Barcan Marcus et al., eds., Logic, Methodology and Philosophy of Science VII

GRAVITY AND SPATIAL GEOMETRY[1]

DAVID MALAMENT*

Dept. of Philosophy, Univ. of Chicago, Chicago, IL 60637, USA

Philosophers of science have written at great length about the geometric structure of physical space. But they have devoted their attention primarily to the question of the epistemic status of our attributions of geometric structure. They have debated whether our attributions are *a priori* truths, empirical discoveries, or, in a special sense, matters of stipulation or convention. It is the goal of this paper to explore a quite different issue — the role played by assumptions of spatial geometry *within physical theory*, specifically within Newtonian gravitational theory.

Standard formulations of Newtonian physics, of course, presuppose that space is Euclidean. But the question arises whether they must do so. After all, the geometric structure of physical space was a topic of intense interest in the 19th century long before Newtonian physics was abandoned. Think of Gauss, Riemann, Helmholtz, and Poincaré. It is probably most natural to assume, and perhaps these men *did* assume, that any hypotheses about spatial geometry function only as inessential auxiliary hypotheses within Newtonian physics — superimposed, as it were, on a core of basic underlying physical principles which themselves are neutral with respect to spatial geometry. Yet it turns out that there is an interesting sense in which this is just not so, a sense which is only revealed when one considers Newtonian gravitational theory from the vantage point of general relativity.

One can, and I think should, construe the former theory as a special limiting form of the latter in which relativistic effects become negligible.

---

[1] The following is extracted from a long, technical paper [3]. Proofs can be found there together with a good deal of supplemental material on spacetime structure in Newtonian physics. The results presented there draw on work of Künzle in [1] and [2].

* I am grateful to Jürgen Ehlers and Robert Geroch for comments on an earlier draft. Ehlers, in particular, saved me from making a number of seriously misleading statements.

That is, one can think of Newtonian gravitational theory as the so-called "classical limit" of general relativity. The big surprise, at least to me, however, is that when one *does* think about it this way one finds that the theory *must* posit that space is Euclidean. It's curious. The very limiting process which produces Newtonian physics and a well-defined, observer invariant spatial structure also generates strong constraints on spatial curvature. These constraints turn out to be *so* strong as to guarantee the Euclidean character of space. That, anyway, will be my principal claim today.

Claim. Insofar as it is the "classical limit" of general relitivity, Newtonian gravitational theory *must* posit that space is Euclidean.

A good bit of differential geometry will be required to make the claim precise. But the underlying idea is quite intuitive. It is absolutely fundamental to relativity theory that there is an upper bound to the speeds with which particles can travel (as measured by an observer). The existence of this upper bound is embodied in the null cones (or light cones) one finds in spacetime diagrams. In classical physics, however, there is no upper bound to particle speeds. The transition from general relativity to Newtonian physics is marked by this all important difference. The maximal particle speed goes to infinity. The transition can be conceived geometrically as a process in which the null cones at all spacetime points "flatten" and eventually become degenerate. In the limit the cones are all tangent to a family of hypersurfaces, each of which represents "space" at a given "time". The curious fact is this. If at every intermediate stage of the collapse process spacetime structure is in conformity with the dynamic constraints of general relativity (as embodied in Einstein's field equation), then the resulting induced hypersurfaces are necessarily flat, i.e. have vanishing Riemann curvature. One can think of it this way — the limiting process which effects the transition from general relativity to Newtonian gravitational theory "squeezes out" all spatial curvature!

The proposition which follows is intended to capture the collapsing light cone picture in a precise statement about relativistic spacetime models.

