Foundations of Physics, Vol. 29, No. 6, 1999

What is the Cause of Inertia?

James F. Woodward[1] **and Thomas Mahood**[2]

Received June 1, 1998; revised May 3, 1999

The question of the cause of inertial reaction forces and the validity of "Mach's principle" are investigated. A recent claim that the cause of inertial reaction forces can be attributed to an interaction of the electrical charge of elementary particles with the hypothetical quantum mechanical "zero-point" fluctuation electromagnetic field is shown to be untenable. It fails to correspond to reality because the coupling of electric charge to the electromagnetic field cannot be made to mimic plausibly the universal coupling of gravity and inertia to the stress–energy–momentum (i.e., matter) tensor. The gravitational explanation of the origin of inertial forces is then briefly laid out, and various important features of it explored in the last half-century are addressed.

1. INTRODUCTION

The cause of inertia has been an issue in theoretical physics since the time of Newton. Bishop George Berkeley, in particular, pointed out in Newton's day that in the absence of other matter in a universe, all discussion of the motion of a single body was meaningless. And since motion (i.e., momentum) is the measure of inertia, by inference objects in otherwise empty universes should have no inertia. Ernst Mach repeated Berkeley's criticism toward the end of the 19th century, adding an insinuendo that some physical agent should be held accountable for forces of inertial reaction. The obvious candidate for the agent of inertial reaction forces is gravity because of its universal coupling to mass, or so it appeared to Einstein and some of his contemporaries. Newtonian gravity cannot be made to account for inertial forces however. But Einstein hoped to be able to encompass, as he named it, "Mach's principle" in his theory of gravity, general relativity

---

[1] Departments of History and Physics, California State University, Fullerton, California 92634.

[2] Department of Physics, California State University, Fullerton, California 92634.

theory (GRT). [Einstein's efforts in this direction are nicely recounted by Julian Barbour(1) and Carl Hoefer(17).]

Einstein's attempts to make GRT Machian were not a complete success. But his insight has fired the imagination of many of the cleverest thinkers of the past century. After all, Mach's principle—the gravitational induction and relativity of inertia—is the most extreme statement of the *principle* of relativity. The idea that inertial forces should be caused by the rest of the matter in the universe, once apprehended, has an ineluctable intuitive obviousness about it that is hard to resist. This has led to continuing debate among experts as to the real meaning of the principle, and arguments over how it is to be formally realized in physical theory. The lack of consensual accord among experts on the meaning of Mach's principle has left room for some, Rueda and Haisch,(24a, b) for example, to claim that inertia should not be understood in the context of gravity theory; that inertia is in fact is a consequence of electromagnetic interactions of accelerating matter with a quantum mechanical vacuum fluctuation electromagnetic field.

We would be the last to suggest that all issues relating to the origin of inertia are really settled. But we think it important to point out that many things about the origin of inertia have, indeed, been sorted out. The appearance of discord among experts is, at least in part, a consequence of the fact that the issues involved have been resolved piecemeal over the past half century. Our aim here is to pull those matters now known to be correct together and lay them out systematically so that those inexpert, but interested, in the origin of inertia can find them presented coherently in one place. Some (but not all) of this paper, therefore, is a review of work already done.

While the origin of inertia remains a foundational issue in the physical sciences, as it has been for the past few centuries, recently a sense of "something magical" being in the air has gotten abroad. The idea that means of manipulating inertia for the purposes of rapid spacetime transport may be imminent is being discussed in some quarters. [In this connection, see Haisch and Rueda(15) and Haisch *et al.*(16b)] That means that serious resources may soon be injected into an otherwise dry (and inexpensive) academic debate. On the practical side, we hope to make clear what's plausible, and what's not, in the matter of inertia manipulation. In particular, we show that schemes based on vacuum fluctuations and manipulations of putative vacuum energy are doomed to failure for simple and compelling reasons. Next we summarize the chief things that have been found out about the origin of inertia and identify some of the strange and contentious aspects of this matter. We then present simple discussions of some of the chief issues related to the origin of inertia.

2. VACUUM FLUCTUATIONS AND THE ORIGIN OF INERTIA

Accompanied by extraordinary prepublication press attention (e.g., Ref. 18), early in 1994 B. Haisch, A. Rueda, and H. Puthoff(16a) (HRP) published "Inertia as a Zero-Point-Field Lorentz Force," in *Physical Review A*. As motivation for their proposal, they offered

> Since the time of Newton there has been only one noteworthy attempt to associate an underlying origin of inertia of an object with something external to that object: Mach's principle .... Mach's principle has remained, however, a philosophical statement rather than a testable scientific proposition .... and while special and general relativity both involve the inertial properties of matter, they provide no deeper insight into an origin of inertia than Newton's definition of inertia as a fundamental property of matter.

This statement is not correct. It has long been known, at least to experts in gravitational physics, that inertial reaction forces (and thus inertia) are accounted for as gravitational effects in general relativity for isotropic universes like ours, as we elaborate below.

