +\infty, \psi) = F(I, -\infty, \psi)$ so that the line element does not jump between $H = +\infty$ and $H = -\infty$. Then the graviton null surface coincides with the physical light cone. In fact, near the graviton null surface $ds^2 \approx F(I, \pm\infty, \psi)\, g_{\alpha\beta}\, dx^\alpha dx^\beta$; i.e., the Finsler geometry induces a Riemann geometry near the light cone because the conformal factor is independent of dx^α there.

Following our program let us construct the matter physics, i.e., Maxwell's equations, the gauge field equations, Weyl's equation, etc., using the effective metric

$$\tilde{g}_{\alpha\beta} \equiv F(I, \infty, \psi)\, g_{\alpha\beta} \tag{13}$$

that has been picked out as special by the Finsler geometry. It is clear that short wavelength solutions of these field equations will propagate on the light cone as defined above simply because their characteristics coincide with the null surface of the metric used to build them. Photons will thus travel on the light cone and so their trajectories will extremize the action S as required of classical particles moving in the Finsler physical geometry. Thus we reach a consistent picture of photon dynamics.

But for massive particles a dichotomy appears. These might be described by the Dirac equation with nonzero rest mass or the massive Klein-Gordon equation. In order that the weak equivalence principle be satisfied, let these field equations be formulated with the same metric $\tilde{g}_{\alpha\beta}$ as used for the other fields. The classical trajectories corresponding to a field equation may be inferred, say, from the the Hamilton-Jacobi equation that results from the eikonal approximation to the field equation. Working out this procedure for the massive Klein-Gordon equation shows the trajectories to be geodesics of the effective metric $\tilde{g}_{\alpha\beta}$.

Unless some very special conditions are satisfied [10,11], these will not be geodesics of the Finsler geometry, i.e., extrema of the action S. An inconsistency thus appears: The two descriptions of particles predict different trajectories. The only way to bring about harmony is to require that $F(I, H, \psi) = F(I, \infty, \psi) = A(I, \psi)$, i.e., that the Finsler geometry reduce to the Riemann geometry defined in Eq. (13). Of course, this is just a special case of the physical geometry obtained in Eq. (11).

The above remarks are not directly applicable to the subcase when the Finsler function vanishes asymptotically as $H \to +\infty$. For then the line element ds^2 remains non-Riemannian on the graviton null surface; i.e., Eq. (13) is not applicable. We are left without a metric with which to build the field equations, so that the matter physics would remain ill defined. We conclude that, physically speaking, this behavior of F must be excluded.

B. Finsler function F with one zero

Case A does not exhaust the possibilities. The function F may have zeros in H (for finite H). Let us consider the case, case B, in which F has one zero, $H = h(I, \psi)$, being positive for $H > h$ and negative for $H < h$. We see that this zero corresponds to the physical light cone. That is, $ds^2 = 0$ when

$$\frac{L^2(\psi_{,\alpha}\, dx^\alpha)^2}{-g_{\mu\nu}\, dx^\mu dx^\nu} = h(I, \psi) \tag{14a}$$

or

$$(h g_{\alpha\beta} + L^2 \psi_{,\alpha}\, \psi_{,\beta})\, dx^\alpha dx^\beta = 0. \tag{14b}$$

In obtaining the second equation use has been made of the assumption that $h \neq \infty$, i.e., $g_{\mu\nu}\, dx^\mu dx^\nu \neq 0$ at the zero of F. Thus, the coordinate increments dx^α which are null with respect to the Riemann metric $h g_{\alpha\beta} + L^2 \psi_{,\alpha}\, \psi_{,\beta}$ have $ds^2 = 0$ with respect to the Finsler geometry. They make up the light cone on which photons propagate.