We take a relativistic spacetime model to be a triple (M, g_{ab}, T_{ab}) where M is a smooth, connected, four-dimensional manifold (representing the totality of all spacetime points); g_{ab} is a smooth Riemannian metric of Lorentz signature $(+1, -1, -1, -1)$ on M (which represents the metric of spacetime); T_{ab} is a smooth, symmetric field on M (which represents the mass-energy density present throughout spacetime); and where Einstein's equation

$$R_{ab} - \tfrac{1}{2}g_{ab}R = 8\pi T_{ab}$$

is satisfied. In the proposition we start with a one-parameter family of such models all sharing the same underlying manifold M:

$$(M, g_{ab}(\lambda), T_{ab}(\lambda)), \quad 0 < \lambda \leq 1.$$

Then we impose two constraints — one on the limiting behavior of the $g_{ab}(\lambda)$ as λ goes to 0, and one on that of the $T_{ab}(\lambda)$. The first guarantees that all null cones open up and become tangent to a family of hypersurfaces. The second guarantees that the limiting values of mass-energy density, momentum density, and material stress (as determined by any one observer) are all finite. Our conclusion is that as a result of the conditions imposed the limiting hypersurfaces have vanishing Riemann curvature.

To motivate the first constraint it will help to consider a special case which should look familiar. In Minkowski spacetime all curvature vanishes. One can find a global t, x, y, z coordinate system in which the metric g_{ab} and its inverse g^{ab} have coefficients

$$g_{ab} = \mathrm{diag}(+1, -1/c^2, -1/c^2, -1/c^2),$$
$$g^{ab} = \mathrm{diag}(+1, -c^2, -c^2, -c^2).$$

(That is, the coefficients of g_{ab} form a 4×4 matrix whose diagonal entries are $+1$, $-1/c^2$, $-1/c^2$, $-1/c^2$, and whose non-diagonal entries are all 0.) Now let us consider these as fields parametrized by c. The first has a limit as c goes to infinity. The other does too after it is suitably rescaled:

$$g_{ab}(c) \rightarrow \mathrm{diag}(+1, 0, 0, 0),$$
$$g^{ab}(c)/c^2 \rightarrow \mathrm{diag}(0, -1, -1, -1).$$

In a sense the limiting process has allowed us to recover separate temporal and spatial metrics. We have pulled apart a non-degenerate metric of signature $(+1, -1, -1, -1)$ to recover its degenerate positive and negative pieces.

This example is special in several respects. The null cones open symmetrically around the "time" axis at each point. The opening occurs uniformly across the manifold. (It is as if the cones were rigidly rigged to each other.) And background affine structure is kept fixed and flat throughout the process. These features cannot be retained when one considers arbitrary (curved) relativistic spacetime models. But the limit existence assertions ***can*** be generalized, and they turn out to be exactly what one needs.

Consider again our parametrized family of metrics. We are not going to regiment how their null cones open. We shall allow, intuitively, that the cones open at different rates at different points, that their axes wiggle as they open, and so forth. Our sole requirement is that, *somehow or other*, the cones do finally become tangent to a family of "constant-time" hypersurfaces, ***and*** that they do so in such a way that, after rescaling, a well-defined spatial metric is induced on the surfaces. Formally the requirement comes out this way. (Here and in what follows, all limits are taken as λ goes to 0.)

(1a) There exists a smooth, non-vanishing, closed field t_a on M such that $g_{ab}(\lambda) \rightarrow t_a t_b$.

(1b) There exists a smooth, non-vanishing field h^{ab} of signature $(0, +1, +1, +1)$ on M such that $\lambda g^{ab}(\lambda) \rightarrow -h^{ab}$.

Clearly the parameter λ corresponds to $1/c^2$.

Let's consider the first clause. I claim that it captures the intended collapsing null cone condition. Suppose t_a is as in (1a). Since it is closed, t_a must be locally exact. That is, at least locally it must be the gradient of some scalar field t on M. It is precisely the hypersurfaces of constant t value to which the cones of the $g_{ab}(\lambda)$ become tangent. [To see this let ∇_a be any derivative operator on M, and let η^a be any vector in the domain of t, tangent to the surface through that point. Then $t_a = \nabla_a t$ and $\eta^a \nabla_a t = 0$. It follows that

$$g_{ab}(\lambda)\eta^a\eta^b \rightarrow t_a t_b \eta^a \eta^b = (\eta^a \nabla_a t)^2 = 0.$$

Thus, in the limit η^a becomes a null vector. The surfaces of constant t value are degenerate null cones!]

