

2

Mach's Principle

So strongly did Einstein believe at that time in the relativity of inertia that in 1918 he stated as being on an equal footing three principles on which a satisfactory theory of gravitation should rest:

1. The principle of relativity as expressed by general covariance.
2. The principle of equivalence.
3. Mach's principle (the first time this term entered the literature):...that the $g_{\mu\nu}$ are completely determined by the mass of bodies, more generally by $T_{\mu\nu}$.

In 1922, Einstein noted that others were satisfied to proceed without this [third] criterion and added, "This contentedness will appear incomprehensible to a later generation however."

...It must be said that, as far as I can see, to this day Mach's principle has not brought physics decisively farther. It must also be said that the origin of inertia is and remains *the* most obscure subject in the theory of particles and fields. Mach's principle may therefore have a future – but not without the quantum theory.

–Abraham Pais, *Subtle is the Lord: the Science and the Life of Albert Einstein*, pp. 287–288.
(Quoted by permission of Oxford University Press, Oxford, 1982)

BACKGROUND

Recapitulating, we have seen that when the implications of the principle of relativity for space and time were understood in the early twentieth century, Einstein quickly apprehended that the quantity of interest in the matter of inertia was not (rest) mass *per se*, rather it was the total non-gravitational energy contained in an object (isolated and at rest). This followed from Einstein's second law, which says:

$$m = \frac{E}{c^2}, \quad (2.1)$$

where m is now understood as the total inertial mass, not just the rest mass of an object, and E is the total non-gravitational energy. If one restricts oneself to Special Relativity Theory (SRT), this is about all one can say about inertial mass. It was Einstein's hope that he could go farther in identifying the origin of inertia in General Relativity Theory (GRT), as is evident in the quote from Pais's biography of Einstein above.

30 Mach's principle

As we have seen in the previous chapter, Einstein didn't need "Mach's principle" to create GRT. Shortly after publishing his first papers on GRT, he did try to incorporate the principle into his theory. He did this by adding the now famous "cosmological constant" term to his field equations. Those equations, as noted in Chapter 1, without the cosmological constant term, are:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu}, \quad (2.2)$$

where $G_{\mu\nu}$ is the Einstein tensor that embodies the geometry of spacetime, $R_{\mu\nu}$ is the "contracted" Ricci tensor (obtained by "contraction" from the Riemann curvature tensor which has four "indexes," each of which can take on values 1–4 for the four dimensions of spacetime), $g_{\mu\nu}$ is the "metric" of spacetime, and $T_{\mu\nu}$ is the "stress-energy-momentum" tensor, that is, the sources of the gravitational field. The cosmological term gets added to $G_{\mu\nu}$. That is, $G_{\mu\nu} \rightarrow G_{\mu\nu} + \lambda g_{\mu\nu}$, where λ is the so-called cosmological constant. We need not worry about the details of these tensor equations. But it's worth remarking here that the coefficient of $T_{\mu\nu}$, with factors of Newton's constant of gravitation G in the numerator and the speed of light c to the fourth power in the denominator, is exceedingly small. This means that the sources of the field must be enormous to produce even modest bending of spacetime. That is why a Jupiter mass of exotic matter is required to make wormholes and warp drives.

Ostensibly, Einstein added the cosmological constant term to make static cosmological solutions possible by including a long-range repulsive force. But he also hoped that the inclusion of the cosmological constant term would render his field equations solutionless in the absence of matter. Willem deSitter quickly showed that Einstein's new equations had an expanding, asymptotically empty solution, one with full inertial structure. And a vacuum solution, too. So Einstein's attempt to include Mach's principle in this way was deemed a failure.

The chief reason for his failure seems to have been the way he defined the principle: that the inertial properties of objects in spacetime should be defined (or caused) by the distribution of matter (and its motions) in the universe. Put a little differently, Einstein wanted the sources of the gravitational field at the global scale to determine the inertia of local objects. He called this "the relativity of inertia." The problem Einstein encountered was that his GRT is a local field theory (like all other field theories), and the field equations of GRT admit global solutions that simply do not satisfy any reasonable formulation of, as he called it, Mach's principle. Even the addition of the "cosmological constant" term to his field equations didn't suffice to suppress the non-Machian solutions.

Alexander Friedmann and Georges Lemaitre worked out cosmological solutions for Einstein's field equations in the 1920s, but cosmology didn't really take off until Edwin Hubble, very late in the decade, showed that almost all galaxies were receding from Earth. Moreover, they obeyed a velocity-distance relationship that suggested that the universe is expanding. From the 1930s onward work on cosmology has progressed more or less steadily. The cosmological models initiated by Friedman, predicated on the homogeneity and isotropy of matter at the cosmic scale, were developed quickly by Robertson and Walker. So now cosmological models with homogeneity and isotropy are called

Friedmann, Robertson, Walker (FRW) cosmologies. One of them is of particular interest: the model wherein space is flat at cosmological scale.

Spatial flatness corresponds to “critical” cosmic matter density $- 2 \times 10^{-29}$ g per cubic centimeter – and has the unfortunate tendency to be unstable. Small deviations from this density lead to rapid evolution away from flatness. Since flatness is the observed fact of our experience and the universe is more than 10 billion years old, how we could be in a spatially flat universe so long after the primeval fireball was considered something of a problem. The advent of “inflationary” cosmologies 20 or so years ago is widely thought to have solved this problem. As we will see shortly, spatial flatness and critical cosmic matter density figure into the answer to the question of the origin of inertia. But we are getting ahead of the story.

MACH'S PRINCIPLE

As the Equivalence Principle makes clear, gravity defines local inertial frames of reference as those in a state of free fall in the vicinity of a local concentration of matter. Moreover, gravity is the only truly “universal” interaction in that gravity acts on everything. For these reasons Einstein was convinced that GRT should also account for inertial phenomena, for inertia, like gravity, is a universal property of matter, though it is normally “inert.” (Good historical articles on his attempts to incorporate Mach's principle in GRT can be found in: *Mach's Principle: From Newton's Bucket to Quantum Gravity*, Brikhauser, Boston, 1995, edited by Julian Barbour and Herbert Pfister.)

Notwithstanding that Willem deSitter shot down his early efforts to build Mach's principle into GRT by adding the “cosmological constant” term to his field equations, Einstein persisted. When he gave a series of lectures on GRT at Princeton in 1921, he included extended remarks on the principle and the issue of inertia in GRT. (These remarks can be found in *The Meaning of Relativity*, 5th ed., Princeton University Press, Princeton, 1955, pp. 99–108). In his words:

[T]he theory of relativity makes it appear probable that Mach was on the right road in his thought that inertia depends upon a mutual action of matter. For we shall show in the following that, according to our equations, inert masses do act upon each other in the sense of the relativity of inertia, even if only very feebly. What is to be expected along the line of Mach's thought?

1. The inertia of a body must increase when ponderable masses are piled up in its neighborhood.
2. A body must experience an accelerating force when neighbouring masses are accelerated, and, in fact, the force must be in the same direction as that acceleration.
3. A rotating hollow body must generate inside of itself a “Coriolis field,” which deflects moving bodies in the sense of the rotation, and a radial centrifugal field as well.

We shall now show that these three effects, which are to be expected in accordance with Mach's ideas, are actually present according to our theory, although their magnitude is so small that confirmation of them by laboratory experiments is not to be thought of. . . .

The first of Einstein's criteria is the idea that when “spectator” matter is present in the vicinity of some massive object, the spectator matter should change the gravitational

potential energy of the object. And since $E = mc^2$, that gravitational potential energy should contribute to E and change the mass of the object.

It turns out that Einstein was wrong about this. Only non-gravitational energies contribute to E when it is measured locally. But the reason why E , locally measured, doesn't include gravity involves a subtlety about the nature of gravity and inertia that is easily missed. The second criterion is the prediction of, as it is now known, "linear accelerative frame dragging," though Einstein states it as the production of a force by the accelerating spectator matter on the body in question, rather than the dragging of local spacetime by the accelerating matter. This, when the action of the universe is considered, turns out to be the nub of Mach's principle. If the universe is accelerated in any direction, it rigidly drags inertial frames of reference along with it in the direction of the acceleration. Consequently, only accelerations relative to the universe are detectable; and inertia is "relative."

Einstein didn't consider the cosmological consequences of this term. But he showed that this term and its effects depends on gravity being at least a vector field theory (analogous to Maxwell's theory of electrodynamics). The effect is not to be found in Newtonian gravity, a scalar field theory (as the field equation can be written in terms of a scalar "potential" alone with the direction and magnitude of gravitational forces recovered using the "gradient operator"). The third criterion is just the Lens-Thirring effect and Gravity Probe B prediction.¹

Solving the full tensor field equations of GRT exactly is notoriously difficult, so Einstein did a calculation in the "weak field" approximation (where the metric tensor $g_{\mu\nu}$ is approximated by $\eta_{\mu\nu} + h_{\mu\nu}$ with $\eta_{\mu\nu}$ the Minkowski tensor of the flat spacetime of SRT and $h_{\mu\nu}$ the tensor that represents the field) and put his results into vector formalism. Suffice it to say, he found results that seemed to support each of his three criteria. (The formal predictions can be found in an excerpt from a paper by Carl Brans on the localization of gravitational energy at the end of this chapter.) His predicted effects are indeed very small when one considers even quite large local concentrations of matter (other than black holes in the vicinity of event horizons, of course).

Why didn't Einstein see that the sort of force that, because of the universality of gravity, is equivalent to frame dragging in his second prediction could explain Mach's principle? At least part of the problem here seems to be that he wasn't thinking cosmologically when looking for predicted quantitative effects – and so little was understood about the structure and size of the universe in the 1920s that there was no plausible basis, other than the most general sorts of considerations, to make inferences about the action of cosmic matter on local objects.

Shortly after Einstein gave his Princeton lectures, he found out, through posthumously reported remarks made by Mach shortly before his death in 1916, that Mach had disavowed any association with Einstein's ideas on relativity and inertia.

¹ Initially conceived of by George Pugh and Leonard Schiff in the 1960s, Gravity Probe B was a collection of high precision gyroscopes flown in a satellite in polar orbit intended to detect the dragging of spacetime caused by the rotation of Earth. The project, which flew several years ago, spanned decades and cost nearly a billion dollars. One noted relativist, queried by the press on the launch of the satellite, was reported to have remarked, "never was so much spent to learn so little." The history of this project is yet to be written. But it will doubtless prove fascinating.

Einstein, not long thereafter, asserted that any correct cosmological model should be spatially closed so that its geometry (the left hand side of his field equations) would be completely determined by its sources (the right hand side of his field equations) without the stipulations of additional boundary conditions and abandoned further work on Mach's principle.

If you are an expert, you may also be thinking, Einstein's calculation was done in the weak field approximation where gravitational effects are small. In cosmological circumstances one can expect gravitational potentials to be very large; indeed, even as large as the square of the speed of light – as is the case near the event horizon of a black hole. Well yes. But the universe isn't like the region of spacetime near to the event horizon of a stellar mass black hole. The sort of curvature encountered there is simply absent in the universe considered at cosmic scale. At cosmic scale, the universe is spatially flat. And absent local concentrations of matter, spacetime looks Minkowskian, notwithstanding that the gravitational potential approaches the square of the speed of light. So using the weak field approximation to compute lowest order gravimagnetic effects is perfectly okay.

THE MACH'S PRINCIPLE REVIVAL

By the early 1950s, the cosmological situation had changed. Significant theoretical work on cosmology had taken place, for example, that of Roberston and Walker in the 1930s and 1940s. Thomas Gold, Herman Bondi, and Fred Hoyle had proposed "steady state" cosmology, and Walter Baade had shown that there were two populations of stars, dramatically increasing the age of the universe for FRW cosmological models. So when Dennis Sciama, one of the very few doctoral students trained by Paul Dirac, came along in the early 1950s, tackling the "problem of the origin of inertia" seemed a reasonable thing to do.

Sciama's approach was to ignore GRT and write down a vector theory of gravity analogous to Maxwell's theory of electrodynamics. He initially thought his vector theory different from GRT. But eventually it was found to be just an approximation to GRT. This, by the way, is an exceedingly important point. Sciama's calculations are *not* optional. They are the exact predictions of GRT when conditions make the vector approximation valid and the idealizations he adopted reasonable.

What Sciama noticed was that when you write out the equation for the gravity field that is the analog of the electric field in electrodynamics, in addition to the commonplace term involving the gradient of a scalar potential, there is a term that is the rate of change of the "vector potential." In electrodynamics, the vector potential is associated with the magnetic field, and the term involving the rate of change of the vector potential that appears in the equation for the electric field means that when the magnetic field changes, it contributes to the electric field, causing it to change, too. Sciama noted that in the analogous case for gravity, the rate of change of the vector potential leads to a term in the "gravelectric" field that depends on acceleration of an object relative to the (on average) uniform bulk of the matter in the universe. That is,

$$\mathbf{E}_g = -\nabla\phi - \frac{1}{c} \frac{\partial \mathbf{A}_g}{\partial t} . \quad (2.3)$$

where \mathbf{E}_g is the gravelectric field strength, c the vacuum speed of light, and ϕ and \mathbf{A}_g the scalar and three-vector gravitational potentials respectively produced by all of the “matter” in the causally connected part of the universe. Matter is in quotes because what counts as matter is not universally agreed upon. We take “matter” to be everything that gravitates. This includes things such as zero-restmass energetic radiation and “dark energy,” which are sometimes excluded as matter. The “del” in front of the scalar potential is the “gradient” operator, which returns the rate of change of the potential in space and its direction. The relationship that allows one to write the change in \mathbf{A}_g terms of the scalar potential and velocity is the fact that \mathbf{A}_g is just the sum over all matter currents in the universe. That is,

$$\mathbf{A}_g = \frac{1}{c} \int_V \frac{\rho \mathbf{v}}{r} dV , \quad (2.4)$$

where ρ is the matter density in the volume element dV , \mathbf{v} the relative velocity of the object and volume element, and r the radial distance to the volume element. The factor of c in the denominator appears because Gaussian units are employed.² Sciama assumed that gravity, like electromagnetism, propagates at speed c , so normally this integration would involve a messy calculation involving retarded Green's functions and other mathematical complications. But because of the extremely simple, idealized conditions Sciama imposed, he saw that he could sidestep all of that messiness by invoking a little trick.