Trajectories whose tangent coordinate increments have $H < 0$ with $H < h$ or $H > h$ with $H > 0$ are physical trajectories of "massive" particles (because they have $ds^2 < 0$). Thus when $h \leq 0$ (the physically interesting case as we shall see in Sec. IV), the whole positive H axis and that part of the negative axis left of h represent trajectories of massive particles. It may be seen that the light cone, at $H = h$, is the boundary of these trajectories (recall that $H = +\infty$ and $H = -\infty$ are to be identified), in accordance with intuition.

If F is bounded as $H \to \pm\infty$, another light cone appears: When $g_{\alpha\beta}\, dx^\alpha dx^\beta = 0$, ds^2 vanishes also. The simultaneous existence of two light cones at one event, with the second one having physically acceptable trajectories on either side of it, is unphysical. We thus require that $F \sim H$ as $H \to \pm\infty$ so that the vanishing of $g_{\alpha\beta}\, dx^\alpha dx^\beta$ is compensated for. In this way $ds^2 \neq 0$ on the null graviton surface.

In order to pick out a special metric from the Finsler geometry, let us expand $F(I, H, \psi)$ in powers of $(H - h)$ about the light cone $H = h$. Retaining only the first term we have

$$ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta\, F'(I, H = h, \psi)\,(H - h). \tag{15}$$

If we now multiply in $g_{\alpha\beta}\, dx^\alpha dx^\beta$ and define

$$B(I, \psi) \equiv -F'(I, h, \psi); \quad A(I, \psi) \equiv h(I, \psi) B, \tag{16}$$

we find that the geometry in the vicinity of the light cone is Riemannian with the metric $\tilde{g}_{\alpha\beta}$ of Eq. (11). Since we are considering the case where $F > 0$ for $H > h$, we must require that $B < 0$ in this case.

How would things change had we assumed that F is negative for $H > h$? In that case the physically allowed trajectories are restricted to the range $[0, h]$ of H. Graviton trajectories (at $H = \pm\infty$) are not contiguous to this physical range so that gravitons travel on trajectories with $ds^2 > 0$. We may thus exclude this case by causality and require

$$B(I, \psi) < 0. \tag{17}$$

The Finsler geometry has thus picked out a special Riemannian metric $\tilde{g}_{\alpha\beta}$, whose null surfaces coincide with the physical light cones. By analogy with the discussion in the last subsection, we must construct all field equations, such as Maxwell's, with $\tilde{g}_{\alpha\beta}$ to ensure that short waves travel on the physical light cone, $\tilde{g}_{\alpha\beta}\, dx^\alpha dx^\beta = 0$. This will then be in harmony with the classical "photon" trajectories derived from the action S. As before, massive particles will be predicted by the field equations to travel on geodesics of $\tilde{g}_{\alpha\beta}$, and by the action S to follow entirely different trajectories—extremal curves in the Finsler geometry. Harmony can be secured only by requiring that the Finsler geometry reduce, for any dx^α, to a Riemann geometry with metric $\tilde{g}_{\alpha\beta}$. Then the expansion Eq. (15) is exact. Note that in this case the requirement that F blow up linearly with H as $H \to \pm\infty$ is met automatically.

C. Other cases

The other cases of the Finsler function F may be characterized by the number and order of the zeros it possesses in the variable H. If there are more than one zero, we return to the problem of multiple light cones, and must thus exclude this case. Even when there is only one zero, at $H = h(I, \psi)$, we must face the possibility that it is a zero of higher order, i.e., that one or several derivatives of F vanish at $H = h$.

Suppose F has a zero of order $n > 1$. In its vicinity we may expand F in powers of $H - h$ and retain the first term:

$$ds^2 = -\frac{\{[h(I, \psi) g_{\alpha\beta} + L^2 \psi_{,\alpha}\, \psi_{,\beta}]\, dx^{\alpha\beta} dx^\beta\}^n}{D(I, \psi)(g_{\alpha\beta} dx^\alpha dx^\beta)^{n-1}}. \tag{18}$$