In their recent work Rueda and Haisch(24a, b) (RH) scale back somewhat the reach of their claims and take cognizance of at least some of the work of others. They, however, identify only a small part of the work that has been done and suggest that some of it has been disproved by experiment. Were these assertions true, of course, gravity in the guise of general relativity theory would be wrong and, therefore, unable to account for inertia correctly in any event. These claims, too, we examine below. First we sketch the crucial physical content of HRP's and RH's scheme and show why it cannot plausibly account for inertial reaction forces.

2.1. A Sketch of HRP's and RH's Claims

The formal starting point for HRP's conjecture on the origin of inertia was earlier work by Boyer(4) in which he examined the response of a dipole oscillator to an acceleration in the presence of a thermal bath of photons—the electromagnetic zero-point field (ZPF) that putatively permeates all spacetime. The equation of motion accordingly contains a term that gives the force exerted by the ZPF on the dipole. Actually, to get the effect in question we do not need to consider a dipole. We need only examine the response of a single elementary charged particle. The justification for this simplification, as HRP remark (Ref. 16, p. 680), is that at the chief, very high interaction energies (frequencies) expected in ZPF processes, even bound charged particles can be regarded as "asymptotically free." If the particle is subjected to an external force that produces a steady acceleration, the particle's equation of motion will include the usual inertial term ($m\ddot{\mathbf{x}}$), a radiation reaction term—from the so-called Lorentz–Dirac equation $[(2e^2/3c^3)\{\dddot{\mathbf{x}} - (\ddot{\mathbf{x}}^2/c^2)\,\dot{\mathbf{x}}\}]$—and a term that accounts for the interaction of the accelerating charge and the putative ZPF $[e\mathbf{E}_{\mathrm{zpf}}]$. For simplicity we consider here motion in the x direction. e is the electric charge of the particle, $\mathbf{E}$ the electric field strength, and overdots represent differentiation with respect to proper time. For the nonrelativistic velocities we are concerned with the distinction between proper and coordinate time can be ignored.

The crux of this whole business lies in HRP's claim that the interaction with the ZPF, eE_{zpf}, that is, can account for the entire inertial reaction force the accelerating particle experiences. To make their argument, they invoke an approximation employed by Einstein and Hopf in calculating the force of radiation on an electric charge. In this approximation one posits that the response of the charge is dominantly to the electric field, which induces motion in the direction of the electric field (in addition to the motion being induced by the external force). The motion of the charge induced by the electric field makes the charge a current, which then interacts with the magnetic part of the radiation field to produce a magnetic force on the charged particle. It is the magnetic part of the action of the radiation field on the charge that allegedly is responsible for the inertial reaction force that the agent producing the external force feels. There are important subtleties in this argument that are required to make it work. Some formalism is needed to see them. So we put this into an admittedly crude, but nonetheless accurate formalism.

HRP engage in elaborate "stochastic" averaging over random phases, integrating over the frequency spectrum of the putative ZPF, and transforming from accelerating frames of reference to the laboratory frame of reference. While all of these procedures are necessary to their purpose, they are not the crucial physics involved in their conjecture. That physics lies in the way that electric charge couples to the electromagnetic radiation field via the Lorentz force. So in the formalism we present here we accept and suppress all of the elaborations just mentioned by absorbing them into a single coefficient in our equations.

Consider an elementary electric charge accelerated by an external force. If the electromagnetic ZPF actually exists, then this charge will experience a flux of electromagnetic radiation directed opposite to the acceleration, and the electric part of the radiation will cause the charge to move in its direction. Formally, the effective electric field response is

$$e\mathbf{E}_{\mathrm{zpf}} \approx k\boldsymbol{v} \tag{2.1}$$

where k is a constant of proportionality with dimensions force per unit velocity. It is into k that all of the suppressed technical elaboration is placed. The velocity v in Eq. (2.1) should be understood to be the typical velocity (an average of some suitable sort) produced by an oscillatory $\mathbf{E}_{\mathrm{zpf}}$ during a half-cycle for which the sign of v doesn't change. [This and the unusual form of Eq. (2.1) are consequences of the suppressed averaging procedures.] Since the response to the electric part of the radiation field is periodic, it time-averages to zero. But the magnetic part of the Lorentz force acting on the particle that HRP claim to be the source of inertial reaction forces may not, for in this approximation,

$$\mathbf{F}_{\mathrm{zpf}} = (e/c)(v \times \mathbf{B}_{\mathrm{zpf}}) \tag{2.2}$$

Up to slightly different notation conventions, Eq. (2.2) is just HRP's Eq. (18). Now substituting for v from Eq. (2.1), we get

$$\mathbf{F}_{\mathrm{zpf}} \approx (e^2/kc)(\mathbf{E}_{\mathrm{zpf}} \times \mathbf{B}_{\mathrm{zpf}}) \tag{2.3}$$