Sciama noted that in the case of an object moving with velocity \mathbf{v} with respect to the rest of the universe, one could change reference frame to the “instantaneous frame of rest” of the object; and in that frame the object is at rest and the rest of the universe moves past it – apparently rigidly – with velocity $-\mathbf{v}$. Since, in this special frame of reference everything in the universe, as detected by the object, is moving with the same velocity $-\mathbf{v}$ – the velocity in the integration of Eq. 2.4 can be removed from the integration, and Eq. 2.4 becomes:

$$\mathbf{A}_g = \frac{1}{c} \int_V \frac{\rho \mathbf{v}}{r} dV = \frac{\mathbf{v}}{c} \int_V \frac{\rho}{r} dV . \quad (2.5)$$

The result of this trick is to transform an integration over matter *current* densities into an integration over matter densities *per se*. Anyone familiar with elementary electrodynamics will instantly recognize this integration as that which gives the *scalar* potential of

² Nowadays in some quarters so-called SI units are used. They make the magnitudes of many things normally encountered in field theory unintuitively large or small. I use the traditional Gaussian units of field theory because there was a good reason why they were adopted decades ago by those who work in this area.

the field – but in this case, it returns the scalar potential of the gravitational field. As a result, for the simple case considered by Sciama, Eq. 2.5 becomes:

$$\mathbf{A}_g = \frac{1}{c} \int_V \frac{\rho \mathbf{v}}{r} dV = \frac{\mathbf{v}}{c} \int_V \frac{\rho}{r} dV \cong \frac{\mathbf{v}}{c} \frac{GM}{R} = \frac{\mathbf{v}\phi}{c}. \quad (2.6)$$

where we have taken r as the radial distance from the local object to a spherical volume element (of thickness dR), G is Newton's constant of gravitation, and M and R are the mass and radius of the universe respectively.

R was taken by Sciama as the radius of the “Hubble sphere,” that is, the product of the speed of light and the age of the universe. A more accurate calculation would have employed the “particle horizon,” the sphere centered on Earth within which signals traveling at the speed of light can reach Earth. The particle horizon encompasses considerably more material than the Hubble sphere. Sciama also neglected the expansion of the universe.

These issues notwithstanding, Sciama's work triggered an at times intense debate about the origin of inertia. Why? Because when we put the result of the integration in Eq. 2.6 back into Eq. 2.3, we get:

$$\mathbf{E}_g = -\nabla\phi - \frac{1}{c} \frac{\partial \mathbf{A}_g}{\partial t} = -\nabla\phi - \frac{\phi}{c^2} \frac{\partial \mathbf{v}}{\partial t}. \quad (2.7)$$

Now, we return to the consideration of our object moving with velocity \mathbf{v} with respect to the homogenous and isotropic universe that we can envisage as moving *rigidly* with velocity $-\mathbf{v}$ past the object which is taken as (instantaneously) at rest. In this case the gradient of the scalar potential vanishes. And if \mathbf{v} is constant or zero, so, too, does the second term – and there is no gravelectric field felt by the object.

However, if the object is accelerating with respect to the rest of the universe (due to the application of some suitable “external” force), then the second term does not vanish as $\partial \mathbf{v} / \partial t = \mathbf{a}$, the acceleration, is not zero. More importantly, from the point of view of the origin of inertia – and inertial reaction forces – if $\phi / c^2 = 1$, then the gravelectric field exactly produces the “equal and opposite” inertial reaction force the accelerating agent experiences. That is, *inertial reaction forces are exclusively gravitational in origin*. The reason why this was so intriguing is that the condition $\phi / c^2 = 1$ has special cosmological significance, as we will consider presently.

Clearly, Sciama's calculation is an approximation. In particular, it is a vector approximation to a field theory that was known to require tensor form in order to be completely general. And it is an idealization. Sciama's assumptions about the distribution and motion of the “matter” sources of the gravelectric field at the object considered are much simpler than reality, even in the early 1950s, was known to be. Nevertheless, Sciama's theory is *not* a “toy model.” Toy models are created by physicists when they can't formulate their theory in tractable form in the full four dimensions of real spacetime. To make their theories tractable, they generate them with one or two spatial dimensions where the math is simple enough to be managed. Sciama's theory is four-dimensional. And the above calculation returns an answer for inertial reaction forces that is essentially correct despite the approximation and idealizations adopted. The part of Sciama's paper “On the Origin of

Inertia” where he calculates this expression is reproduced as Addendum #1 at the end of this chapter.

It is worth noting here that an important feature of inertial reaction forces is present in Eq. 2.7, and it was noted by Sciama. The two terms on the right hand side of the equation have different dependencies on distance. The scalar potential depends on the inverse first power of the distance. The gradient of the scalar potential, when you are far enough away from a body of arbitrary shape so that it can be approximated as a sphere, depends on the inverse second power of the distance. That is, Newtonian gravitational force exerted by a body on another sufficiently distant goes as the inverse square of the distance separating them.³

When you are calculating the effect of distant matter on a local object, inverse square dependence applies for the gradient of the scalar potential. And it drops off fairly quickly. The term arises from the time-derivative of the vector potential scales with the scalar potential, not its gradient. So the distance dependence of this term is inverse first power. When the distances involved in a situation are small, this difference between the terms may be unimportant. When the distances are large, the difference is crucial. The term arising from the vector potential dominates because it doesn't decrease nearly as rapidly as the Newtonian term does for large distances. This is the reason why the inertia of local objects is due almost exclusively to the action of distant matter.

The inverse first power of the distance dependence of the term from the vector potential that causes inertial forces also signals that the interaction is “radiative.” That is, the interactions that arise from this term involve propagating disturbances in the gravity field. They do *not* arise from instantaneously communicated effects or the passive action of a pre-existing field. So inertial forces would seem to be gravity “radiation reaction” effects. This poses a problem, for an inertial reaction force appears at the instant an accelerating force is applied to an object. How can that be true if the inertial reaction force involves an active communication with chiefly the most distant matter in the universe, and communication with the stuff out there takes place at the speed of light?

If reaction forces were produced by the interaction with a passive, locally present pre-existing field, this would not be a problem. But that is not what is calculated in Sciama's treatment. The trick of using the instantaneous frame of rest where the universe very obviously appears to be moving rigidly past the accelerating object not only sidesteps a messy calculation involving Green's functions; it blurs the issue of instantaneity of reaction forces. This is arguably the most difficult aspect of coming to grips with the origin of inertia.

You may be wondering, if this sort of thing happens with gravity, why don't we see the same sort of behavior in electromagnetism? After all, if we accept Sciama's theory as the vector approximation to GRT that it is, they are both vector field theories with essentially

³Newton is routinely credited with the discovery of the inverse square law of universal gravitation. But his contemporary Robert Hooke claimed to have independently discovered the inverse square law before Newton made public his claim. Newton refused the presidency of the Royal Society until shortly after Hooke's death. Shortly thereafter, the Royal Society moved to new quarters, and Hooke's papers from the 1680s were lost in the move. Whether Hooke actually discovered the inverse square nature of gravity, absent his papers, is a matter of conjecture. It seems unlikely, though, that he discovered the universal nature of the interaction.

the same field equations. Ironically, as it turns out, the problems of the origin of inertia – in the form of electrical “self-energy” – and “radiation reaction” have plagued electrodynamics for years, too. It just hasn’t been discussed much in recent years. But infinite “self-energies” of point particles was the motivation, for example, for the invention the “renormalization” program of quantum field theory, and of string theory.

We’ll be looking at these issues in later chapters in some detail. Here we note that although the vector field formalisms for gravity and electromagnetism are essentially the same, this type of gravitational force from the action of cosmic matter does not arise in electrodynamics – because on average the universe is electric charge neutral, so cosmic electric charge currents sum to zero everywhere. More specifically, since on average there is as much negative electric charge as positive in any region of spacetime, the total charge density is zero. So, in the calculation of the vector potential – as in Eq. 2.5 – since ρ is zero, the integral for the potential vanishes. This means that in everyday electrodynamics you never have to deal with the action of distant electric charge and currents of any significance. But in gravity, you do.

Sciama’s calculation is not optional. It is a prediction of GRT providing that $\phi/c^2 = 1$. Is $\phi/c^2 = 1$ true?

Yes. When is $\phi/c^2 = 1$? When “critical cosmic matter density” is reached, and space at the cosmic scale is flat. Sciama didn’t know if this were true. Indeed, even in the 1950s it was thought that the amount of luminous matter in the universe was not sufficient to be “critical.” So Sciama did not make a bald-faced claim that he could fully account for inertial reaction forces. But space at the cosmic scale sure looked pretty flat. And it was known that if cosmic scale space deviated from flatness, it would quickly evolve to far greater distortion. As the universe was at least billions of years old and still flat, most cosmologists assumed that space really was flat, and that critical cosmic matter density was obtained. And the fact that luminous matter was less than 10% of the critical value came to be called the “missing mass” problem.⁴ Only after the turn of the century was space at the cosmic scale measured – by the Wilkinson Microwave Anisotropy Probe (WMAP) about a decade ago. So we know whether or not cosmic scale space is flat. It is.

You may be wondering, if we know that space at the cosmic scale is flat, why isn’t it common knowledge that inertial reaction forces are caused by the gravitational interaction of local accelerating objects with chiefly cosmic matter? Well, two issues figure into the answer to this question. One is the consequence of an analysis done by Carl Brans in the early 1960s. (Excerpts from Brans’ paper are to be found at the end of this chapter.) And the other, related to Brans’ argument, is the business about there being no “real” gravitational forces. Brans showed that if the presence of “spectator” matter (concentrations of matter nearby to a laboratory that shields the stuff in it from all external influences except gravity, which cannot be shielded) were to change the gravitational potential energies of objects in the shielded laboratory, you could always

⁴ Actually, the “missing mass” problem was first identified in the 1930s by Fritz Zwicky by applying the “virial theorem” to clusters of galaxies. The virial theorem says that on average, the kinetic and potential energies of galaxies in clusters should be the same. So, by measuring the motions of galaxies in a cluster, you can estimate the mass of the cluster. It leads to galaxy cluster mass estimates 10–100 times greater than the light emitted suggests is present. Only later was it extended to encompass cosmology, too.

tell whether you were in a gravity field or an accelerating lab in deep space by performing only local experiments.

In particular, the gravitationally induced changes in the masses of elementary particles in the lab would change their charge to mass ratios, and this would be locally detectable. No such changes in charge to mass ratios would occur in an accelerated reference frame in deep space. As a result, a gravity field could always be discriminated from an acceleration with local experiments. Since this would be a violation of the Equivalence Principle, Brans asserted that gravitational potential energy cannot be “localized.” That is, the scalar gravitational potential must have exactly the same value, whatever it might be, everywhere in the laboratory, no matter where the lab is located or how it is accelerating. As Brans noted, this condition on gravitational potential energy reveals Einstein’s first prediction quoted above as wrong. Evidently, it appears that the distribution of matter outside of the lab cannot have any identifiable effect on the contents of the lab. Mach’s principle, however, would seem to suggest the opposite should be the case. And it was easy to infer that Mach’s principle was not contained in pristine GRT.

The inference that Mach’s principle is not contained in GRT, however, is mistaken. If you take account of the role of the vector potential in Sciama’s gravelectric field equation,⁵ it is clear that should spectator matter outside the lab be accelerated, it will have an effect on the contents of the lab, changing what are perceived to be the local inertial frames of reference. This is the action of Mach’s principle. But as the accelerating spectator matter will act on all of the contents of the lab equally, for inertial forces are “fictitious,” they produce the same acceleration irrespective of the mass of the objects acted upon. So, using local measurements in the lab it will not be discernible either as a force of gravity or a change in the acceleration of the lab. And it will not change the gravitational potential energies of the contents of the lab.

Brans’ argument about the localizability of gravitational potential energy has an even more radical consequence – one found in the excerpt from Misner, Thorne, and Wheeler on energy localization in the gravitational field found in the previous chapter. If you can eliminate the action of the gravitational field point by point throughout the laboratory by a careful choice of geometry that, for us external observers, has the effect of setting inertial frames of reference into accelerated motion with respect to the walls, floor and ceiling of the lab, it seems reasonable to say that there is no gravitational field, in the usual sense of the word, present in the lab. This is what is meant when people say that GRT “geometrizes” the gravitational field. In this view there are no gravitational forces. Gravity merely distorts spacetime, and objects in inertial motion follow the geodesics of the distorted spacetime. The only real forces in this view are non-gravitational. Inertia, of course, is a real force. But if you believe that there aren’t any real gravitational forces, then the origin of inertia remains “obscure” – as Abraham Pais remarked in the quote at the outset of this chapter – for it isn’t a result of the electromagnetic, weak, or strong interactions (and can’t be because they are not universal), and that leaves only gravity.

⁵ Or Einstein’s vector approximation equation for the force exerted by spectator matter that is accelerating on other local objects.