The manifold defined by the vanishing of the curly brackets in this expression is the light cone. We see that in its vicinity the line element is not even approximately Riemannian. Of course we may still use the metric $h g_{\alpha\beta} + L^2 \psi_{,\alpha}\, \psi_{,\beta}$ to construct Maxwell's equations, etc., and this will give photons which travel on its null surface. However, the usual problem will arise for classical massive particles: The field equations suggest they follow geodesics of $h g_{\alpha\beta} + L^2 \psi_{,\alpha}\, \psi_{,\beta}$, but the action S predicts they follow geodesics of the Finsler geometry. We cannot bring about harmony here by having the Finsler geometry reduce to Riemann geometry because the higher order zero means that the Finsler geometry is never close to a Riemann one, at least not near the light cone. We must, therefore, exclude the case with a higher order zero.

So far we have concentrated on special Riemann metrics picked out by the Finsler geometry on account of its behavior near the light cone. Of course, there may be other Riemann metrics which might be of relevance to our quest and which are not so characterized. However, because they would have separate origin, it seems implausible that such metrics would serve to construct matter field equations whose behavior at the light cone is compatible with the required photon trajectories. In view of this it seems we are restricted to the Riemann metrics given by Eq. (11).

IV. DISFORMAL TRANSFORMATION AND GRAVITATIONAL THEORY

We may subsume both cases A and B by adopting the form Eq. (11) for the physical metric together with the stipulation that $B(I, \psi)$ cannot be positive. Then $\tilde{g}_{\alpha\beta}$ in Eq. (11) with $B = 0$ coincides with the physical metric for case A. Thus on the basis of weak equivalence, causality, and consistency between the predictions of matter field equations and classical action principle for the trajectories of classical particles, we have made it plausible that Eq. (11) is the most general relation between gravitational metric $g_{\alpha\beta}$ and physical metric $\tilde{g}_{\alpha\beta}$ when the only additional structure admitted is a scalar field.

One more refinement is in order. The metric $\tilde{g}_{\alpha\beta}$ depends explicitly on ψ as well as on its derivatives. In general this would mean that one cannot change the zero of ψ without changing the metric in a non-negligible way. One can conceive that it would be useful to retain the property of translation in ψ which exists in other contexts in physics. We may secure this by requiring that both A and B in the metric depend on ψ only through a common factor of the form $\exp(2\psi)$. Any redefinition of the zero of ψ then amounts to a multiplication of the physical metric by a constant, i.e., to a global change of units which is physically irrelevant. Implementing this factoring leads us to our final expression for the physical metric:

$$\tilde{g}_{\alpha\beta} \equiv e^{2\psi}[A(I)g_{\alpha\beta} + L^2 B(I)\psi_{,\alpha}\,\psi_{,\beta}], \tag{19}$$

with

$$B(I) \leq 0. \tag{20}$$

We recognize Eq. (20) as the condition imposed by causality. It has earlier been derived in a different way in Ref. [12].

We call the sort of relation between $g_{\alpha\beta}$ and $\tilde{g}_{\alpha\beta}$ in Eq. (19) a *disformal* transformation. The term is meant as a contrasting one to *conformal* transformation which is the special case with $B = 0$ of Eq. (19). When $B = 0$ the transformation of a region of spacetime implied by $g_{\alpha\beta} \to \tilde{g}_{\alpha\beta}$ leaves all shapes invariant and merely stretches all spacetime directions equally. When $B \neq 0$ the stretch in the direction parallel to $\psi_{,\alpha}$ is by a different factor from that in the other spacetime directions and shapes are distorted. Maxwell's equations, the Weyl equation for spinors, gauge field equations, etc., are all invariant under the transformation with $B = 0$, but not generally under $g_{\alpha\beta} \to \tilde{g}_{\alpha\beta}$ with $B \neq 0$. Of course, the physical context in which we have introduced the disformal transformation requires that all the mentioned equations be written at the outset with the metric $\tilde{g}_{\alpha\beta}$.