As long as k is finite and nonzero, $\mathbf{F}_{\mathrm{zpf}}$ may be finite and nonzero too. Inasmuch as $\mathbf{E}_{\mathrm{zpf}} \times \mathbf{B}_{\mathrm{zpf}}$ is the Poynting vector—that is, energy flux density—of the putative ZPF seen by the accelerating charge, should it have a nonzero time average, a nonzero time-averaged momentum flux density will also be present. And the accelerating charged particle should experience a force the direction opposite to its acceleration. It has been known since the 1970s that an accelerating observer will detect a thermal bath of evidently real photons (the Davies–Unruh effect). But the Poynting vector in a thermal bath, on average at least, vanishes. HRP (and RH), however, find that the time-averaged ZPF Poynting vector does not vanish in an accelerating frame of reference, so $\mathbf{F}_{\mathrm{zpf}}$ does not vanish. Because of the properties of relativistic transformations from accelerating frames of reference to ones at rest [*viz*., HRP's Eqs. (9a) and (9b)], the expression they find for $\mathbf{F}_{\mathrm{zpf}}$ ends up with a factor of the acceleration, $\mathbf{a}$, in it [*viz*., HRP's Eqs. (91) to (94)].

A few comments on Eq. (2.1) are in order. Hidden in this innocuous equation is important physics encompassing much of the calculations HRP do at considerable length. To appreciate the significance of Eqs. (2.1) to (2.3), and why k must be finite and nonzero, we digress to consider the action of an electromagnetic wave on a free electrically charged particle in a little more detail. We do so because the HRP conjecture amounts to the assertion that accelerating charges see incident waves (made real by the acceleration) that act on them via the Lorentz force. As far as the interaction is concerned, this is completely equivalent to a charged particle initially at rest (as the accelerating charge is in its frame of instantaneous rest) acted upon by an incident electromagnetic wave. We keep things simple by considering only a linearly polarized monochromatic plane wave. A ZPF interaction will of course, be more complicated because of the presence of a spectrum of frequencies. But the crucial physics can be seen with a single frequency.

When the plane wave first acts, the electrically charged particle responds only to the electric field of the wave since it is at rest. This causes the particle to execute a sinusoidal oscillation in the direction of the electric field, that is, perpendicular to the direction of propagation of the wave. If we ignore radiative damping of the motion of the particle, the equation of motion for the response to the electrical field is

$$|m\mathbf{a}| = eE_o \sin(\omega t) \tag{2.4}$$

where m is the mass of the particle, ω the angular frequency of the incident wave, and E_o its amplitude. This may be integrated to give the velocity, v, of the particle,

$$|v| = -(e/\omega m)\, E_o \cos(\omega t) \tag{2.5}$$

which may now be substituted into the magnetic part of the Lorentz force to get the effect of the magnetic part of the incident wave. Since $\mathbf{B} = B_o \sin(\omega t)$ and v and $\mathbf{B}$ are orthogonal, we immediately obtain

$$|\mathbf{F}_{\text{mag}}| = -(e^2/2\omega mc)\, E_o B_o \sin(2\omega t) \tag{2.6}$$

in the direction of the incident wave. So, at this approximation, the magnetic part of the Lorentz force makes the charged particle execute an oscillation in the propagation direction of the wave, but no stationary force is exerted on the particle, for the force in Eq. (2.6) time-averages to zero.[3]

In order to get a stationary component of $\mathbf{F}_{\text{mag}}$ that might be interpretable as an inertial reaction force, we *must* include radiative damping. The oscillation of the particle driven by the incident wave must in turn drive the emission of radiation by the oscillating particle, as, of course, indeed it does. Now elementary considerations of energy and momentum balance of the sort customarily invoked in discussions of radiation pressure tell us that the particle will experience a stationary force under the action of the incident wave. Part of the energy absorbed by the particle from the

---

[3] The combined action of the electric and magnetic parts of the Lorentz force predict that the motion of the charge in response to the electromagnetic wave will be a “figure eight.” This is in fact true, for the scattered radiation that results from such motion contains harmonics that have recently been detected when a plasma was subjected to intense bursts of laser light.$^{(6)}$

incident wave is reradiated. But the momentum flux in the incident wave does not correspond to the momentum flux in the stimulated radiation (plus the unabsorbed flux), so the particle must experience recoil to carry away the momentum needed to conserve total momentum in the interaction.

We now ask how, at the level of an incident wave acting on a free point particle, the stationary force on the particle comes about. The mechanism that one might guess is that the radiation reaction force produced by the stimulated radiation emitted by the particle might account for the stationary force that must be present. But at nonrelativistic velocities at least, the stimulated radiation is symmetrical about the direction of acceleration of the particle induced by the electric field of the incident wave, so there is no net radiation reaction force in the direction of the incident wave. If the stimulated radiation, however, induces a small phase lag between the incident electric field strength $|\mathbf{E}| = E_o \sin(\omega t)$ and the induced velocity (via the induced acceleration), then a stationary force results. In other words, if $|v| = -(e/\omega m)\, E_o \cos(\omega t + \phi)$, then $\mathbf{F}_{\text{mag}}$ acquires a time-independent term, which depends on the sine of ϕ, in the direction of the incident wave. The phase lag between $\mathbf{E}$ and v is precisely the sort of effect that one would expect when a dissipative force is coupled to a fluctuational process. The effect of the stochastic averaging that HRP do at great length is to introduce a small phase lag between the electric component of the ZPF radiation incident on the charge and the motion it induces that leads to a stationary component of $\mathbf{F}_{\text{mag}}$, just as adopting Eq. (2.1) above with k finite and not equal to zero does when the time-averaged Poynting vector of the ZPF seen by the particle is nonvanishing. Note, too, that a radiation reaction-induced phase lag between $\mathbf{E}$ and v must also be included in Eq. (2.2) to recover a stationary force from Eq. (2.3) (unless $\mathbf{E}_{\text{zpf}}$ and $\mathbf{B}_{\text{zpf}}$ are separately taken to be nonvanishing time averages). Radiation reaction, therefore, we see to be *crucial* to the electromagnetic ZPF conjecture on the origin of inertial reaction forces. It is not sufficient to talk about fluxes of radiation impinging on electric charges in general terms if one is to specify how such fluxes communicate forces to those charges. Coupling of the flux to the charge, including radiation damping, must be specified.