But we've excluded gravity because we know that there aren't any gravitational forces. And the origin of inertia remains a mystery.

There may not be any "real" gravitational forces in GRT, but there is "frame dragging." That is, in the conventional view, matter can exert a force on spacetime to produce frame dragging, but it can't act directly on the matter in the possibly dragged spacetime. If this sounds a bit convoluted, that's because it is. Let's illustrate this point.

About the time that Thorne and his graduate students were introducing the rest of us to traversable wormholes, a committee of the National Academy of Sciences was doing a decadal review of the state of physics, producing recommendations on the areas of physics that should be supported with real money. One of their recommendations was that Gravity Probe B should be supported because, allegedly, no other test of "gravitomagnetism" was contemplated, and this was an important, if difficult and expensive, test of GRT.

Ken Nordtvedt, a physicist with impeccable credentials who had proposed the "Nordtvedt effect,"⁶ then being tested by ranging the distance of the Moon with a laser, but who had not been a member of the decadal survey committee, pointed out that the claim was just wrong. He noted that even in doing routine orbit calculations, unless care was taken to use special frames of reference, one had to take account of gravimagnetic effects to get reasonable results. Using "parameterized post Newtonian" (PPN) formulation of gravity, a formalism that he and others had developed as a tool to investigate a variety of theories of gravity some 20 years earlier, he showed explicitly how this came about.

In the course of his treatment of orbital motion, Nordtvedt drew attention to the fact that gravity predicts that linearly accelerated objects should drag the spacetime in their environs along with themselves since the gravitational vector potential does not vanish.⁷ Nordtvedt's 1988 paper on the "Existence of the Gravitomagnetic Interaction" where he discussed all this is excerpted in [Addendum #3](#) at the end of this chapter. In effect, he recovered the same basic result as Einstein and Sciama, only where they had talked about gravitational forces acting on local objects, Nordtvedt put this in terms of "frame dragging."⁸

Are they the same thing? Well, yes, of course they are. The reason why you may find this confusing is because in the case of everything except gravity, one talks about the sources of fields, the fields the sources create, and the actions of fields *in* spacetime on other sources. That is, spacetime is a *background* in which sources and fields exist and interact. In GRT spacetime itself is the field. There is no background spacetime in which the gravitational field exists and acts. Since there is no background spacetime, GRT is called a "background independent" theory.

⁶ The Nordtvedt effect proposes that gravitational potential energies do contribute to the mass-energy of things and predicts (small) deviations from the predictions of GRT that would follow. Such effects have not been observed.

⁷ He also predicted that the masses of things should vary as they are accelerated, an effect of the sort that we'll be looking at in the next chapter.

⁸ Nordtvedt considered only a rigid sphere of uniform density of modest dimensions. He did not extend the argument to the case where the sphere is the entire universe, as did Sciama.

It is this background independence that makes gravity and GRT fundamentally different from all other fields. And it is the reason why “frame dragging” is fully equivalent to the action of a gravitational force. If you want to preserve the configuration of a system before some nearby objects are accelerated, when the nearby objects begin to accelerate you have to exert a force that counteracts the effect of the frame dragging produced by the acceleration of the nearby objects. When you do that, what do you feel? An inertial reaction force – the force produced by the action of the dragged spacetime, which is produced by the gravitational action of the accelerated nearby objects. By interposing frame dragging we’ve made it appear that no gravitational force is acting. But of course gravity is acting, notwithstanding that we’ve introduced the intermediary of frame dragging to make it appear otherwise.

When only nearby objects are accelerated to produce frame dragging, as Einstein noted for the equivalent force he expected, the predicted effects are quite small. When it is the universe that is accelerated, it is the full normal inertial reaction force that is felt if you constrain some object to not accelerate with the universe. Why the difference? Because when the entire universe is “rigidly” accelerated, the interior spacetime is rigidly dragged with it, whereas nearby objects, even with very large masses, produce only small, partial dragging.

You may be thinking, yeah, right, rigidly accelerating the whole universe. That would be a neat trick. Getting the timing right would be an insuperable task. The fact of the matter, nonetheless, is that you can do this. We all do. All the time. All we have to do is accelerate a local object. Your fist or foot, for example. The principle of relativity requires that such local accelerations be equivalent to considering the local object as at rest with the whole universe being accelerated in the opposite direction. And the calculation using the PPN formalism for frame dragging (with GRT values for the coefficients in the equation assumed) bears this out. At the end of his paper on gravimagnetism Nordtvedt showed that a sphere of radius R and mass M subjected to an acceleration \mathbf{a} drags the inertial space within it as:

$$\delta\mathbf{a}(\mathbf{r}, t) = -\left(2 + 2\gamma + \frac{\alpha_1}{2}\right) \frac{U(\mathbf{r}, t)}{c^2} \mathbf{a} \quad (2.8)$$

where the PPN coefficients have the values $\gamma = 2$ and $\alpha_1 = 0$ for the case of GRT and $U(\mathbf{r}, t)$ is the Newtonian scalar potential, that is, $U = GM/R$. So we have four times ϕ (changing back to the notation of Sciamma’s work on Mach’s principle) equal to c^2 to make $\delta\mathbf{a} = \mathbf{a}$ in Eq. 2.8; that is, if the universe is accelerated in any direction, spacetime is rigidly dragged with it, making the acceleration locally undetectable.

You may be concerned by the difference of a factor of 4 between the Nordtvedt result and Sciamma’s calculation. Factors of 2 and 4 are often encountered when doing calculations in GRT and comparing them with calculations done with approximations in, in effect, flat spacetime. In this case, resolution of the discrepancy was recently provided by Sultana and Kazanas, who did a detailed calculation of the contributions to the scalar potential using the features of modern “precision” cosmology (including things like dark matter and dark energy, and using the particle horizon rather than the Hubble sphere), but merely postulating the “Sciamma force,” which, of course, did not include the factor of 4 recovered in Nordtvedt’s calculation. They, in their relativistically correct

calculation, found ϕ to have only a quarter of the required value to make the coefficient of the acceleration equal to one. Using the general relativistic calculation, with its factor of 4 in the coefficient, makes the full coefficient of the acceleration almost exactly equal to one – as expected if Mach's principle is true.⁹

You might think that having established the equivalence of frame dragging by the universe and the action of inertial forces, we'd be done with the issue of inertia. Alas, such optimism is premature. A few issues remain to be dealt with. Chief among them is that if $\phi = GM/R$, since at least R is changing (because of the expansion of the universe), it would seem that $\phi = c^2$ must just be an accident of our present epoch. However, if the laws of physics are to be true everywhere and during every time period, and inertial reaction forces are gravitational, then it must be the case that $\phi = c^2$ everywhere and at all times if Newton's third law of mechanics is to be universally valid.

Well, we know that the principle of relativity requires that c , when it is locally measured, has this property – it is a “locally measured invariant.” So, perhaps it is not much of a stretch to accept that ϕ is a locally measured invariant, too. After all, GM/R has dimensions of velocity squared. No fudging is needed to get that to work out right. But there is an even more fundamental and important reason to accept the locally measured invariance of ϕ : it is the central feature of the “Einstein Equivalence Principle” (EEP) that is required to construct GRT. As is universally known, the EEP prohibits the “localization” of gravitational potential energy. That is, it requires that whenever you make a local determination of the total scalar gravitational potential, you get the same number, whatever it may happen to be (but we know in fact to be equal to c^2). Note that this does *not* mean that the gravitational potential must everywhere have the same value, for distant observers may measure different values at different places – just as they do for the speed of light when it is present in the gravity fields of local objects. Indeed, this is not an accident, because ϕ and c are related, one being the square of the other.

Should you be inclined to blow all of this off as some sort of sophistry, keep in mind that there is a compelling argument for the EEP and the locally measured invariance of ϕ – the one constructed by Carl Brans in 1962 that we've already invoked. If you view the gravitational field as an entity that is present in a (presumably flat) background spacetime – as opposed to the chief property of spacetime itself (as it is in GRT) – it is easy to believe that gravitational potential energies should be “localizable” – that is, gravitational potentials should have effects that can be detected by local measurements. Brans pointed out that were this true, it would be a violation of the principle of relativity as contained in the Equivalence Principle. Why? Because, as mentioned above, you would always, with some appropriate local experiment, be able to distinguish a gravitational field from accelerated frames of reference.

⁹See: J. Sultana and D. Kazanas, arXiv:1104.1306v1 (astro-ph.CO, later published in the *Journal of Modern Physics D*). They find that the “Sciama” force is one quarter of that needed for an exact inertial reaction force. The factor of 4 discrepancy arises from the fact that Sultana and Kazanas simply assumed the “Sciama” force without deriving it from GRT, and Sciama's calculation is not exactly equivalent to a general relativistic calculation like Nordtvedt's. The difference is the factor of 4 that when multiplied times their result returns 1 almost exactly.

42 Mach's principle

Brans' way was to measure the charge to mass ratios of elementary particles. An even simpler, cruder way to make the discrimination between gravity field and accelerated reference frame is to drop stuff. You won't be able to tell the difference between a gravity field and accelerated frame of reference by the way things "fall" since they all "fall" with the same acceleration in both cases, irrespective of their masses or compositions. But you will be able to tell by how big a dent in the floor they make – because their masses are presumably different when gravity is present, versus when it is not, and bigger masses make bigger dents. Brans' argument makes clear that the EEP must be correct if the principle of relativity is correct – and that Einstein was wrong in 1921 when he assumed that the piling up of spectator matter would change the masses of local objects. Notwithstanding that the non-localizability of gravitational potential energies, however, the fact that inertial reaction forces are independent of time and place requires that the masses of things be equal to their total gravitational potential energies. That is, $E = mc^2$ and $E_{grav} = m\phi$, so if $E = E_{grav}$ and $\phi = c^2$ as Mach's principle demands, we have a simple identity.

ANOTHER EXAMPLE

To bring home the full import of the foregoing discussion of GRT and Mach's principle, we briefly consider a slightly more complicated example than that used so far. Instead of considering a test body in an otherwise uniform universe, we look at the behavior of a test object (with negligible mass) in the vicinity of Earth. In Newtonian physics we say that the mass of Earth produces a gravitational field in its vicinity that exerts a force on the test object. If the test object is unconstrained, it falls toward the center of Earth with an acceleration of one "gee." We can arrest this motion by applying an upward force with equal magnitude, balancing the "force" of gravity. The agent applying the upward balancing force, of course, experiences the downward force which he or she attributes to Earth's gravity. This is the commonplace explanation of these circumstances that even relativists intuitively recognize.

The general relativistic explanation of the circumstances of our test body in proximity to Earth, however, is *fundamentally* different. Earth does *not* produce a gravity field that acts to produce a force on the test body. Earth does produce a local distortion of spacetime (which is the gravity field), changing the local inertial structure of spacetime from the otherwise flat character it would have (as measured by the WMAP project). As a result, if our test body engages in unconstrained motion, it responds inertially and finds itself in a state of free fall. Despite the fact that the test body appears to us to be accelerating, and we intuitively assume that accelerations are the consequence of the application of forces, *no forces act on the falling test body*.

What happens, then, when we apply a constraining force to the test body to stop its free fall acceleration? Does this somehow turn on Earth's gravity force to balance the constraining force we have applied? No. You can't turn gravity off and on (yet). *The balancing force that you feel is the inertial reaction force that arises in response to the "arresting" force that you have applied to the test object*. Your arresting force has actually produced acceleration of the test object – with respect to local inertial frames of reference

that are in free fall. The force that we normally ascribe to the gravitational action of Earth, which is quite real, is not produced by Earth. It is produced chiefly by the distant matter in the universe. The reason why we associate it with the action of Earth is because Earth determines the extent of the local distortion of inertial spacetime, and thus the amount of acceleration required to arrest the inertial motion of objects in the vicinity of Earth's surface.

One may ask: is it really necessary to adopt this arguably very odd way of looking at the circumstances that seem to make such intuitive sense when viewed from the Newtonian point of view? That is, can we in some sense accept GRT, but take the above description as an "equivalent representation" to the Newtonian viewpoint with its objective gravity field that produces forces on nearby objects? No. The representations are in no sense equivalent. The reason why is the EEP. The geometrization of the gravitational field in GRT depends on the complete indistinguishability of accelerated reference frames from the local action of gravity fields.

There are those who argue that the presence of tidal effects in all but (unphysical) uniform gravity fields always allow us to distinguish gravity fields from accelerated reference frames, but this is a red herring. We can always choose our local Lorentz frame sufficiently small so as to reduce tidal effects to insignificant levels, making the two types of frames indistinguishable. Were gravitational potential energies localizable, however, we would be faced with a real violation of the indistinguishability condition that would vitiate field geometrization. Using either Brans' charge to mass ratios, or the cruder dents criterion, no matter how small we make the region considered, we can always make determinations that tell us whether we are dealing with a gravity field or an accelerated reference frame, because, unlike tidal forces, charge to mass ratios and dents don't depend on the size of the region considered. They are so-called "first" or "lowest" order effects.

The foregoing considerations are sufficient in themselves to reject attempts to "objectify" static gravity fields. But they are attended by an even stronger argument. If local gravitational potential energies really did contribute to locally observable phenomena, then $\phi/c^2 = 1$ everywhere and at all times would *not* in general be true. Consequently, inertial reaction forces would not always equal "external" applied forces, and Newton's third law would be false. That would open the way to violations of the conservation of energy and momentum. If you're trying to make revolutionary spacecraft, you may not think this necessarily bad. It is.