One can also easily verify that the scalar field t gives limiting values of elapsed proper time. [Suppose that $\gamma : [a, b] \rightarrow M$ is a timelike curve with respect to all the $g_{ab}(\lambda)$, and its image falls within the domain of t. The elapsed proper time between $\gamma(a)$ and $\gamma(b)$ along γ relative to $g_{ab}(\lambda)$ is given by

$$PT(\gamma, \lambda) = \int_a^b [g_{mn}(\lambda)\eta^m\eta^n]^{1/2} \mathrm{d}s$$

where η^a is the tangent field to γ. As λ goes to 0 we have

$$PT(\gamma, \lambda) \rightarrow \int_a^b (t_n \eta^n)\mathrm{d}s = \int_a^b (\eta^n \nabla_n t)\mathrm{d}s = t(\gamma(b)) - t(\gamma(a)).$$

Thus the limiting value of proper time is independent of the choice of timelike curve connecting $\gamma(a)$ to $\gamma(b)$. It is given, simply, by the t-coordinate interval between the two points.]

It remains now to consider the constraint to be imposed on the mass-energy tensor fields $T_{ab}(\lambda)$. Suppose (M, g_{ab}, T_{ab}) is a relativistic spacetime model, and ξ^a is a unit timelike vector at some point of M representing an observer 0. 0 will decompose T_{ab} at the point into its temporal and spatial parts by contracting each index with $\xi^a\xi^m$ or $(\xi^a\xi^m - g^{am})$. (The latter is the "spatial metric" as determined by 0.) The components he determines have the following physical interpretation:[2]

$$T_{ab}\xi^a\xi^b = \text{mass-energy density relative to } 0,$$

$$T_{ab}\xi^a(\xi^b\xi^n - g^{bn}) = \text{three-momentum density relative to } 0,$$

$$T_{ab}(\xi^a\xi^m - g^{am})(\xi^b\xi^n - g^{bn}) = \text{three-dimensional stress tensor relative to } 0.$$

We shall require of the limiting process that it assign (finite) limiting values to these quantities as determined by some observer 0. The condition comes out as follows.

(2) There exists a smooth field T^{ab} on M such that $T^{ab}(\lambda) \to T^{ab}$.

Here $T^{ab}(\lambda) = T_{mn}(\lambda)g^{ma}(\lambda)g^{nb}(\lambda)$. [The condition is stronger than the requirement that the $T_{ab}(\lambda)$ have a finite limit. To see where it comes from, consider a family of coaligned vectors $\xi^a(\lambda)$, each of unit length with respect to $g_{ab}(\lambda)$. For each λ, perform the decomposition above. If $T_{ab}(\lambda)\xi^a(\lambda)\xi^b(\lambda)$, $T_{ab}(\lambda)\xi^a(\lambda)[\xi^b(\lambda)\xi^n(\lambda) - g^{bn}(\lambda)]$, and $T_{ab}(\lambda)[\xi^a(\lambda)\xi^m(\lambda) - g^{am}(\lambda)][\xi^b(\lambda)\xi^n(\lambda) - g^{bn}(\lambda)]$ are all to have finite limits, it follows that $T_{ab}(\lambda)g^{am}(\lambda)g^{bn}(\lambda)$ must have one too.] Now we can formulate the proposition.