Inspection of Eq. (2.3) makes clear that the interaction of an accelerating charge with the hypothetical ZPF will be inertia-like in that the force, being proportional to e^2, does not depend on the sign of the interacting electric charge. And since the force, as we have just seen, is a radiation pressure, it always acts to oppose the externally induced acceleration. In order to see if this effect can plausibly account for gross inertial reaction forces all we need do is evaluate k. HRP, in effect, do this, finding that k depends on the radiation damping that the charge would experience even if there were no ZPF present at all, as should be expected in light of our above remarks. In particular, after a lengthy calculation they find

$$\mathbf{F}_{\mathrm{zpf}} \approx (\Gamma \hbar \omega_c^2 / 2\pi c^2)\, \mathbf{a} \tag{2.7}$$

where $\hbar$ is Planck's constant divided by 2π, and ω_c the "cutoff" frequency for interacting radiation. Note that ω_c must be finite (and nonzero) to recover a realistic force from Eq. (2.7). Γ is the coefficient of radiation damping:

$$\Gamma = 2e^2 / 3m_o c^3 \tag{2.8}$$

where m_o is an *unobservable* "bare mass" of the interacting particle. (Note that m_o must be included to keep the charge from responding to the ZPF with infinite acceleration. That is, inertial mass, allegedly being explained at the "observable" level, must be put in by hand at the "unobservable" level to make this conjecture work.) HRP suggest that the parenthetical factor on the RHS of Eq. (2.7) be taken to be the inertial mass of the charged particle and engage in speculation on the relationship of the cutoff frequency and "bare mass" to Planck-scale phenomena.

Since m_o is unobservable and freely adjustable, you can make anything you want of this calculation should you so choose. Of course, that does not really qualify as an explanation of inertia. To be considered a candidate explanation of inertia, compelling arguments would have to be presented showing that the factor in parentheses in Eq. (2.7) really does correspond to the observed masses of things. But even were that possible, it would not be a successful candidate explanation for the reasons laid out in the next section.

A point related to the parenthetical expression in Eq. (2.7), however, is worth mentioning. In Appendix C of HRP(16a) they attempt to tie their expression [their Eq. (110)] for the inertial mass of an accelerated particle to a radiative mass shift predicted by standard theory. In fact, the mass shift in question is a consequence of demanding that radiation reaction processes obey the conservation of energy and momentum locally at each instant. It follows from the $\ddot{\mathbf{x}}^2$ term in the Lorentz–Dirac equation of motion, producing a small effective increase in the restmass of radiating charges (and, equivalently, charges subjected to intense radiation fields). The effective restmass increase reduces the kinetic energy acquired by the charge during acceleration so that the change in kinetic energy of the charge plus radiated energy is equal to the work done by the accelerating force. It is rather small—hardly the origin of gross inertial reaction effects. HRP's identification of their result with this effect, therefore, is puzzling at best. But if their m_o is taken to be the same as the m_o of the calculation in their Appendix C (the observed particle restmass), then the effects are the same as they claim. This suggests that HRP's conjecture allows one to calculate this minuscule, plausibly real effect with the techniques of "stochastic electrodynamics" they employ and that this effect has nothing to do with normal inertial reaction forces.

In their most recent papers RH have steered clear of schemes that involve explicit couplings of the putative ZPF to the accelerated charges allegedly affected. Instead, they calculate the momentum flux of the putative ZPF seen in an accelerating frame of reference and simply assert that it is responsible for the inertial reaction force. (In this respect, their treatment parallels the elementary treatments of radiation pressure alluded to above.) The result is an expression for the inertial mass of an accelerating particle as a function of the spectral energy density of the ZPF that does not explicitly include the electric charge of the object acted upon, because, "on purpose, no interaction features were included in the analysis."(24) Moreover, they decline to carry through the integration over the frequency spectrum of the ZPF, thereby suppressing the cutoff frequency in their expression for inertial mass [in this connection see Eq. (30) in Ref. 24a]. They do, however, include an undefined frequency-dependent coupling coefficient in anticipation of problems that arise when coupling to charge is to be introduced. We find this approach something less than candid at best, for the physical manifestation of inertia is reaction *forces* that arise from the coupling of charges to fields. Momentum fluxes of fields without specified coupling to their charges cannot reasonably be claimed to account for inertial reaction forces. This point is crucial. It is the nature of the coupling of fields to charges that reveals the assertion that inertia arises from any ZPF (other perhaps than that of quantum gravity, which is not yet, and may never be, invented) as untenable.