As we have now seen, the principle of relativity has present within it a collection of interlocking principles – one of which is Mach's principle, which says both that inertial reaction forces are the gravitational action of everything in the universe, and the inertia of objects is just their total gravitational potential energy (divided by c^2). Objects are to be understood as including everything that gravitates (including things we do not yet understand in detail like dark matter and dark energy). Are these principles ones that can be individually rejected if we don't like them without screwing up everything else? No. If the principle of relativity is correct, then the EEP and Mach's principle follow inexorably. If either the EEP or Mach's principle is false, then so, too, is the principle of relativity – and Newton's laws of mechanics. That's a pretty high price to pay for rejecting a principle you may not care for.

Two issues remain to be addressed in a little detail. One is the instantaneity of inertial reaction forces. The other is how Mach's principle squares with traditional gravity wave physics. We address inertial reaction forces and how they relate to gravity wave physics first.

INERTIAL REACTION FORCES AND GRAVITY WAVE PHYSICS

It has been known since Einstein created GRT in 1915 that his theory predicted propagating disturbances in the gravitational field, that is, it predicted "gravitational waves." The whole business of gravity waves and how they are generated by and interact with matter sources, however, was at times quite contentious. Should you want to know the details of how all of this developed, Dan Kennefick has written an outstanding history of the subject: *Traveling at the Speed of Thought: Einstein and the Quest for Gravitational Waves*.

Most, if not all, of the issues of debate were settled many years ago now. One of the issues was the manner in which the prediction is calculated. As noted above, exact solutions of the full non-linear Einstein field equations are few and far between. One of the standard techniques for dealing with this is to invoke the "weak field approximation," where you assume that the Einstein tensor (describing the geometry of spacetime) can be written as the "Minkowski" metric of flat spacetime with an added "perturbation" metric field that accounts for gravity, as mentioned earlier in this chapter. Since the flat spacetime metric in this approach is effectively a "background" spacetime unaffected by the presence of matter and gravity fields, Einstein's theory is effectively "linearized" by this procedure. With a few further assumptions, Einstein's field equations can be put in a form that closely resemble Maxwell's equations for the electromagnetic field – as Einstein himself did in his discussion of Mach's principle mentioned above, and Sciama and Nordtvedt (among many others) subsequently did.

Solutions of Maxwell's equations have been explored in great detail in the roughly century and a half since their creation. The standard techniques include classification according to the disposition of the sources of the fields and their behavior (how they move). This leads to what is called a "multipole expansion" of the field, each component of the field being related to a particular aspect of the distribution and motion of its sources. The simplest part of the field in this decomposition is the so-called "monopole" component, where the sources can be viewed as consisting of a single "charge" located at one point in spacetime.

In electromagnetism the next least complicated source distribution is the so-called "dipole" component. Electrical charges come in two varieties: positive and negative, and the dipole component of a multipole expansion consists of the part that can be characterized by a positive charge located at one point and a negative charge located somewhere else in spacetime. The measure of this charge distribution is called its dipole "moment," defined as the product of the charges times the separation distance between them. If the dipole moment of the dipole component of the field is made to change, the changes in the surrounding field are found to propagate away from the charges at the speed of light. The propagating disturbance in the field is the "radiation" field. Non-propagating

fields are called “induction” fields, as they are induced by the presence of sources of the field and do not depend on their moments changing.

The next term in the multipole expansion for source distributions and their associated field components is the so-called “quadrupole” term. It is the part of the field that takes into account the simplest charge distribution for sources of the same sign (positive or negative in the case of electromagnetism) that cannot be covered by the monopole term. It corresponds to two charges of the same sign separated, like the dipole distribution, by some distance in spacetime. Just as there is a dipole moment, so, too, is there a quadrupole moment. And if the quadrupole moment changes, like the dipole term, a propagating disturbance in the field is produced.

Since there are no negative masses (yet), and the vector approximation of GRT is a vector theory analogous to Maxwell’s equations for electrodynamics, it is found that the “lowest order” radiative component of the gravitational field is that produced by sources with time-varying quadrupole moments. An example is a dumbbell spinning about the axis of symmetry that passes perpendicularly through the bar separating the bells. Another more fashionable example is a pair of black holes in orbit around each other. An example that does not involve spinning stuff is two masses separated by a spring that are set into oscillatory motion along their line of centers. Even in the case of orbiting black holes, the amount of momenergy involved in the gravitational radiation is exceedingly minute. (This is the stuff being sought with the Laser Interferometer Gravitational wave Observatory, with a price tag now approaching a gigabuck.) Laboratory scale gravitational quadrupoles, even operating at very high frequencies, produce hopelessly undetectable amounts of gravitational radiation.¹⁰

What does all this have to do with inertial reaction forces? Well, as Sciama was at pains to point out, his calculation of those forces show two things: one, they depend on the acceleration of sources; and two, their dependence on distance in his gravelectric field equation goes as the inverse first power, not inverse square. These are the well-known signatures of radiative interactions. It would seem then that inertial reaction forces should involve radiation, and that they should be called radiation reaction forces. But there is a problem. The quadrupole radiation given off by an accelerating massive object is incredibly minute. And the monopole component of the field in electrodynamics is non-radiating. How can this be squared with the fact that inertial reaction forces are, by comparison, enormous, decades of orders of magnitude larger than quadrupole radiation reaction? To answer this question we must first tackle the instantaneity of inertial reaction forces.

¹⁰ The field strength of gravitational radiation depends on the frequency at which it is emitted. Gravitational waves, all other things held constant, depend on the fifth power of the emission frequency. This strong frequency dependence has led some to speculate that very high frequency gravitational waves might be used for propulsive purposes. Since the momenergy in gravity waves produced by human scale sources is so hopelessly minute, even allowing for unrealistically high frequency sources, gravity waves hold out no promise of practical scale effects.

THE INSTANTANEITY OF INERTIAL REACTION FORCES

The immediate fact of inertial reaction forces is that they respond to applied forces instantaneously. Why? Well, if you believe, as Newton and legions after him have, that inertia is an inherent property of material objects needing no further explanation, then this question needs no answer. The problem with this view, of course, is the fact noted famously by Mach that inertial frames of reference seem to be those in inertial motion with respect to the “fixed stars.” Today we would say inertial motion with respect to the local cosmic frame of rest, and that, remarkably, isn't rotating. This suggests that the stuff out there has something to do with inertia. But it is so far away, typically billions of light-years distant. How can that produce instantaneous effects?

The easy answer to this question is to assert that the distant stuff produces a gravity field, which we know to *be* spacetime in GRT, here, and when we try to accelerate anything in spacetime, spacetime pushes back. Since the local spacetime is the gravity field of the distant stuff, obviously we should expect local inertia to be related to the distant stuff. This is the “local pre-existing field” argument.

Sounds good, doesn't it? It is, however, a flawed view of things, as was made evident by Sciama's argument back in the early 1950s. As we've noted already, Sciama used a little trick to avoid a tedious calculation involving Lienard-Wiechert potentials, Green's functions, and a lot of associated mathematical machinery. To calculate the effect of very distant matter on a local accelerating body, he noted that from the perspective of the local body, the entire universe appears to be accelerating rigidly in the opposite direction. The apparent rigid motion provides the justification for removing the velocity from the integral for the vector potential. Sciama, of course, knew that this was just a trick to avoid a messy integration, for, as already mentioned, he was quick to point out that distance dependence of the scalar gravitational potential was inverse *first* power, rather than the inverse *second* power of Newtonian gravity. Those familiar with the process of radiation immediately recognize the inverse first power as the signature of a radiative interaction. What Sciama's calculation (and those of Einstein, Nordtvedt, and others) shows is that inertial reaction forces are conveyed by a radiative process. Inertial forces are not the simple passive action of a pre-existing field that acts when local objects are accelerated.

A way to visualize what's going on here is to consider what happens to the spacetime surrounding a local object that is given a quick impulsive acceleration. Before the acceleration, its gravity field is symmetrically centered on it. The same is true shortly after the impulse. But the impulse displaces the center of symmetry of the field from the prior center of symmetry. That produces a “kink” in the gravity field, like that shown in Fig. 2.1. The radiative nature of the interaction means that the kink induced in the field by the impulsive acceleration¹¹ propagates outward from the object during the acceleration at the speed of light.

It is the production of the kink in the field by the source, not the field itself, that produces the inertial reaction force on the source and accelerating agent. In electrodynamics, this is known as the problem of “radiation reaction.” Should you trouble yourself to

¹¹ The technical term for such an acceleration is a “jerk.”

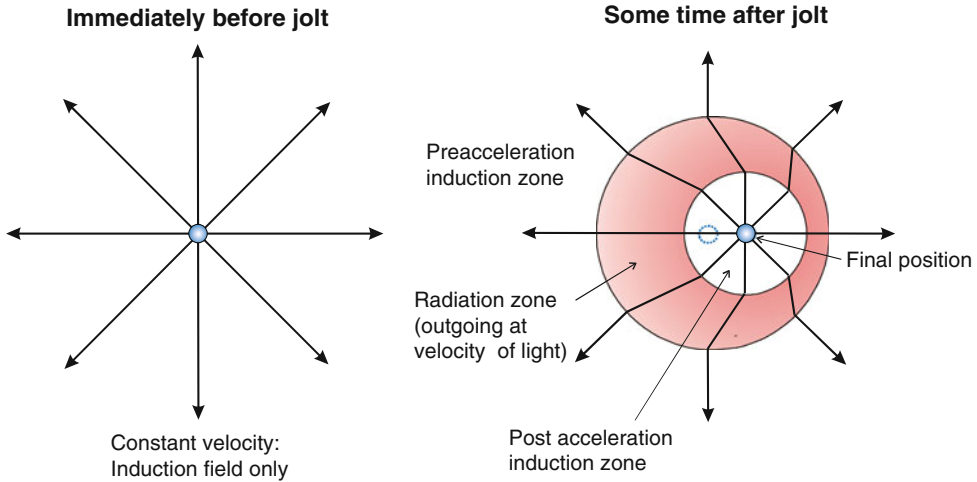


Fig. 2.1 The “kink” diagram. When a source of the gravitational field is in a state of inertial motion, it carries its field, represented by the *lines* radiating from the source’s location, along with it without distortion. If the source is sharply accelerated, and then decelerated, so that it moves after the “jolt” as it did before, a “kink” is introduced into the field lines. The kink does not move to infinity at once. It propagates outward at the speed of light

read up on this in, say, Feynman’s *Lectures on Physics*, or pretty much any advanced text on electrodynamics, you’ll find that this is a messy problem with some very curious features, for example, “pre-acceleration,” where an object starts to accelerate *before* the force producing the acceleration acts (as Dirac showed in a classic paper on electromagnetic radiation reaction published in 1938). All those problems carry over to the gravity case if inertial reaction forces are forces of radiative reaction – as seems to be the case now that the WMAP results are known.

Now, there are two problems here. The first is that the kink in the field is normally taken as due to the monopole term in the multipole expansion, and it is allegedly non-radiative. We will deal with this issue presently. The second problem is that if the coupling between the test object and the distant matter in the universe is carried by the kink in the field propagating at the speed of light, it will take billions of years for the kink to reach the distant matter, and billions of years for a return signal to get back to the accelerating object. Inertial reaction forces, however, are instantaneous. Push something and it pushes back immediately. How can the distant matter in the universe act instantly on an object when it is accelerated by an external force without violating the speed limit, c , of SRT?

ACTION AT A DISTANCE AND “ADVANCED” WAVES

The simplest, most elegant way to deal with the problems just mentioned was worked out for electrodynamics by John Wheeler and Richard Feynman in the 1940s. Their theory, intended to deal with the problems attending classical electron theory (infinite self-

energies,¹² radiation reaction, and so forth), goes by the name “action-at-a-distance” or “absorber” electrodynamics. It is a scheme designed to account for seemingly instantaneous radiation reaction forces that are produced by an interaction with a distant “absorber.” To do this, Wheeler and Feynman noted that the propagating solutions to “classical” wave equations can either be “retarded” – that is, propagate forward in time – or “advanced” – that is, propagate backward in time.

Physically and mathematically, there is no discernible difference between the two classes of solutions. Since it appears to us that waves propagate into the future, we just ignore the solutions that propagate backward in time. After all, we do not appear to be constantly buffeted by waves coming back from the future.

The business of advanced waves can be a bit confusing, so we make a brief foray into this topic to ensure that we are all on the same page. The usual story about the role of time in the laws of physics is that the laws of physics possess a property called “time reversal symmetry.” That is, you can replace the time t with $-t$ everywhere in your equations, and the processes described by the time-reversed equations are just as valid as the original equations. Another way this is sometimes illustrated is to film some process running forward in time, and then point out that if the film is run backward, the processes depicted also obey the laws of physics, albeit the time-reversed laws.

The fact that the laws of physics are time-reversal invariant has led to endless speculations on “the arrow of time,” and how time could be asymmetric given the symmetry of the underlying laws. Philosophers, and physicists with a philosophical bent, seem to be those most prone to delving into the mysteries of time. We’ll be concerned here with a much more mundane problem: How exactly do advanced waves work?

A commonplace example used to illustrate advanced waves is the spreading of ripples on a pond when a rock is thrown into the middle. When the rock hits the water, it sets up a series of waves that propagate from the point of impact in symmetrical circles toward the shoreline. If we make a film of this sequence of events and run it backward, we will see the waves forming near the shoreline, and then moving in concentric circles of decreasing diameters toward the center. And when the waves arrive at the center, the rock will emerge from the water as though thrust from the depths by the waves. This wave behavior is illustrated in Fig. 2.2 as sequences of time-lapsed pictures of waves, with time proceeding from left to right. The normal view of things is shown in the upper strip of pictures, and the reversed in the lower strip.