The disformal transformation was introduced in Ref. [12] where restrictions on the various functions appearing in it were summarized. One is condition (20). We reiterate some of the others. For this purpose let us define

$$C(I) \equiv A(I) + I\,B(I). \tag{21}$$

The ratio C/A quantifies the anisotropy of the disformal transformation: The direction along $\psi_{,\alpha}$ is stretched by a factor C/A as large as that for the other three directions. Because we think of $g_{\alpha\beta}$ as a gravitational metric, it is clear its signature must be $\{-,+,+,+\}$ globally (or $\{+,-,-,-\}$ in the competing convention). Otherwise, the lack of global hyperbolicity will make the setting up of the initial value problem for the metric $g_{\alpha\beta}$ impossible. Furthermore, the physical metric $\tilde{g}_{\alpha\beta}$ must also have signature $\{-,+,+,+\}$ globally in order that it may be able to reduce to a Minkowski metric at every spacetime event. By considering these conditions in a local frame with one axis aligned with $\psi_{,\alpha}$, it is possible to show that [12]

$$A(I) > 0, \qquad C(I) > 0. \tag{22}$$

Comparing Eqs. (16), (20), and (22), we see that necessarily h depends only on I and $h(I) < 0$.

To be a *bona fide* metric, the physical metric $\tilde{g}_{\alpha\beta}$ must be invertible: There must be an inverse metric $\tilde{g}_{\alpha\beta}$ at every spacetime point. If it exists it must be of the form

$$\tilde{g}^{\alpha\beta} = e^{-2\psi}A(I)^{-1}[g^{\alpha\beta} - L^2 B(I)C(I)^{-1}g^{\alpha\mu}g^{\beta\nu}\psi_{,\mu}\,\psi_{,\nu}]. \tag{23}$$

The conditions (22) guarantee that this inverse is well defined everywhere.

One more condition can be obtained if it is agreed that there should be a one-to-one correspondence between gravitational metric and physical metric. Suppose one contracts Eq. (23) with $\psi_{,\beta}$. The result is

$$g^{\alpha\beta}\psi_{,\beta} = C(I)\,e^{2\psi}\tilde{g}^{\alpha\beta}\psi_{,\beta}\,. \tag{24}$$

A further contraction with $\psi_{,\alpha}$ gives

$$J \equiv L^2\,\tilde{g}^{\alpha\beta}\psi_{,\alpha}\,\psi_{,\beta} = e^{-2\psi}IC(I)^{-1}, \tag{25}$$

where it is clear that J is defined as the analogue of I written with metric $\tilde{g}^{\alpha\beta}$. In principle it is possible to solve Eq. (25) for I. If we solve Eq. (19) for $g_{\alpha\beta}$ and eliminate I everywhere, we get

$$g_{\alpha\beta} = A[I(J\,e^{2\psi})]^{-1}\{e^{-2\psi}\tilde{g}_{\alpha\beta} - L^2 B[I(J\,e^{2\psi})]\psi_{,\alpha}\,\psi_{,\beta}\}. \tag{26}$$

It can be seen that if $I(J\,e^{2\psi})$ were to be multiply valued, there would be several gravitational metrics for each physical metric, which would be unphysical. The way to avoid this is to require that J be a monotonic function of I for fixed ψ. We shall thus require

$$\frac{d}{dI}\left[\frac{I}{C(I)}\right] > 0. \tag{27}$$

We have required I/C to be increasing because the opposite assumption runs counter to the situation in many known theories (see below).

V. CONCLUSIONS AND QUESTIONS

We thus conclude that subject to conditions (20), (22), and (27), Eq. (19) is the most general relation between physical and gravitational metrics which respects the weak equivalence principle, ordinary notions of causality, and which is insensitive to a change of zero for the sole auxiliary scalar field ψ.

It should be noticed that Finsler geometry served mostly a negative role in our argument. Although it is more general than Riemann geometry, we found it rather unpromising for building dynamical equations for matter. It did point us to certain Riemannian geometries as candidates for the physical geometry. Had one proceeded entirely within the framework of Riemann geometry, one might not have thought of a relation such as Eq. (19) between gravitational and physical geometries. Thus the line of thought developed here opens up for discussion a broader class of physical geometries than those conformal to the gravitational geometry.