2.2. Why Inertia Is Not Caused by an Electromagnetic ZPF

The obvious reason to reject the electromagnetic ZPF as the cause of inertial reaction forces, known all along, is the fact that if the ZPF really did exist, the gravitational effect of the energy resident in it would curl up the universe into a minute ball. In fairness to HRP, we note that this is a fundamental problem for all relativistic quantum field theories. It is the so-called "cosmological constant problem." (See Ref. 3, pp. 162–164.) Up to the present at least, deeply implausible "fine-tuning" has been required to get rid of the predicted vacuum energy that is plainly not there. Proponents of the HRP scheme, including HRP themselves, have advanced arguments to deflect this objection. While we find these arguments unconvincing, evidently others are less troubled by this defect of their conjecture. So, rather than just take issue with the viability of their proposals to justify the dismissal of the physical effects of infinite vacuum energy densities, we also give a different argument that shows beyond any reasonable doubt that inertia is not an electromagnetic ZPF effect. Since it does not require belief in the correctness of general relativity theory, perhaps it will find wider acceptance. Before proceeding to that argument, however, we briefly recount HRP's arguments relating to the energy density of the vacuum and explain why we think them unconvincing.

When HRP first advanced their proposal that inertia was due to an electromagnetic interaction with the ZPF, it was suggested that one might be able to deal with the formally infinite energy density of the vacuum entailed by their ideas (and other quantum field theories) by invoking gauge invariance of the first kind. In the case of Newtonian gravity, this type of gauge invariance allows one to add any arbitrary value of the gravitational potential to the computed value of the potential—*as long as the additive value is applied globally* (everywhere). This can be done because Newtonian gravitational force depends only on the gradient of the potential, not its absolute value. Linear theories usually possess this type of gauge invariance. The problem in the case of gravity, as GRT makes plain, is that gravity/inertia is not linear. And since any alternative theory of gravity/inertia must mimic GRT to account for GRT's known effects, it, too, will be nonlinear in the last analysis. As Peters(21) pointed out many years ago, nonlinear theories of gravity do not admit gauge invariance of the first kind. So the gravitational effects of the energy resident in any ZPF cannot be gotten rid of in this way.

Since potentials (and their energies) associated with gravity/inertia cannot be scaled away by a global gauge transformation, another justification for ignoring the energy density of the quantum vacuum must be found. The argument sometimes advanced by supporters of quantum vacuum fluctuations (including a referee of this paper) is that if gravity and inertia arise from electromagnetic interactions as claimed, then energy *per se* ceases to be a source of gravity. So one can have as much energy in the vacuum as one wants. As long as it does not lead to an interaction between electromagnetic radiation and electric charges that violates observed facts, the energy resident in the vacuum has no physical effect and can be ignored. In particular, it does not curl up the universe. In GRT such energy must do so, for spacetime geometry is determined by the matter tensor (whose dominant component in most cases is mass-energy).

Appealing though this argument may seem to those enamored of the quantum vacuum, it conflicts with known facts. For example, the fact of relativistic mass increase with *velocity* cannot be consistently explained on the HRP hypothesis. This is a consequence of the fact that electric charge, unlike relativistic mass, is a Lorentz invariant quantity. It does not depend on the velocity of the particle that carries it. So, in HRP's expression for the inertial mass of an electron (or "parton" as they call it) moving with constant velocity *in the lab frame of reference*:

$$m_i = (e^2 \hbar \omega_c^2 / 3\pi m_o c^5) \tag{2.9}$$

e is independent of the particle velocity. Of course, one could "freely adjust" the unobservable bare mass m_o to make things work out right. But there is no plausible physical reason to do so other than that something must be done to get relativistic mass dependence.

The only other way to get relativistic mass dependence in Eq. (2.9) with this hypothesis is to assume that the cutoff frequency ω_c is velocity dependent. As it turns out, this is sometimes done in quantum field theoretic arguments. (This was one of the reasons Dirac was convinced to his dying day that something was fundamentally wrong with quantum field theory.) In this case, however, this cannot be done. This follows from the fact that in order to preserve Lorentz invariance—that is, to not violate special relativity theory—one must assume that the spectral energy density of the ZPF goes as ω^3. With this spectral energy density, an electric charge moving with constant velocity, whatever it may be, always sees the same ZPF. So an electron, say, moving with large velocity in our lab sees exactly the same ZPF as one at rest. In particular, the integrated momentum flux of the ZPF impinging on any particle in *any* state of inertial motion is zero, irrespective of the inertial frame of the observer. But the mass energy of the particle is velocity dependent, notwithstanding that the total incident momentum flux is zero in all inertial frames of reference. How, without exerting an external force that induces an acceleration, can we ascertain this? Gravity. Relativistic mass is a source of the gravitational field of the particle that can be probed by a distant "test mass" that, being arbitrarily small, has a negligible effect on the particle.