The problem with this picture is that when we run the movie backward to supposedly reverse the direction of time, what we really do – since we can only run the movie forward in time, regardless of which end of the movie we start with – is run the waves *backward in space* as the movie runs *forward in time*. A true advanced wave starts in the future at the shoreline and propagates backward in time toward the center of the pond, something we

¹²Self energy in electrodynamics arises because the parts of an electric charge repel the other parts of the charge, and work must be done to compress the parts into a compact structure. The energy expended to affect the assembly is stored in the field of the charge. When the electron was discovered by J. J. Thomson in 1897, it was not long until H. A. Lorentz and others suggested that the electron’s mass might be nothing more than the energy stored in its electric field (divided by c^2). They used this conjecture to calculate the so-called “classical electron radius” that turns out to be about 10^{-13} cm. But should you assume that the size of the electron is zero, the energy of assembly turns out to be infinite.

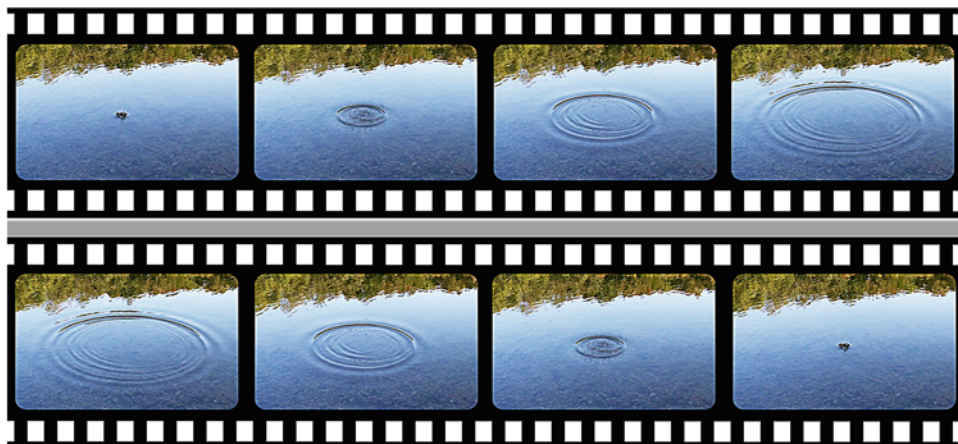


Fig. 2.2 The *top set of frames*, reading left to right, show waves propagating forward in time and space as they spread from a rock being thrown into a pond. When people talk about “advanced” waves, they often remark that waves propagating backward in time are those seen by running a movie of the waves in reverse, producing the sequence of pictures in the *bottom row*. However, the *bottom row* shows waves propagating backward in space as time goes forward

cannot actually see from the present. So, when we watch a movie running backward, we must imagine that we are running backward in time, notwithstanding that we are actually “moving” forward in time.

What we do see, moving forward in time, when an advanced wave comes back from the future is a wave that appears to be propagating away from the impact of the rock toward the shoreline of the pond. That is, the advanced wave looks exactly like a retarded wave. As long as the advanced wave coming back from the future didn’t propagate farther into the past than the rock hitting the water that initiated all of the waves, neither you nor I could tell whether the waves in the pond had any advanced component. So, using retarded and advanced waves to get distant objects to “instantaneously” affect local objects becomes finding a solution for wave action that cancels the advanced waves at the source (the rock hitting the water) to keep them from traveling farther into the past.

What Wheeler and Feynman noted was that if a forward in time propagating wave in the electromagnetic field was eventually absorbed by enough material out there in the distant universe, and as it was absorbed it produced an “advanced” wave propagating backward in time, all of the contributions from all of the parts of the absorber would just get back to the source at exactly the right time to produce the apparent force of radiative reaction. And as they passed the origin of the waves into the past, if the waves were half advanced and half retarded, they would cancel out the “advanced” wave propagating from the source into the past. So future events would not indiscriminately screw up the past (and our present). But the half-advanced waves coming back from the future provide a way for arbitrarily distant objects to affect events in the present seemingly instantaneously. In the case of gravity, this allows the whole universe to act on any object that’s accelerated by an external (non-gravitational) force with an equal and opposite force. This solution to the problems of radiation reaction is so neat it almost has the appearance of a cheap tourist trick, too good to be true. But it actually works.

Some exceedingly important features of action-at-a-distance electrodynamics must be mentioned, as they figure critically into the understanding of inertial reaction forces when the theory is extended to include gravity. Of these, far and away the most important is the fact that *there is no radiation as understood in conventional electrodynamics in the action-at-a-distance version*. It has not been mentioned yet, but in addition to acceleration dependence and inverse first power of the distance dependence of the “amplitude” or field strength¹³ of the radiation field, there is another condition that a radiation field must satisfy: it must have a “freely propagating” non-vanishing energy density as it approaches “asymptotic infinity.” This property gives the field “independent degrees of freedom.”

What this means, in simple physical language, is that once a radiation field has been launched by the acceleration of some charges of the field, the radiation is “decoupled” from both the source (which can no longer affect it) *and the sinks (just sources soaking up the field), if any, that ultimately absorb it*. Note that the launching of the radiation does *not* depend on it ever being absorbed by sinks out there somewhere in the future. That’s what “freely propagating at asymptotic infinity” means. Note, too, that there are no classical radiation fields in action-at-a-distance electrodynamics, for no electromagnetic disturbances (that might be considered radiation in classical theory) are ever launched without the circumstances of their eventual absorption being established *before* they are launched. That is, there are no field “modes” with “independent degrees of freedom,” no loose radiation that might make it to “asymptotic infinity.”

Why is this the case? Because the theory only works if the eventual absorption of all disturbances is guaranteed so that the requisite “advanced” disturbances, needed to combine with the “retarded” disturbances, are present to yield the world as we see it. What this means it that if your field theory is an action-at-a-distance theory, you can have “monopole” propagating disturbances in the field that carry energy and momentum – as the “kink” diagram suggests ought to be possible – and that they can have the acceleration and inverse first power of the distance characteristics of classical radiation, but they will not be considered “radiation” by those ignorant of action-at-a-distance theory.

You may ask at this point, how can such radically different results be obtained from action-at-a-distance and classical field theory? The answer is really quite simple. Michael Faraday, the pre-eminent experimental physicist of the nineteenth century, hated action-at-a-distance. In his day, it was the chief feature of Newtonian gravitation, and even Newton himself had thought that instantaneous action of gravity over arbitrarily large distances stupid.¹⁴ Indeed, Newton’s famous “hypotheses non fingo” [I make no hypotheses {about

¹³ The “amplitude” (for an oscillatory field) or “field strength” (the magnitude of the scalar potential or field vector) is not the same as the “intensity” of the field. The intensity is proportional to the square of the field strength. So, a field whose strength decreases as $1/r$ has an intensity that decreases as $1/r^2$, as does electromagnetic radiation (light), for example. When the intensity decreases at this rate, some energy just barely makes it to “asymptotic infinity.” If the intensity decreases faster than $1/r^2$, as it does for any field whose strength decreases more quickly than $1/r$, then no freely propagating energy makes it to asymptotic infinity.

¹⁴ As Faraday discovered in the early 1840s when Newton’s “third letter to Bentley” was first published. Hitherto, Newton’s true views on action-at-a-distance were not generally known. After reading Newton’s letter, it is said that Faraday became positively boorish regaling everyone with the news that Newton rejected action-at-a-distance.

the mechanism of gravity]] remark was his response to critics who assailed him about action-at-a-distance.

Faraday, despite repeated attempts, never found a way to rid gravity of action-at-a-distance. But he invented the field concept for electrodynamics to head off a similar fate for electrical and magnetic phenomena. Maxwell incorporated Faraday’s concept into his elaboration of the equations of electrodynamics. If you look at reality through “local” eyes, this approach makes eminent good sense. After all, you can wiggle some electric charges and launch an electromagnetic wave without giving any thought at all to what eventually happens to the wave. For all you know, it may well end up propagating freely at asymptotic infinity. If all you know about the totality of reality is emerging astronomical knowledge of the galaxy, as was the case through the early twentieth century, this is perfectly reasonable. But when you know more about cosmology, the know-nothing strictly local view is not so obviously reasonable.

How can the classical “local” view be squared with the action-at-a-distance picture? Well, we can’t just take some source distribution with a quadrupole moment, say, a dumbbell, create a time varying quadrupole moment by spinning the dumbbell or making the masses of the dumbbell accelerate along their line of centers with respect to each other. That will just give us back the idiotically small radiation calculated by gravity wave physicists. What’s missing in the dumbbell picture? The rest of the universe. How can we include it? By taking note of the fact that it acts seemingly instantaneously, so we can imagine that some non-negligible part of the whole universe is located in very close proximity to one (or the other) of our dumbbell masses.

The dumbbell mass, if you will, anchors the local system in the universe. And this anchoring mass must be present in any real system in order to accelerate the primary mass to produce the “monopole” kink in the field depicted in Fig. 2.1. That is, the idealization of a single mass that is accelerated is unrealizable, as there must always be a second reaction mass against which the accelerating agent acts to produce the acceleration of the primary mass. So all real accelerations necessarily involve quadrupoles.¹⁵ But when we are talking about the monopole kink in the field of one of the masses, the second mass of the equivalent quadrupole is a significant part of the mass of the universe. We can consider the mass of the universe effectively present at the second dumbbell mass because of the instantaneous action-at-a-distance character of inertial effects. The radiation produced by this quadrupole is decades of orders of magnitude larger than that for the local dumbbell quadrupole taken by itself. The reaction to the quadrupole radiation produced by the effective universe-dumbbell system is the inertial reaction force that acts on the dumbbell mass being accelerated.

There are obvious problems with carrying through a calculation of the sort just sketched. Concentrating a large fraction of the mass of the universe at a point in proximity to anything will recreate the initial singularity, and so on. But the point nonetheless

¹⁵ Two exceptions to this rule should be noted. First, a spherical object whose parts are undergoing a uniform radial acceleration does not radiate as the quadrupole moment is and remains zero. While such an expansion changes the radial tension in the field, it produces no “kink” in the field of the sort shown in Fig. 2.1. Second, there are those who hope to find a way to couple an object directly to the distant matter in the universe and produce accelerations without the need for an anchoring local mass. Such speculations are sometimes referred to as “field effect” propulsion. Hope springs eternal.

remains that if you insist on doing a standard quadrupole calculation, you've got to get the right quadrupole if you expect to get reasonable results. When you are considering inertial reaction forces, the right quadrupole always includes the effective universe, and it acts immediately and as if it were very, very nearby.

Wheeler and Feynman's elegant solution to the problem of radiation reaction is the only apparent way to get seemingly instantaneous reaction forces that depend on distant matter *without screwing up the dictum of the principle of relativity that limits signal propagation velocities to the speed of light*. Feynman may have harbored similar views, for he devoted the first part of his Nobel address to absorber electrodynamics.¹⁶ In electrodynamics you can hold either view, for the two are fully equivalent. But when you come to grips with Mach's principle, you find that this is the only convincing way to deal with inertial reaction forces while preserving the finite signal velocity required by the principle of relativity.

When Mach's principle was hotly debated in the 1960s, Fred Hoyle and Jayant Narlikar figured this out and wrote papers and a book on the subject. No one paid much attention, it seems.¹⁷ Their contemporaries may have been influenced by Hoyle's support for the "steady state" cosmology, which was then losing credibility. Wheeler's last book in the mid-1990s was an attempt to evade action-at-a-distance by invoking "constraint" equations on "initial data" that have instantaneous propagation (because they are "elliptic" rather than "hyperbolic"). Wheeler had abandoned the action-at-a-distance theory that he and Feynman had developed 50 years earlier. However, this should be evaluated keeping in mind that the *propagating kink in the field* is the field response to the acceleration of sources. Inertia is not just the action of the pre-existing gravity field on sources as they accelerate.

In the 1980s, John Cramer adapted the Wheeler-Feynman theory to quantum mechanics to explain "entanglement," another instance of seemingly instantaneous signal propagation that is customarily explained away in less than completely convincing ways. Cramer's "transactional interpretation" of quantum mechanics has not yet attracted widespread adherents. The culture of "shut up and calculate" has softened over the years. But serious examination of alternate interpretations of quantum mechanics has yet to make it into the mainstream of physics pedagogy.

Before Hoyle, Narlikar, and Cramer were others who saw the writing on the wall. Herman Weyl, the father of "gauge theory," famously remarked shortly after the first of the Wheeler-Feynman papers on action-at-a-distance electrodynamics, "Reality simply *is*, it does not *happen*." And Olivier Costa de Beauregard made early attempts to apply it to quantum theory.

The reason why the action-at-a-distance view of radiation reaction meets such stiff resistance is captured in Weyl's remark just quoted. The passage of time is an illusion.

¹⁶ When I read it as a grad student in the 1960s, I thought he was nuts. But Feynman knew what he was doing. Frank Wilczek recounts (in *The Lightness of Being*, pp. 83–84) a conversation with Feynman in 1982 about fields: "... He had hoped that by formulating his theory directly in terms of paths of particles in space-time – Feynman graphs – he would avoid the field concept and construct something essentially new. For a while, he thought he had. Why did he want to get rid of fields? 'I had a slogan, ... The vacuum doesn't weigh anything [dramatic pause] *because nothing's there!* ...'" Feynman initially thought that his path integral approach captured the chief feature of the action at a distance theory: no freely propagating radiation in spacetime.

¹⁷ Paul Davies, author of many popular books on physics, however, recounts in his *About Time* that it was attendance at one of Hoyle's lectures on this topic that set him on his early research career.