A conformal transformation between metrics can be interpreted as a change of local units of length [3,5]: The ratio between gravitational and material units varies from event to event. By analogy we may interpret the disformal transformation as a change of local units of length for which the units for intervals along the gradient of ψ are different than those for intervals orthogonal to it. Conformal transformations have traditionally been the source of insight into field theory. It appears likely that disformal transformations will help to supplement those insights. At the most immediate level, they provide a method for constructing novel gravitational theories based on pairs of disformally related geometries. As far as we are aware, such theories have been considered only once before [12], and the motivation there was to relate the standard interpretation of the data from gravitational lenses to the modified gravity resolution of the missing mass puzzle [13]. The present work thus provides a theoretical backbone for those studies.

Thus far we have only described a framework; concrete theories will arise when dynamics are specified for ψ. GR is the trivial case; it corresponds to the requirement that $\psi = \text{const}$. By an appropriate choice of units it can be arranged that $\exp(2\psi)\,A(0) = 1$, so that, as expected, GR equates physical and gravitational geometry. The value of B is plainly irrelevant and may be set to zero by fiat. Brans-Dicke theory (in Dicke's form) [3], Dirac's theory [5], and the variable mass theory [4] all prescribe nontrivial dynamics for ψ, but choose (in the appropriate units) $A(I) = 1$ and $B(I) = 0$. One can further conceive of theories with $A(I) \neq \text{const}$ and $B(I) < 0$ in which gravitons travel slower than photons, a feature which might be subject to direct experimental test. One such theory has been studied in detail in Ref. [12].

If the true gravitational theory is of the conventional type ($A = 1$ and $B = 0$), the question arises what symmetry or selection principle forces A and B into these trivial values when they could be functions of the invariant I? Conversely, if nature has taken advantage of the wider possibilities for A and B, in what ways are the intuitions about gravity that have been molded by conventional theories to be modified? Further work will take up these and other questions.

ACKNOWLEDGMENTS

I would like to thank Jim Hartle and Gary Horowitz for hospitality at Santa Barbara, and the Israel Science Foundation for support.

[1] A. Schild, in *Relativity Theory and Astrophysics*, edited by J. Ehlers (American Mathematical Society, Providence, Rhode Island, 1967).
[2] G. Nordström, Ann. Phys. (Leipzig) **42**, 533 (1913).
[3] C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961); R. H. Dicke, *ibid.* **125**, 2163 (1962).
[4] J. Bekenstein, Phys. Rev. D **15**, 1458 (1977); J. D. Bekenstein and A. Meisels, *ibid.* **18**, 1313 (1980).
[5] P. A. M. Dirac, Proc. R. Soc. London **A333**, 403 (1973).
[6] C. G. Callan, R. C. Myers, and M. J. Perry, Nucl. Phys. **B311**, 673 (1988).
[7] E. Cartan, *Les Espaces de Finsler* (Herman, Paris, 1934).
[8] I. W. Roxburgh, Gen. Relativ. Gravit. **23**, 1071 (1991).
[9] C. F. Will, Int. J. Mod. Phys. D **1**, 13 (1992).
[10] I. W. Roxburgh, Gen. Relativ. Gravit. **24**, 4191 (1992).
[11] R. K. Tavakol and N. van den Bergh, Gen. Relativ. Gravit. **18**, 849 (1986).
[12] J. D. Bekenstein, in *Sixth Marcel Grossmann Meeting on General Relativity*, Kyoto, Japan, 1992, edited by H. Sato and T. Nakamura (World Scientific, Singapore, 1992), p. 905.
[13] M. Milgrom, Astrophys. J. **270**, 365 (1983); J. D. Bekenstein and M. Milgrom, *ibid.* **286**, 7 (1984); R. H. Sanders, Astron. Astrophys. Rev. **2**, 1 (1990).