THE FOUNDATIONS OF THE THEORY OF GRAVITATION*

Sir Hermann Bondi

Churchill College
Cambridge CB3 0DS, U.K.

(Received June 29, 1999)

It is emphasized that Einstein's theory of gravitation has its physical and logical roots, firstly, in Newton's theory, namely in the existence of Newtonian tidal forces, and, secondly, in the requirement that these forces be compatible with the theory of relativity. Furthermore, it is pointed out that the nonexistence of any covariant description of energy in Einstein's theory is deeply rooted in a local 'unindentifiability' of the gravitational force in Newton's theory, although this fact is irrelevant in that theory.

PACS numbers: 04.20.Cv

Many physicists think of Einstein's theory of gravitation as a very separate part of their subject based somewhat obscurely on an alleged 'principle of general relativity' and a sophisticated 'principle of equivalence'. It is my purpose to show that Einstein's theory can be directly inferred by normal physical arguments from Galileo's discovery, confirmed to high accuracy by modern measurements, that all bodies fall equally, irrespective of shape, composition, etc. What then is the *universal* observable of this phenomenon? It is evidently not weight which only arises when we stand an a solid body like the Earth. But what is always present, even for freely falling bodies, is the tide-raising force so brilliantly explained by Newton as due to the non-uniformity of gravitation. Thus the universal observable of gravitation is *the relative acceleration of neighbouring particles.* (Note that with this definition an alleged 'uniform gravitational field' is no gravitational field at all and arguments about falling lifts etc. are pointless.)

To put this mathematically, the observable of gravitation is thus the link between the relative acceleration vector δf^i of two neighbouring particles

* Presented at "The Infeld Centennial Meeting", Warsaw, Poland, June 22–23, 1998.

and their relative displacement vector δx^i. This is described by the relation

$$\delta f^i = a^i{}_j \, \delta x^j. \tag{1}$$

Therefore the tensor a_{ij} is the observable of gravitation. Note that any anti-symmetry in a would imply that gravitation could spin up bodies without limit. As this is not observed, a_{ij} must be a symmetrical tensor. In fact Newtonian theory asserts that

$$a_{ij} = -\frac{\partial^2 V}{\partial x^i \partial x^j} \tag{2}$$

and is completed by Poisson's equation linking gravitation to its sources

$$a^i{}_i = -\nabla^2 V = -4\pi \, G \, \rho. \tag{3}$$

Excellent as Newton's theory has proved to be, it is unacceptable since it is non-relativistic. This is clear from (1), since if one particle moved at just below the speed of light, there will exist displacement vectors implying that the other particle is accelerated through the speed of light. Thus a must depend on the velocity v^k. Interpreting (1) as a four-dimensional equation, this dependence must ensure that the resulting acceleration is necessarily orthogonal to the velocity, as this excludes acceleration through the speed of light. Accordingly there needs to be an anti-symmetry in the connection. To combine this with the previously required symmetry needs a slightly involved algebraic consideration. It emerges that the only simple solution is to put

$$\delta f^i = c^i{}_{jkl} \, \delta x^j v^k v^l, \tag{4}$$

where the four suffix tensor c_{ijkl} is anti-symmetric for an interchange of the second and third suffix, but symmetric for the *double* interchange of the first and second *and* third and fourth suffixes. (Anti-symmetry for the interchange of the first and fourth suffixes follows.) Thus the tensor c constitutes the most straightforward way of making Newton's tidal relation (1) compatible with special relativity. A further physical consideration is now helpful to elucidate the significance of c.

There are many ways to derive the gravitational or Einstein red shift, which occurs when the emitter of radiation is *below* the receiver. (My favourite derivation uses a tower on the Earth carrying a closed chain of buckets, filled with atoms of some species. On one side they are all in the ground state, on the other side in a specified excited state, which makes them more massive and therefore heavier. The ensuing motion makes the need for the red shift patent.) Spectral lines being the basis of time keeping, the red shift implies that two clocks that keep the same time when side by side, keep time differently when one is stationed above the other. Apply this to a spherical Earth and attempt to describe the situation by a relativistic metric. Because of the red shift, the coefficient of the time term must depend on the radial distance. Without any infringement of generality one arrives therefore at

$$ds^2 = g(r)\,dt^2 - h(r)\,dr^2 - r^2\left(d\theta^2 + \sin^2\theta\,d\varphi^2\right), \tag{5}$$

where g, because of the red shift, is not constant. This cannot occur in a flat 4-space. Thus basic physical considerations prove that a relativistic theory of gravitation implies a non-Euclidean space. The simplest such geometry is Riemannian and it is therefore cogent to use it.

The most important quantity in Riemannian geometry, specifying its deviation from flatness, is the curvature tensor R_{ijkl}, which incorporates the symmetry relations found above (equation (4)) for the tensor c. The fundamental equation of geodesic deviation is effectively identical with (4). Thus the direct inference from purely physical considerations is that, in order to make Newton's theory relativistic, one has to adopt a Riemannian geometry in which the paths of free particles are geodesics and the curvature tensor is the observable of gravitation.

Note that in this derivation of Einstein's theory there is no mention of any 'general relativity'. Also, gravitation is adequately characterised by Galileo's statement that all bodies fall equally.

From this point on the further development of the theory is conventional. It may, however, be useful to make some comments on the physics of energy in the theory of gravitation. In Newtonian theory, gravitational potential energy plays an essential part in the conservation of energy, whenever there is an interaction between non-gravitational ('tangible') forces and the 'intangible' force of gravitation. (This terminology aims to indicate that locally gravitational force is not identifiable, since one cannot compare the motion of a particle affected by gravitation with one not so affected, whereas for example in an electric field one can readily distinguish the motion of a charged particle from that of a neutral one.) As a helpful Newtonian example, consider two bodies of similar masses describing eccentric elliptic orbits about their common centre of mass. In each orbit, the (tangible) kinetic energy waxes and wanes; the potential energy correspondingly waning and waxing. The kinetic energy of each body belongs to that body and resides in it and the kinetic energy of the system is their sum. However, the potential energy belongs to the system as a whole and cannot meaningfully be shared out between the two bodies or located in them. The fact that no position can be ascribed to gravitational potential energy is irrelevant in Newtonian theory. However, in a relativistic theory energy has mass and therefore an unlocalisable energy is unacceptable. Accordingly there is an excellent *physical* reason why there is no covariant description of gravitational energy in Einstein's theory.

But without such a measure of gravitational energy there is no conservation of energy. It is therefore unsound to call the field equations of the theory 'conservation laws'. Mathematically this is clear since, unlike an ordinary divergence, a covariant one does not lead to a Green's type of integral formulation. Physically it follows from what has been said above. The field equations should therefore be called laws of the *non-conservation* of the energy of the tangible forces arising from their interaction with gravitation.

Are there indeed any laws of the conservation of energy in the theory? There are none locally (except under very special conditions), but, given suitable boundary conditions at infinity, sophisticated considerations show that there are global conservation rules. This is most easily seen in the spherically symmetric case: Consider an isolated mass contracting slowly and heating up in the process. This will increase its thermal energy which is observable and therefore tangible. In the Newtonian framework we would say that this is balanced by a decrease in the (unobservable) gravitational potential energy. This is an irrelevance in relativistic theory, but we know from Birkhoff's theorem that the mass of the entire system, as seen from the outside, is constant. Thus there is conservation.

I have tried to make evident that some of the peculiar features of Einstein's theory of gravitation are not 'accidental' results of the mathematical apparatus, but are necessary physical consequences of constructing a relativistic theory based on Galileo's discovery.