Indeed, “persistent illusion” was exactly the way Einstein characterized our notions of past, present, and future and the passage of time to the relatives of his lifelong friend Michel Besso after Besso’s death, shortly before his own. The past and the future are really out there. Really. Not probably. You may think that this must all be a lot of nonsense dreamed up by people who don’t have enough real work to fill their time. But let me point out that if absurdly benign wormholes are ever to be built and actually work, then this worldview *must* be correct. The past and the future must really “already” be out there. How can you travel to a past or future that doesn’t “already” exist?

THE “RELATIONAL” AND “PHYSICAL” VERSIONS OF MACH’S PRINCIPLE

Should you find the forgoing confusing and contentious, you’ll doubtless be disappointed to learn that we haven’t yet covered the full range of arguments involving Mach’s principle. As arguments about Mach’s principle developed over the decades of the 1950s, 1960s, and 1970s, two distinct ways of “interpreting” the principle emerged. One came to be called the “relationalist” view, and the other we shall call the “physical” view.

Serious arguments about Mach’s principle ceased to be fashionable in the mid-1970s. A few hardy souls wrote about the principle in the late 1970s and 1980s, but no one paid them much mind. Mach’s principle became fashionable again in the early 1990s, and Julian Barbour and Herbert Pfister organized a conference of experts in the field held in Tübingen in the summer of 1993. The proceedings of the conference were published as volume six of the *Einstein Studies* series with the title: *Mach’s Principle: From Newton’s Bucket to Quantum Gravity* (Birkhauser, Boston, 1994). This is an outstanding book, not least because the questions, comments, and dialog were published, as well as the technical papers presented.

Both the relationalist and physical positions on Mach’s principle were on display at the conference. Many of the attendees seem to have been convinced relationalists. The essence of the relationalist position is that all discussion of the motion of massive objects should be related to other massive objects; that relating the motion of objects to spacetime itself is not legitimate. This probably doesn’t sound very much like our discussion of Mach’s principle here. That’s because it isn’t. The relationalist approach says nothing at all about the origin of inertial reaction forces. The physical view of Mach’s principle, however, does. After the conference, one of the leading critics of Mach’s principle, Wolfgang Rindler, wrote a paper alleging that Mach’s principle was false, for it led to the prediction of the motion of satellites in orbit around planets that is not observed – that is, the motion was in the opposite direction from that predicted by GRT. It was 3 years before Herman Bondi and Joseph Samuel’s response to Rindler was published. They pointed out that while Rindler’s argument was correct, it was based on the relationalist interpretation of Mach’s principle. They argued that the physical interpretation that they took to be exemplified by GRT and Sciama’s model for inertia gave correct predictions. Therefore, Mach’s principle could not be dismissed as incorrect on the basis of satellite motion, as Rindler had hoped to do. It seems that Einstein was right in 1922, and Pais in 1982, when they remarked that Mach’s principle was a missing piece of the puzzle of the origin of inertia. We should now know better. After all, the WMAP results show that *as a matter of fact* space is flat, and it is certainly not empty, so if the principle of relativity, introduced by

Galileo, is right, then Mach's principle is correct, too. And we should simply drop all of the arguments and assumptions that distract us from this conclusion.¹⁸

MACH'S PRINCIPLE, STARSHIPS, AND STARGATES

You may be thinking that all of this Mach's principle stuff is just too confusing and contentious to take seriously. There must be another way – one with simple principles that no one argues about – to make starships and stargates. Sorry. No such luck. The only way to build starships and stargates is by making traversable absurdly benign wormholes. That can only happen when we understand the role of inertia in gravity. It might seem to you, if this is true, that we are doomed never to build such devices. It's now a 100 years since Einstein first tried to model Mach's ideas in a vector theory of gravity, and we seem no closer to getting a version of Mach's principle that might collect a consensus.

The problem here is that Mach's principle has been understood from the first days of general relativity to be essentially a cosmological problem. Look at Einstein's statement of the principle in the quote at the beginning of this chapter. The geometry must be fully specified in terms of the sources – that is, no solutions of the field equations should exist when there are no sources, or when other “non-Machian” conditions (like rotation of the universe) exist. The fact of the matter is that non-Machian, self-consistent solutions of Einstein's equations do exist. This has led some to the view that the principle should be taken to be a boundary condition on the cosmological solutions of Einstein's equations. But even this approach yields equivocal results.

Let's look at an example of what we're talking about. In the years before Alan Guth and others proposed the cosmological models containing the process of “inflation,” one of the outstanding issues of cosmology was the so-called “flatness” problem. The then prevailing preferred cosmological models – Friedman-Robertson-Walker (FRW) cosmologies – could be classified as “open” [expands forever] or “closed” [expands to some finite radius and then collapses] separated by a model that expands forever, but tends to zero expansion at temporal asymptotic infinity. The separating model is characterized by spatial flatness (and “critical” cosmic matter density) at all times. Even then (and now more so given the WMAP results), the universe looked very flat at cosmological scale. As noted above, the problem with the spatially flat model is that it is unstable. The slightest deviation from exact flatness produces very rapid evolution away from flatness – but the universe has been around for billions of years. The inflationary scenario invented by Guth and others, in fact, was intended to address precisely this problem.

¹⁸In this connection, Paul Davies relates an apposite story: “. . . I ventured: “What is the origin of the random phase assumption?” To my astonishment and dismay, [David] Bohm merely shrugged and muttered: “Who knows?”

“But you can't make much progress in physics without making that assumption,” I protested.

“In my opinion,” replied Bohm, “progress in science is usually made by *dropping* assumptions!”

This seemed like a humiliating put-down at the time, but I have always remembered these words of David Bohm. History shows he is right. . . .

Most cosmologists accept the inflationary model. But it doesn't have the status of a paradigm, not yet anyway. Other cosmological models are offered for other reasons. And there is a camp that argues that the consensus cosmology is wrong for other reasons. Hoping for a consensus to emerge on Mach's principle in such circumstances is simply not realistic. Frankly, the technical details of fashionable cosmological models are not important here. If we want to build starships and stargates, do we need to wait until cosmologists decide on some model and then see if it includes Mach's principle? No! Whatever that model, if it is ever found, turns out to be, it will be one with spatial flatness. Why? Because spatial flatness is *measured* to be the fact of our reality. Spatial flatness in FRW cosmologies guarantees "critical" cosmic matter density obtains, and that guarantees $\phi = c^2$. We know that for the EEP, Mach's principle, and Newton's third law to be true, this condition must be true everywhere and at every time in local measurements. And this *must* be true no matter what cosmological model you choose to believe in.

Now, building starships and stargates is not a matter of cosmology. It is a matter of using the law of gravity and inertia at the local level. We want to find a way to manipulate stuff we can lay our hands on and figure out how to make it produce effects that will make it possible to effectively induce outrageous amounts of exotic matter. We may have to pay attention to cosmological scale effects in some circumstances. But whether the fashionable cosmological model is explicitly Machian is really irrelevant to what we are up to. So we accept the physical version of Mach's principle – the assertion that inertial reaction forces are gravitational, and mass is just the total gravitational potential energy divided by the square of the speed of light – and ask: does the principle lead to any effects that we might be able to use to make starships and stargates? We address the answer to this question in the next chapter. Here, to sum up, we note that one way to suppress the confusion surrounding Mach's principle is to codify the principle in the form of a simple law or two. Imagine trying to do mechanics without Newton's laws, or electrodynamics without Maxwell's equations, or relativity without Einstein's laws. Therefore let's propose the adoption of the Mach-Einstein-Sciama laws of inertia:

First law: $\phi = c^2$ locally always; or, inertial reaction forces are due to the gravitational action of causally connected "matter", where matter is understood as everything that gravitates.

Second law: $m = E/\phi$, or the mass of an entity (isolated and at rest) is equal to its non-gravitational energy divided by the locally measured total gravitational potential.

A zeroth law might be added: Inertial reaction forces are instantaneous. But that is arguably belaboring the obvious. The first and second laws, in contradistinction, are not obvious. We will use these laws, and mostly ignore cosmology, to derive some interesting local effects that may make stargates possible. Cosmology will only come back into our consideration after those effects have been derived and some experimental work aimed at detecting them has been presented.

ADDENDA

Addendum #1: On the Origin of Inertia Article

It is convenient to begin by calculating the potential at a test-particle that is at rest in a universe containing no irregularities. Since our field equations have the same form as Maxwell's, we can use electrodynamic formulae to calculate the potential, and to bring out the analogy with electrodynamics we use a similar notation and terminology, but we emphasize that in this paper we shall be concerned with purely gravitational phenomena.

Retardation effects are taken to arise in the same way as in electrodynamics, so that the contribution of any region of the universe to the potential at a point P at time t is computed by ascribing to that region just the properties that are observed at P at time t .

We thus have for the scalar potential (8)

$$\Phi = - \int \frac{\rho}{r} dV. \quad (1)$$

We use the minus sign in (1) because inertial mass then turns out to be positive, but in fact either sign can be used (Section 4(vii)). The vector potential \mathbf{A} vanishes by symmetry.

We shall assume that matter receding with velocity greater than that of light makes no contribution to the potential, so that the integral in (1) is taken over the spherical volume of radius $c\tau$. An assumption of this sort is necessary since we have naïvely extrapolated the Hubble law without considering relativistic effects, and should give the correct order of magnitude. A relativistic treatment is given in II.

Since the density is supposed uniform, (1) gives

$$\Phi = -2\pi\rho c^2\tau^2. \quad (2)$$

Owing to our assumptions, the numerical factor 2π is only approximate.

We now calculate the potentials for the simple case when the particle moves relative to the smoothed-out universe with the small rectilinear velocity $-\mathbf{v}(t)$. In the rest-frame of the particle the universe moves rectilinearly with velocity $\mathbf{v}(t)$. Now at time t there will be observable at the particle, in addition to the Hubble effect, a Doppler shift corresponding to $\mathbf{v}(t)$ from all parts of the universe. Hence, in computing the potential in the rest-frame of the particle at time t , we must ascribe to every region of the universe the velocity that is observed at time t , that is, $\mathbf{v}(t) + \mathbf{r}/\tau$.

Neglecting terms of order v^2/c^2 , we have

$$\Phi = -2\pi\rho c^2\tau^2$$

as before. The vector potential no longer vanishes, but has the value

$$\mathbf{A} = - \int \frac{\mathbf{v}\rho}{r} dV. \quad (3)$$

Since \mathbf{v} is independent of \mathbf{r} , we can take it outside the integral. We then obtain

$$\mathbf{A} = \frac{\Phi}{c} \mathbf{v}(t).$$

Since the change of ρ with time is very small, the gravelectric part of the field is approximately

$$\begin{aligned} \mathbf{E} &= -\text{grad } \Phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \\ &= -\frac{\Phi}{c^2} \frac{\partial \mathbf{v}}{\partial t}, \end{aligned}$$

while the gravomagnetic field is

$$\mathbf{H} = \text{curl } \mathbf{A} = 0.$$

Addendum #2: Brans on Gravitational Energy Localization

PHYSICAL REVIEW

VOLUME 125, NUMBER 1

JANUARY 1, 1962

Mach's Principle and the Locally Measured Gravitational Constant in General Relativity*

CARL H. BRANS†

Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received August 14, 1961)

It has been conjectured that a "Mach's principle" might lead to a dependence of the local Newtonian gravitational constant, K , on universe structure, $K^{-1} \sim M/R$. Einstein and others have suggested that general relativity predicts such a result. A closer analysis, however, including the carrying out of the geodesic equations to second order, seems to indicate that this is not true and that the apparent "Mach's principle" terms involving total universe structure are really only coordinate effects. Further, the measure of gravitating mass obtained in a local, proper Newtonian gravitational experiment is compared in a coordinate-free way to an experimentally measurable inertial mass and found to be related to it in a way independent of the rest of the universe. A generalization of these results is given. It is based on the fact that in general relativity the only way the universe can influence experiments done in an electrically shielded laboratory is through the metric and that this can be "transformed away" to any degree of accuracy for a sufficiently small laboratory. Consequences of this are summarized in Dicke's "strong principle of equivalence." It is noted, however, that there are other statements which might be called "Mach's principles" which are satisfied in general relativity.

I. INTRODUCTION

THE principal idea which guided Einstein in formulating the general theory of relativity was the local equivalence of gravitational and inertial effects, that is, the equivalence of a uniform gravitational force field and a constant acceleration of the reference frame. Another idea relating gravity and inertia is Mach's principle. This is less precisely formulated but suggests that the inertial properties of a body are determined by the distribution of matter in the universe. Since the gravitational field interacts with *all* matter, one could hope to see the Mach principle relationship between inertial and distant matter described in terms of the gravitational field. To state this in a way independent of units, consider the ratio of the inertial mass of a body to its active gravitational mass.¹

In particular, let us see that this ratio might be in a static universe consisting only of a mass shell of radius R and inertial mass M together with a relatively small body of inertial mass m at its center. If we probe the gravitational field of m with a small test particle, we might expect from the Eötvös experiment that the acceleration of the test particle is independent of its mass. It certainly depends, however, on m and r and conceivably on M and R . The fact that the Newtonian theory of gravity is valid to a high degree of accuracy suggests that for $m \ll M$, $r \ll R$, the acceleration is

$$a = -[m/r^2 F(M, R)], \quad (1.1)$$

where F is a function of dimensions mass over length (velocity of light $c=1$). Dimensional analysis then suggests

$$F = AM/R, \quad (1.2)$$

where A is a constant dimensionless number. For a more general type of universe with masses m_a at distances r_a from some point x , this might be extended to

$$F(x) = A \sum_a m_a / r_a. \quad (1.3)$$

Until recently, experimental determinations of F from (1.1) were possible only on the earth. The value found is not inconsistent with (1.3), a positive value of A in the neighborhood of 10^0 or 10^1 , and present astronomical knowledge of m_a and r_a . It is clear that in a uniform universe, $m_a \sim r_a^2$, so that the dominant contribution to the sum on the right side of (1.3) comes from distant matter and the resulting $F(x)$ is fairly constant in space and time. This also is consistent with present observations.