Now, for the moving electron to have a higher cutoff frequency than the one at rest, since the ZPF seen is velocity independent, some physical property of the moving electron that determines the cutoff frequency must be different from that for the electron at rest. In the HRP conjecture, the cutoff frequency is determined by the size (very small) of the particle since, they claim, the particle cannot respond to waves with frequencies corresponding to wavelengths much smaller than the particle itself. Consider the response of our two particles to an accelerating force in the direction of the velocity of the moving particle. The aspect of the size of the particle that is relevant to the cutoff here is its lateral dimension (the direction of the excursion induced by $\mathbf{E}_{\mathrm{zpf}}$). Since lateral dimensions are invariant under Lorentz boosts, the cutoff for both particles is the same. So no relativistic inertial mass increase in the moving particle should occur on HRP's hypothesis, contrary to observed fact. The situation is different when the acceleration is orthogonal to the velocity, and Lorentz contraction of the particle will lead to an increased cutoff frequency. But that does not obviate the fact that no inertial mass increase is expected when the acceleration is in the direction of motion. The usual notion of relativistic mass does not suffer from this defect.

In their more recent work RH, attuned to this problem, suppress the coupling of the ZPF to the accelerating charge(s) that experiences the force exerted by the ZPF that is to be interpreted as inertia. They thereby obtain a force that is equal to a Lorentz factor times an expression they assert to be the mass times acceleration. And they claim that they have recovered correct relativistic mass behavior because of the presence of the Lorentz factor. The expression for the mass is the appropriately averaged Poynting vector for ZPF modes without the integral over frequency that leads to the cutoff being performed. Encouraging though this may appear, when the integration is performed the cutoff must be introduced to recover a finite result. And coupling with radiation damping must be introduced, as we have seen above, to recover an inertia-like force. So, the Lorentz factor that makes it appear that relativistic mass behavior is present notwithstanding, the foregoing argument still applies.

We turn now to the argument, unrelated to vacuum energy density, that brings out another inconsistency of the electromagnetic ZPF conjecture on the origin of inertia. HRP and RH repeatedly remark in the course of developing their claims that a detailed understanding of inertial reaction forces must depend on an analysis of the coupling of leptons and quarks to the ZPF. We agree. And such an analysis is much simpler than evidently they assume. All we need do is note that $\mathbf{F}_{\mathrm{zpf}}$ is proportional to e^2/k, or, should you prefer, e^2/m_o. Since k (or m_o) is an undefined, adjustable parameter, we will not get very far if we look only at individual leptons or quarks. But by looking at collections of them we can sidestep this problem. For example, consider the neutron and proton, each a different collection of up and down quarks. We start out assuming matters to be simple. We posit that k is a constant. From Eq. (2.3) we know that $\mathbf{F}_{\mathrm{zpf}}$ acting on each of the constituent quarks will be proportional to e^2, where e is now the quark charge. Next we note that a proton consists of two up quarks and a down quark, whereas a neutron contains two down quarks and an up quark. Now the electric charge of an up quark is $+\frac{2}{3}$ electron charge, and that of a down quark $-\frac{1}{3}$ electron charge. Summing the squares of the quark charges in the proton and neutron, we find $\frac{9}{9}$ and $\frac{6}{9}$, respectively. Because the quarks are asymptotically free for frequencies near the cutoff frequency, the total force on the nucleon is just the linear sum of the separate forces on each of the quarks, each of which goes as its own e^2. It follows immediately that if the ZPF acts on the constituent quarks of the proton and neutron as RH claim, the inertial mass of the proton should be 1.5 times that of the neutron. This is not right.

What can we do to salvage HRP's hypothesis? Well, a few lines of algebra reveal that to get the observed inertial masses of the neutron and proton, we must assume that the inertial masses of the up and down quarks are almost exactly the same. No other proposition will work. So we might argue that the ZPF does not interact with the observed electric charges of the quarks. Rather, we might suppose that it interacts with the "bare" and "dressing" charges that quarks putatively possess. These both might be enormous, and since it is the sum of their squares that determines the inertial mass, it is conceivable that the plainly different observed charges are insignificant by comparison with the bare and dressing charges they consist of. We can accommodate this in our formalism by positing that k (or m_o) has a value for each type of quark such that e^2/k has the same value for both the up and the down quark—a value that yields $\mathbf{F}_{\mathrm{zpf}}$ for each quark that is one-third of the neutron/proton inertial reaction force.

We have avoided the problem of the neutron and proton mass discrepancy by invoking k, to put it colloquially, as a fudge factor. This, however, does not solve all of our problems. Up and down quarks are also the constituents of pions. The charged pions, for example, consist of either an up and an antidown quark or a down and antiup quark. Where the nucleons with their three quarks have a mass of 939 MeV, the charged pion's mass is 140 MeV. The nucleon mass requires that the quark mass be 313 MeV regardless of whether it is an up or down quark. Should this be true, the only way to recover the observed pion mass is to assume that the antiquark mass is -173 MeV. This is not good news. If inertial mass arises from the interaction of electric charge with the ZPF as claimed, then, since it depends on the square of the electric charge, inertial mass should be positive definite. It follows immediately that the antiquark mass should be the same as the quark mass, as it must be to account for the masses of the antinucleons in any event. Perhaps a miracle occurs when antiquarks form mesons that transforms the sign and value of our adjustable factor k (or m_o). We identify this transformation as miraculous because physically it involves the annihilation of a large amount of electric charge—a violation of charge conservation. Miraculous transformations of this sort, however, are not usually considered positive features of "natural" explanations of things.