PROPOSITION. *Suppose that for all $\lambda \in (0,1]$, $(M, g_{ab}(\lambda), T_{ab}(\lambda))$ is a relativistic spacetime model. Further suppose that conditions* (1) *and* (2) *above are satisfied. Finally suppose that S is any spacelike hypersurface in M as determined by t_a (i.e. any imbedded three-dimensional submanifold of M satisfying $t_a\eta^a = 0$ for all vectors η^a tangent to S). Then if $\mathcal{R}^a{}_{bcd}(\lambda)$ is the three-dimensional Riemann curvature tensor field on S induced by $g_{ab}(\lambda)$, $\mathcal{R}^a{}_{bcd}(\lambda) \to 0$.*

[2] See, e.g., MISNER, THORNE, and WHEELER [4], p. 131

A proof is given in considerable detail in [3]. Here we simply indicate the structure of the argument. It proceeds in two stages. Suppose that for each λ, $\nabla_a(\lambda)$ is the unique derivative operator (or affine connection) on M compatible with $g_{ab}(\lambda)$. Further suppose that ρ is taken to be the scalar field $T^{ab}t_a t_b$. First one shows that there must exist a derivative operator ∇_a on M such that $\nabla_a(\lambda) \rightarrow \nabla_a$,[3] and such that the structure $(M, t_a, h^{ab}, \nabla_a, \rho)$ satisfies the conditions:

Compatibility	$\nabla_a t_b = 0 = \nabla_a h^{bc}$,
Orthogonality	$t_a h^{ab} = 0$,
Poisson's Equation	$R_{ab} = 4\pi\rho t_a t_b$,
Integrability	$R^{[a}{}_{(b}{}^{c]}{}_{d)} = 0$.

These conditions characterize a kind of generalized Newtonian spacetime structure introduced by Künzle in [1] and [2]. Thus the first stage of the argument is of interest in its own right. It makes precise one sense in which a generalized version of Newtonian gravitational theory is the "classical limit" of general relativity.[4] In particular it shows that Poisson's equation is a limiting form of Einstein's equation.

The second stage of the argument makes the connection with spatial geometry. It certainly need not be the case that the four-dimensional Riemann tensor field $R^a{}_{bcd}$ on M determined by ∇_a vanishes. But Poisson's equation (in the presence of the Compatibility and Orthogonality conditions) *does* imply that the three-dimensional Riemann field $\mathcal{R}^a{}_{bcd}$ induced on any spacelike hypersurface S does so. (The claim is that *space*, not *spacetime*, is necessarily flat in the "classical limit" of general relativity.) Once the dust clears, this second stage of the argument turns on a simple linear algebraic fact. In three dimensions (but not higher) the Ricci tensor field cannot vanish without the full Riemann tensor field doing so as well.

One has $\mathcal{R}^a{}_{bcd} = 0$; and $\mathcal{R}^a{}_{bcd}(\lambda) \rightarrow \mathcal{R}^a{}_{bcd}$ follows easily from $\nabla_a(\lambda) \rightarrow \nabla_a$. So the proposition follows.

Edmund Whittaker once said that "gravitation simply represents a continual effort of the universe to straighten itself out". I have tried to show that at least in the limiting Newtonian context that straightening process is so complete as to rule out any spatial curvature whatsoever.

[3] The condition $\nabla_a(\lambda) \rightarrow \nabla_a$ can be taken to mean that for any smooth vector field η^a on M, $\nabla_a(\lambda)\eta^b \rightarrow \nabla_a\eta^b$. See [3] for a detailed discussion of limit relations between tensor fields.

[4] The argument in [3] is a variant of that given by Künzle in [2].

References

[1] KÜNZLE, H., 1972, *Galilei and Lorentz structures on space-time: Comparison of the corresponding geometry and physics*, Ann. Inst. Henri Poincaré 17, p. 337.
[2] KÜNZLE, H., 1976, *Covariant Newtonian limit of Lorentz space-times*, General Relativity and Gravitation 7, p. 445.
[3] MALAMENT, D., *Newtonian gravity, limits, and the geometry of space*, forthcoming in: Pittsburgh Studies in the Philosophy of Science.
[4] MISNER, C., THORNE, K. and WHEELER, J., 1973, *Gravitation* (W. H. Freeman, San Francisco, CA).