A comparison of (1.1) with the standard classical Newtonian theory of gravity shows that F^{-1} plays the role of Newton's "universal gravitational constant." However, if (1.3) is true, this number is not a universal constant but depends on the distribution of mass in the universe about the point where it is measured. To investigate possible resulting changes in value of this number, it is convenient to introduce a standard value and refer variations to it. Specifically, let $K_0/8\pi$ be defined as the presently observed terrestrial value of $F(x)^{-1}$. $K_0/8\pi$ is thus a constant number of dimensions length over mass. Then rewrite (1.1) as

$$a = -\frac{K_0}{8\pi r^2} \left(\frac{8\pi m}{K_0 F} \right). \quad (1.4)$$

This equation is identical with Newton's if the quantity in parentheses, $m_g \equiv 8\pi m / K_0 F$, is taken to be the active gravitational mass⁴ associated with m . Notice that by definition of K_0 , this gives $m_g = m$ at the present time on earth. However, if (1.3) is true, a Cavendish-type experiment interpreted in the context of a Newtonian theory with fixed gravitational constant $K_0/8\pi$ would give a measurement of active gravitational mass m_g

* Based on part of a multilithed Ph.D. thesis submitted to Princeton University.

† National Science Foundation Predoctoral Fellow, 1957-1960. Now at Loyola University, New Orleans, Louisiana.

¹ H. Bondi, *Revs. Modern Phys.* 29, 423 (1957).

MACH'S PRINCIPLE IN GENERAL RELATIVITY

yielding a ratio

$$\frac{m}{m_0} = A \sum_a \frac{K_0 m_a}{8\pi r_a}, \tag{1.5}$$

which would not necessarily always be unity.

Einstein² claims to find such a result in general relativity. In order to study this problem, consider the creation of relatively small masses m_a' , at distances r_a' , from the present standard laboratory in which, prior to the creation of m_a' , $m/m_0=1$ by definition of K_0 . With m_a' present, however, (1.5) then yields

$$\frac{m}{m_0} = 1 + A \sum_{\text{"new matter"}} \frac{K_0 m_a'}{8\pi r_a'}. \tag{1.6}$$

If it is assumed that each $K_0 m_a'/8\pi r_a'$ is small compared to unity, the weak-field equations might be used to check (1.6). Einstein does this and arrives at

$$\left(1 + \sum_{\text{"new matter"}} \frac{K_0 m_a}{8\pi r_a}\right) a \cong -\frac{K_0 m}{8\pi r^2}. \tag{1.7}$$

Thus

$$\frac{m}{m_0} = 1 + \sum_{\text{"new matter"}} \frac{K_0 m_a}{8\pi r_a}, \tag{1.8}$$

which is identical with (1.6) if $A=1$. Einstein argued from this that since some matter contributes to the ratio, m/m_0 , all the universe probably does (Sec. II). There has been some discussion³ of what the numerical coefficient A of the sum in the right side of (1.8) should be, and indeed the first approximation procedure seems inadequate to resolve this. Consequently, the equations of motion through second order will be applied to this problem in Sec. II.

This result (1.8), or its corrected form (2.11), is clearly coordinate dependent, however. Hence the relationship between its numerical description of the path of a particle and the actually observed path is not defined without further analysis. The usual interpretation of general relativity is based on the identification of the *invariant* theoretical measure of an interval, proper time, with time experimentally measured in some fundamental way, e.g., on an atomic clock. An invariant measure of distance and thus acceleration can be obtained from this by setting the velocity of light equal to one. When this is done, the invariant description of the path of a test particle relative to a central mass is found to be approximately Newtonian with coefficients independent of the rest of the universe. (See Sec. III.)

However, the number m appearing in the left side of (1.5) has not yet been related to an experimentally measured inertial mass. To remedy this, a description of a process for invariantly studying the acceleration of charged bodies in a known electric field is given. The

resultant ratio of "force" to acceleration is defined as the inertial mass. For a simple theory of matter m_{inert} is found to be just the m appearing in (1.5) (Sec. IV). This procedure assumes given standards of charge and time interval.

The independence of the relationship between the two numbers, m_0 and m_{inert} , from the rest of the universe is more generally true than the above special case might indicate. In fact, assume that the space in the neighborhood of an electrically shielded laboratory is sufficiently flat that in a certain coordinate system the differences between the metric components and those of the Minkowskian, together with the first two derivatives of these differences, are negligible over the laboratory. Then, according to general relativity, if small masses, charged or uncharged, are introduced into the laboratory, the description of their motions and interactions in this coordinate system is independent of the rest of the universe. This is due to the fact that once the laboratory is shielded, the only way the rest of the universe could influence it, according to general relativity, is through the metric. If this is sensibly flat within, its influence can be transformed away by a coordinate transformation, thus eliminating any effects from the rest of the universe. This is Dicke's "strong principle of equivalence."⁴ (See Sec. V.)

There are, however, other statements which might be considered Mach's principles. These are based on the fact that in general relativity gravitational and inertial forces have the same formal origin. (See Sec. V.)

II. EINSTEIN'S RESULTS

Gravity and general relativity being largely concerned with the interaction between masses as masses, Einstein was naturally interested in whether or not Mach's principle as discussed in Sec. I above was satisfied in general relativity. Specifically, is the attraction and resultant relative motion of two gravitating bodies influenced by the rest of the universe?

Einstein investigated this in the weak-field approximation.² The metric he found to represent the gravitational field due to a distribution of small masses corresponding to a "density" σ and having small velocities, dx^i/ds , can be written as

$$\begin{aligned} g_{00} &= 1 - \frac{K}{4\pi} \int \frac{\sigma dV}{r}, \\ g_{0i} &= \frac{K}{2\pi} \int \frac{\sigma(dx^i/ds)}{r} dV, \\ g_{ij} &= -\delta_{ij} \left(1 + \frac{K}{4\pi} \int \frac{\sigma dV}{r} \right), \end{aligned} \tag{2.1}$$

² A. Einstein, *The Meaning of Relativity* (Princeton University Press, Princeton, New Jersey, 1955), 5th ed., pp. 99-108.

³ W. Davidson, *Monthly Notices Roy. Astron. Soc.* 117, 212 (1957).

⁴ R. H. Dicke, *Science* 129, 621 (1959). See also *Revs. Modern Phys.* 29, 355 (1957); *J. Wash. Acad. Sci.* 48, 213 (1958); *Am. J. Phys.* 28, 344 (1960).

CARL H. BRANS

on replacing Einstein's imaginary time x^4 by the real $x^0 = -ix^4$. Here K is just the constant introduced in the Einstein field equations and thus not yet related to K_0 or other observed numbers. Equation (2.1) is correct only to first order in $K \int \sigma dV/r$, and dx^i/ds . The geodesic equation for a test particle in this field becomes

$$\frac{d}{dx^0}[(1+\bar{\sigma})\mathbf{v}] = \nabla\bar{\sigma} + \frac{\partial\mathbf{A}}{\partial x^0} + (\nabla \times \mathbf{A}) \times \mathbf{v}, \quad (2.2)$$

where

$$\begin{aligned} \mathbf{v} &\equiv d\mathbf{x}/ds, \\ \bar{\sigma} &\equiv \frac{K}{8\pi} \int \frac{\sigma dV}{r}, \\ \mathbf{A} &\equiv \frac{K}{2\pi} \int \frac{\sigma \mathbf{v}}{r} dV. \end{aligned} \quad (2.3)$$

For simplicity, consider the application of these results to the case of the motion of a test particle near a small mass m at rest at the origin, all inside a static, spherical shell of mass M_s and radius R_s ,⁵ (2.2) now becomes

$$\frac{d}{dx^0} \left[\left(1 + \frac{KM_s}{8\pi R_s} + \frac{Km}{8\pi r} \right) v^i \right] = \frac{Km}{8\pi} \frac{\partial}{\partial x^i} \left(\frac{1}{r} \right). \quad (2.4)$$

Thus, $(1 + KM_s/8\pi R_s + Km/8\pi r)$ times the coordinate acceleration of the test particle is just the Newtonian term, to this approximation. Einstein interpreted this by saying that the "inert mass is proportional to $1 + \bar{\sigma}$,"² or in (2.4) to $1 + (K/8\pi)(M_s/R_s + m/r)$. However, an equivalent statement, more convenient for this discussion and in keeping with that of Sec. I, can be made. Specifically, dividing (2.4) by $[1 + (K/8\pi) \times (M_s/R_s + m/r)]$ gives, for v_i instantaneously zero,

$$\frac{d}{dx^0} v^i = \frac{Km}{8\pi [1 + (K/8\pi)(M_s/R_s + m/r)]} \frac{\partial}{\partial x^i} \left(\frac{1}{r} \right). \quad (2.5)$$

This, in keeping with Einstein's interpretation above, would suggest that the locally measured Newtonian active gravitational mass of m is

$$m_a = m / [1 + (K/8\pi)(M_s/R_s + m/r)], \quad (2.6)$$

or that the effective, locally measured Newtonian gravitational constant is

$$K_E = K / [1 + (K/8\pi)(M_s/R_s + m/r)]. \quad (2.7)$$

If this is true, a comparison of (2.6) with (1.5) would show that a Mach's principle in the sense of Sec. I would be satisfied in general relativity, since the number K_E in (2.7) measuring the attraction of m for

⁵ This example, while admittedly rather specialized, is sufficient to illustrate the ideas under consideration. It should also be noted that here $KM_s/R_s \ll 1$ so that this does not correspond to the total "universe mass shell" discussed in Sec. I and for which $KM/R \sim 1$.

test particles would depend on the mass distribution M_s/R_s in the rest of the universe. To clarify the relation of (2.6) and (2.7) to the discussion in Sec. I, it is necessary to consider M_s and m as small additions to a background universe [i.e., as the m_s ' were in the discussion preceding (1.6) above]. For the background universe assume that K has been chosen equal to K_0 . Thus, (2.6) will coincide with (1.6) if $A = 1$ in the latter.

V. SUMMARY AND GENERALIZATION

This section will be mainly concerned with investigating some of the consequences of the fact that in general relativity the entire gravitational interaction between masses is carried by the metric tensor which can be "transformed away" to any desired degree of accuracy over a sufficiently small neighborhood of any point. This fact leads naturally to the following definition relating a standard physical laboratory to a mathematical "coordinate patch." A locally almost Minkowskian coordinate system is one in which test particles of any velocity experience no observable acceleration when there is no matter or radiation present in the laboratory. The description of experiments done in a standard physical laboratory is assumed to correspond to the mathematical description given by such a coordinate system.

Using this definition, Dicke's⁴ strong principle of equivalence can be defined as the assertion that as far as inertial and gravitational effects are concerned, the numerical content of experiments described in a locally almost Minkowskian coordinate system is independent of any characteristics of the mass distribution in the rest of the universe. It is important to realize that this is a definite extension of such results of the Eötvös experiment as generalized in the weak principle, i.e., the assertion that the acceleration of a test particle instantaneously at rest relative to a small gravitating body is independent of the mass of the test particle in the limit as this mass goes to zero. In other words, the Eötvös experiment suggests that the acceleration effects of an external gravitating body on a sufficiently

small laboratory can be at least approximately eliminated by allowing the laboratory to fall "freely" since it seems to imply that all parts of the laboratory would fall with very little, if any, relative acceleration. However, it contains nothing to suggest that the only effect of the gravitating body on the laboratory is accelerative, which is the basis for the strong principle.

A sketch of an argument generalizing the results of Sec. IV and suggesting the validity of a strong principle in general relativity follows.

Consider a region having space-time dimensions, in arbitrary but fixed units, bounded by a number ϵ . This region is to represent the space contained in a laboratory in which standard experiments are to be performed. Let the matter tensor in the laboratory be represented by λT_L (here and in the following, to avoid unnecessary clutter, tensor indices will be suppressed when no confusion will arise), where λ is a positive number. Further, let the matter tensor for the rest of the universe be T_U and assume that $T_U=0$ within the laboratory, while $T_L=0$ outside it. The total matter tensor is thus $T_U+\lambda T_L$ everywhere. The purpose of the following discussion is then to show that under certain conditions the influence of the "rest of the universe" on real, proper experiments done in such a laboratory can be made arbitrarily small by making ϵ sufficiently small. The crux of the argument is the fact that the observable outcome of such experiments cannot depend on the purely mathematical choice of coordinate systems in which the calculations are performed.

To this end, let ρ (again suppressing indices) stand for all the matter variables other than the metric, ρ_L referring to matter in the laboratory, and ρ_U to all other matter. Thus, λT_L is a function of ρ_L and T_U is a function of ρ_U . Assume the variables satisfy "equations of motion"

$$f(\rho, g, g')=0, \quad (5.1)$$

where g' stands for all first derivatives of g . Further, let the metric, g , be written as the sum of two parts ${}^0g+\gamma(\lambda)$, with 0g independent of λ and where $\lim_{\lambda \rightarrow 0} \gamma(\lambda) = 0$ as $\lambda \rightarrow 0$.