A proponent of the ZPF conjecture on the origin of inertia might object that the model we have used here is excessively naive. For instance, one might claim that nucleons and mesons are complicated entities with the electric charges bound to each other and so forth. Indeed, HRP(16b) make just such a claim which we examine below. It is worth recalling at this point that HRP themselves state that because the dominant part of the ZPF interaction occurs near the cutoff frequency, which they argue to be of the order of the exceedingly high Planck frequency, the elementary charged particles that make up bound systems can be considered asymptotically free. And when one examines the facts of the structure of hadrons, it is evident that the foregoing analysis captures much of reality. The masses of constituent quarks in a nucleon are measured to be about 10 MeV each out of the roughly 940 MeV nucleon mass. This means that almost all of the nucleon mass is to be found in the restmassless, zero-electric charge, photonlike gluons that bind the quarks.

On the HRP hypothesis, the gluons that bind the quarks in hadrons would have to consist of precisely equal, enormous amounts of electrically charged dust that propagates across the nucleon at the speed of light. The gluons, thus, would have to have the chief characteristics we attributed to the bare and dressing charges above. And because the gluons propagate at light speed (being restmassless), there is yet another problem. The proper acceleration of things that travel at light speed is formally infinite (since they cover arbitrarily large distances in zero proper time) (in this connection see Ref. 23, p. 34). In order to get the observed inertial mass of hadrons we must believe that ZPF photons propagating at light speed can act on gluons, also propagating at light speed, and change their already infinite accelerations.

Is there any other way to deal with the problem of the hadron masses in the scheme of the electromagnetic ZPF conjecture? Well, shortly after one of us (J.F.W.) posted an early version of the hadron mass problem to his website (spring 1997), as mentioned above, HRP(16b) addressed this problem in a general sort of way. This is what they had to say.

> This [the foregoing hadron mass discrepancy presentation] is clearly a naive argument given the conceptual uncertainty of what “mass” actually means for an individual quark which cannot exist in isolation. Nonetheless, taking this paradox at face value does offers [sic] a useful perspective for speculation.
>
> The expression [Eq. (2.9) here] for m_{zp} of an individual particle as derived by HRP [1994] involves two free parameters, Γ_z [given by Eq. (2.8) above] and ω_c. In HRP [1994] we assumed that ω_c was some cutoff frequency dictated either by an actual cutoff of the ZPF spectrum (such as the Planck frequency) or by a minimum size of a particle (such as the Planck length). Let us assume

> that in place of a cutoff frequency there is a resonance frequency which is specific to a given particle, call it ω_c.
>
> ... It would not be surprising that a bound triad of quarks such as the uud or the udd would have a radically different resonance as an ensemble. The resonance of a mechanical system bears no simple relationship to the resonances of its component parts. On this basis it would be easy to see how the same three quarks could have a totally different mass collectively than individually.

Is such a scheme actually plausible? Well, ω_c is not a "resonance" frequency. It is the upper limit in the integration over the frequency spectrum of the ZPF, and if that limit is not imposed, the result of the integration, and the inertial mass of the particle, is infinite irrespective of any resonances that may be present at finite frequencies. Remember, the spectral energy density of the ZPF goes as ω^3, so invoking a "low"-frequency resonance will not suppress the cutoff unless the cutoff is assumed to lie quite close to the resonance frequency. A plausible justification for this assumption would have to be provided for it to be taken seriously. But setting this problem aside, let us examine this conjecture to see if it really makes any sense.

We first note that the masses of the quarks, 5 MeV for the up and 10 MeV for the down, and the mass of the electron, 0.5 MeV, all arise allegedly from the interaction with the electromagnetic ZPF at or very near the Planck frequency, not at some "resonance" frequency. In the case of the electron there can be no question that this must be true, for there is no evidence that the electron is a composite system that might have such a resonance. In the case of quarks, their masses are determined by scattering techniques that probe their behavior in circumstances where they are "asymptotically free" and evidently structureless. So their masses of a few mega-electron volts, at least roughly consistent with the electron mass, seem to arise in the same way as the free electron mass. [Here, by the way, we have another problem: the 5 MeV up quark has a $+(\frac{2}{3})e$ charge and the 10 MeV down quark a $-(\frac{1}{3})e$ charge—exactly the opposite of expectation on the HRP hypothesis.] So the resonance hypothesis must account for constituents with masses of a few mega-electron volts behaving with a mass of nearly a thousand mega-electron volts when they "resonate," due presumably to their being coupled to each other via gluons.