Let ${}^0\rho$ represent the functional form of ρ when $\lambda=0$. Hence, when there is no matter within the laboratory, $\lambda=0$, and ${}^0\rho$ and 0g satisfy

$$f({}^0\rho, {}^0g, {}^0g')=0, \quad (5.2)$$

$$S({}^0g)-T_U({}^0\rho)=0 \quad (S^{\alpha\beta} \equiv R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R). \quad (5.3)$$

If ρ_{null} represents the form of ρ corresponding to the vacuum and η is the Minkowski metric, then it will be assumed that

$$f(\rho_{\text{null}}, \eta, 0)=0, \quad (5.4)$$

$$T_U(\rho_{\text{null}})=0. \quad (5.5)$$

The two most important assumptions will now be made. Within the space of the laboratory it is assumed

that (1) ${}^0\rho=\rho_{\text{null}}$, and (2) the differences between ${}^0g+\gamma(\lambda)$ and η together with the first two derivatives of ${}^0g+\gamma(\lambda)$ go continuously to zero with ϵ and λ . The first assumption is simply that when $\lambda=0$ there are really no matter or fields within the laboratory. In other words, it ensures that when the tensor for matter in the laboratory, λT_L , is zero, the matter variables, ρ_L , actually correspond to the vacuum. This assumption is probably unnecessary for the ordinary descriptions of matter. The second assumption may seem strong in its requirement on the second derivatives of the metric. However, it will be used in the argument following Eq. (5.13).

Similarly, let ${}^0\rho_L$ and $\eta+{}^0\gamma$ be the matter variables and metric describing the situation inside the laboratory in the absence of any matter outside, i.e., when $\rho_U = \rho_{\text{null}}$. Thus, by definition,

$$f({}^0\rho_L, \eta+{}^0\gamma, {}^0\gamma')=0, \quad (5.6)$$

$$S(\eta+{}^0\gamma)-\lambda T_L({}^0\rho_L)=0. \quad (5.7)$$

Finally, the full field equations can be written

$$f(\rho, {}^0g+\gamma, {}^0g'+\gamma')=0, \quad (5.8)$$

$$S({}^0g+\gamma)-\lambda T_L(\rho_L)-T_U(\rho_U)=0. \quad (5.9)$$

In particular, within the laboratory,

$$f(\rho_L, {}^0g+\gamma, {}^0g'+\gamma')=0, \quad (5.10)$$

$$S({}^0g+\gamma)-\lambda T_L(\rho_L)=0. \quad (5.11)$$

However, by assumption, within the laboratory 0g differs from η , and its first two derivatives from zero, only by numbers which go to zero as $\epsilon \rightarrow 0$. Thus (5.10) and (5.11) can be rewritten as

$$f(\rho_L, \eta+\gamma, \gamma')=H, \quad (5.12)$$

$$S(\eta+\gamma)-\lambda T_L(\rho_L)=E, \quad (5.13)$$

where $H \rightarrow 0$ as $\epsilon \rightarrow 0$ and $E \rightarrow 0$ as $\epsilon \rightarrow 0$. Notice that since S depends on the second derivatives of ${}^0g+\gamma$, it is sufficient that these vanish as $\epsilon \rightarrow 0$ for $E \rightarrow 0$ as $\epsilon \rightarrow 0$. Actually, this condition may not also be necessary, but this point is irrelevant to the main argument.

The final result is thus that the variables, ρ_L and $\eta+\gamma$, satisfy, within the laboratory, Eqs. (5.12) and (5.13) which differ from those, (5.6) and (5.7), satisfied by the corresponding variables in the absence of matter in the rest of the universe only by functions H and E which can be made arbitrarily small by making ϵ sufficiently small.

Thus, it seems reasonable to expect that for each λ , the solutions with matter in the rest of the universe, ${}^0g+\gamma$ and ρ_L , and those with matter only in the laboratory, $\eta+{}^0\gamma$ and ${}^0\rho_L$, can be brought arbitrarily close together by making the laboratory sufficiently small. Further, the outcome of proper, local experiments done in such a laboratory can depend only on the behavior of the metric and matter variables within it. Thus, the

CARL H. BRANS

results of such experiments can be made as nearly independent of the matter in the rest of the universe as desired by making ϵ sufficiently small.

Of course, the definition of quantities to be measured¹¹ and local laws to be tested within the laboratory may require $\lambda \rightarrow 0$. It might then be thought that for λ small enough the effects of the matter in the rest of the universe would become comparable to those of matter within the laboratory, vitiating the above argument. To prevent this, a lower limit for λ is demanded. This limit could be determined by the lower bound of available experimental accuracy for the measurements requiring $\lambda \rightarrow 0$. That is, values of λ below this limit would not produce observable differences in measurements. For this fixed λ , ϵ can then be determined as above.

There are, however, other statements which might possibly be called "Mach's principles" which are valid in general relativity. For example, inertial and gravitational forces have a common formal origin in general relativity. Specifically, for a test particle of mass m and velocity w^β ,

$$F^\mu \equiv -m\Gamma_{\alpha\beta}{}^\mu w^\alpha w^\beta \tag{5.14}$$

might be identified with the gravitational force acting on m . On the other hand, this quantity transforms just as an inertial force should, i.e., in going to a relatively accelerated system, the acceleration enters F^μ linearly. For example, in a coordinate system rotating relatively to a Lorentz system in a flat space, F^μ as defined in (5.14) contains the centrifugal and Coriolis forces experienced by particles in this rotating system.

¹¹ For example, inertial mass. See Eq. (4.23) and the discussion following it.

Thus, F^μ might also be identified with "inertial force" acting on m . Inertial coordinate systems would then be those in which F^μ vanishes or equivalently, those in which "free" uncharged test particles are unaccelerated. This coincides with the definition of locally almost Minkowskian coordinate systems above. Another way of saying this is that the locally almost Minkowskian or inertial coordinate systems are those in which the total gravitational force vanishes.

If suitable boundary conditions could then be exhibited for a general type of universe, the Einstein equation would predict the over-all state of motion of inertial frames relative to the total mass distribution in the universe. This statement alone has been mentioned as a "Mach's principle."¹² However, once it is required that fundamental, standard experiments be done within such frames, the rest of the universe cannot, in general relativity, influence their results.

Another paper¹³ will discuss modifications of general relativity violating the strong principle of equivalence by the introduction of a variable gravitational "constant" determined through field equations by the mass distribution in the universe.

ACKNOWLEDGMENTS

The author is grateful to R. H. Dicke and C. W. Misner for many helpful discussions of these problems. He also wishes to thank B. Block, W. Morgan, and J. Peebles for their many helpful comments.

¹² For a general discussion see F. A. E. Pirani, *Helv. Phys. Acta*, Suppl. IV, 198 (1956). Actually, Pirani's "Mach's principle" is stronger than that mentioned above. His requires that inertial systems be nonrotating relative to some average mass density in the universe.

¹³ C. H. Brans and R. H. Dicke, *Phys. Rev.* 124, 925 (1961).

Reprinted excerpt with permission from C.H. Brans, "Mach's principle and the Locally Measured Gravitational Constant in General Relativity," *Physical Review*, vol. 125, pp. 388-396 (1962). Copyright 1962 by the American Physical Society.

Addenda #3: Excerpt from Nordtvedt

International Journal of Theoretical Physics, Vol. 27, No. 11, 1988

Existence of the Gravitomagnetic Interaction

Ken Nordtvedt¹

Received September 25, 1987

The point of view expressed in the literature that gravitomagnetism has not yet been observed or measured is not entirely correct. Observations of gravitational phenomena are reviewed in which the gravitomagnetic interaction—a post-Newtonian gravitational force between moving matter—has participated and which has been measured to 1 part in 1000. Gravitomagnetism is shown to be ubiquitous in gravitational phenomena and is a necessary ingredient in the equations of motion, without which the most basic gravitational dynamical effects (including Newtonian gravity) could not be consistently calculated by different inertial observers.

1. INTRODUCTION

In the overview *Physics Through the 1960s*, the National Academy of Sciences (1986) review of opportunities for experimental tests of general relativity, they declare that “At present there is no experimental evidence arguing for or against the existence of the gravitomagnetic effects predicted by general relativity. This fundamental part of the theory remains untested.” Similar points of view have been expressed elsewhere in promotion of various experiments designed to “see” gravitomagnetism.

In this paper I make two points on this issue, which together lead to a position contrary to the viewpoint summarized by the above statement.

1. The gravitomagnetic interaction is a consequence of the gravitational vector potential. This vector potential plays a crucial, unavoidable role in gravitation; without the gravitational vector potential the simplest gravitational phenomena—the Newtonian-order Keplerian orbit and the deflection of light by a central body—cannot be consistently calculated in two or more inertial frames of observation. Gravitation without the vector potential is an incomplete, ambiguous theory in the most fundamental sense.

¹Physics Department, Montana State University, Bozeman, Montana 59717.

5. DRAGGING OF INERTIAL FRAMES AND MACH'S IDEAS

What seems to have especially caught the interest of physicists in searching for the spin-spin interaction in gravity is that this would seem to be a manifestation of ideas of Mach, who a century ago believed that inertia was caused, in some sense, by the universe's matter distribution. Lense and Thirring later showed that, indeed, in general relativity rotating matter would drag the inertial frame around at a slow rate which fell off with distance from the rotating matter,

$$\Omega = \frac{G}{c^3} \left(\frac{\mathbf{J} - 3\mathbf{J} \cdot \hat{\mathbf{r}}\hat{\mathbf{r}}}{r^3} \right) \tag{16}$$

\mathbf{J} is the angular momentum of the spinning body and \mathbf{r} is the distance to the point of space in question, $\Omega(\mathbf{r})$ is the rotation rate and rotation axis for the inertial space at that point of space which is induced by the spinning source. Equation (16) follows from (12) with choice of PPN coefficients appropriate to general relativity, and the identification

$$\Omega = -\frac{c}{2} \nabla \times \mathbf{h}$$

Looking at the general case, one can ask what is the complete effect of the gravitational vector potential in dragging inertial frames? This question can be addressed by calculating the contribution of \mathbf{h} in establishing the geodesic coordinate frames (inertial frames). The general formula

$$[x^\gamma - x_{(0)}^\gamma]' = [x^\gamma - x_{(0)}^\gamma] + \frac{1}{2} \Gamma_{\alpha\beta}^\gamma [x^\alpha - x_{(0)}^\alpha][x^\beta - x_{(0)}^\beta] \tag{17}$$

in which $\Gamma_{\alpha\beta}^\gamma$ are the Christoffel symbols produced from first derivatives of the gravitational metric field, gives the transformation from original space-time coordinates x^γ to inertial (geodesic) coordinates x'^γ in the vicinity of any chosen space-time point $x^\gamma(0)$. Examining solely the vector potential (g_{0i}) contribution to (17) yields

$$[\mathbf{r} - \mathbf{r}_{(0)}]' = [\mathbf{r} - \mathbf{r}_{(0)}] - c \left[\frac{1}{2} \frac{\partial \mathbf{h}}{\partial t} (t - t_0)^2 + \left(\frac{\nabla \times \mathbf{h}}{2} \right) \times (\mathbf{r} - \mathbf{r}_{(0)}) (t - t_0) \right] \tag{18}$$

The gravitational vector potential produces in this general case a "dragging" of inertial space at each locality with both an acceleration of the inertial frame at rate

$$\mathbf{a}(\mathbf{r}, t) = -c \partial \mathbf{h} / \partial t \tag{19a}$$

and a rotation of the inertial frame at angular rate and axis

$$\Omega(\mathbf{r}, t) = -\frac{1}{2} c \nabla \times \mathbf{h} \tag{19b}$$

If we return to the problem of light deflection by a body moving at speed w and employ the vector potential given by (7), we find that (19a) gives no contribution to the light ray deflection; however, (19b) produces a rotational dragging of inertial frames at a rate

$$\Omega(r, t) = (1 + \gamma) \frac{GMDw}{c^2} \frac{1}{|\mathbf{r} - \mathbf{w}t|^3}$$

and in a counterclockwise sense. The time integral of this rotation rate over the entire trajectory of the light ray produces the total deflection or rotation angle

$$\delta\theta = -\frac{2w}{c} \theta_0$$

which is what is needed to obtain agreement with (5) as discussed in Section 2.

The periastron precession of the binary pulsar orbit discussed previously received contributions of inertial frame dragging from both (19a) and (19b). The situation can be viewed this way; part of the motion of the two bodies in the binary pulsar results from the “Coriolis” acceleration that each body experiences because the motion of the other body is producing rotational dragging of the inertial frame at the locality of each body in question.

Finally, the accelerated celestial body mentioned previously drags the inertial frames through (19a), with the resulting acceleration of inertial space being

$$\delta\mathbf{a}(\mathbf{r}, t) = -\left(2\dot{\gamma} + 2\gamma + \frac{\alpha_1}{2}\right) \frac{U(\mathbf{r}, t)}{c^2} \mathbf{a}$$

in which $U(\mathbf{r})$ is the Newtonian potential function of that body's mass distribution and \mathbf{a} is the body's acceleration.

6. CONCLUSION

The gravitomagnetic interaction—the post-Newtonian gravitational interaction between moving masses—has been observed and measured in a number of different phenomena. The strength of this interaction is now known to an accuracy of 1 part in 1000. The gravitomagnetic interaction is also required in order to have a complete and consistent theory of gravity at all: even static source gravitational effects when viewed in another inertial frame require the gravitomagnetic interaction in order for basic consistency of a theory's equations of motion. Just as in electromagnetic theory, there is no absolute separation of “electric” and “magnetic” effects; such a division is inertial frame dependent.



<http://www.springer.com/978-1-4614-5622-3>

Making Starships and Stargates

The Science of Interstellar Transport and Absurdly Benign
Wormholes

Woodward, J.F.

2013, XXVI, 279 p. 92 illus., 85 illus. in color., Softcover

ISBN: 978-1-4614-5622-3