Since inertial reaction forces according to the HRP hypothesis are attributed to the magnetic part of the Lorentz force, as embodied in Eq. (2.2) above, we see that the supposed resonance behavior must do one, or both, of two things: either it must increase the velocity induced by the electric field of the incident radiation and/or it must increase the effective amplitude of the magnetic field seen by the charges. As we are trying to account for the hadron mass problem in terms of a resonance of several quarks, we reject out-of-hand any idea that Γ (because its only adjustable parameter is the "bare" mass m_o), and thus k, could change in the circumstances contemplated. What about the velocity? Well, this is limited by relativity. It cannot exceed c. So nothing is to be gained here (since this is already built into the Planck frequency—c divided by the Planck length). What about the amplitude of $\mathbf{B}_{\mathrm{zpf}}$? This will depend on the energy density of the ZPF at the resonance frequency. The resonance frequency, since we are assuming the typical velocity of the resonance excursion to be c, will be determined by the physical dimensions of the composite system. In the case of hadrons, this is of the order of 10^{-13} cm, 20 orders of magnitude larger than the Planck length. Accordingly, the base resonance frequency at hadron dimension is about 20 orders of magnitude lower than the Planck frequency.

Does the fact that the resonance frequency of the assembly of quarks in hadrons is 20 orders of magnitude lower than the Planck frequency make any difference? Unfortunately for HRP, yes. In order for the ZPF to remain invariant under Lorentz boosts, its spectral energy density must go as frequency cubed. This means that at the hadron resonance frequency the ZPF energy density will be 60 orders of magnitude down from the energy density at the Planck frequency. And since the amplitude of the field strength goes as the square root of the energy density, $\mathbf{B}_{\mathrm{zpf}}$ at the hadron resonance frequency will be down by 30 orders of magnitude from its value at the Planck frequency, which allegedly leads to the masses of the free quarks and electrons. Even allowing for higher than first-order resonances, there simply is not enough energy present in the ZPF at the hadron resonance frequency to account for inertial reaction forces an order of magnitude and more larger than those produced at the Planck scale that act on the quarks individually.

Naiveté, like beauty, is at least to some extent in the eye of the beholder. So electromagnetic ZPF inertia advocates might want to dismiss the arguments we have presented above as naïve (as HRP apparently have). But we think that the chief point of the foregoing analysis should not require much elaboration. Inertia is *not* caused by the interaction of electric charge with a putative electromagnetic ZPF. Ignoring the issue of the energy density of the vacuum, we see that the ZPF scheme of HRP fails for a simple reason: the coupling of electromagnetic fields to electric charge cannot be made to mimic the universal coupling of gravity to mass-energy while preserving a correct representation of known electromagnetic phenomena without invoking quite implausible *ad hoc* hypotheses. The general conclusion to be drawn from all this is that this will also be true of any other ZPF since none of them, save gravity, have the right coupling behavior. So, at least as far as inertia is concerned, evidently the vacuum is not a seething reservoir of unlimited electromagnetic energy ready to oppose the acceleration of electric charges.

3. WHAT IS THE CAUSE OF INERTIA?

Gravity. RH give some arguments that purport to show that gravity is not the cause of inertia. We address them in Section 4 below. But first we show that when gravitomagnetism is taken into account, inertial reaction forces are easily accounted for. This was first demonstrated explicitly by Sciama in 1953.(25a) Earlier work by Einstein and later work by Nordtvedt lead to exactly the same conclusion. Subtleties that attend this business, worked out by eminent researchers over the years, we summarize later. Since our target audience includes nonexperts, we keep things simple. We consider only translational accelerations and employ conventional vector notation.

3.1. Sciama's Calculation

From the mid-1920s to the early 1950s, not much attention was paid to Mach's principle. Einstein, frustrated by his inability to include this principle in GRT, asserted that closure of the universe was the solution to the problem and attended to other matters. ("Closing" the universe forces one to accept that the properties of its contents cannot be related to anything "outside," for there is no outside.) Others were preoccupied with different concerns. In 1953, Dennis Sciama published a little paper, "On the Origin of Inertia," that revived interest in the subject. Save for a hiatus during the 1980s, Mach's principle has been a topic of contention ever since. Arguably, the persistence of this issue is attributable in no small part to the fact that Sciama's presentation was so simple, obvious (in retrospect anyway), and elegant. With but a few superficial modifications, we replicate Sciama's argument for the case of a test body translationally accelerated in a universe of constant matter density to illustrate the force of his argument. We ignore Hubble flow and other complications that becloud the underlying arguments.

It seems not to have been widely appreciated in the early 1950s that in the linear, weak field limit (where space is very nearly flat), the field equations of GRT can be approximated by equations that have essentially Maxwellian form. For that reason, evidently, Sciama made no appeal to GRT in his discussion of the origin of inertia. Instead, he simply asserted as an hypothesis that the field equations for gravity should be Maxwellian and then, using gravitoelectric and gravitomagnetic potentials and fields, went on to show that inertia could be explained as a gravitational force. In analogy with electrodynamics, Sciama took the gravitoelectric scalar potential ϕ to be

$$\phi = -\int (G\rho/r)\, dV \tag{3.1}$$

where G is the Newtonian constant of gravitation, ρ the matter density, r the distance from the test body to a volume element being considered, and the integration extends over all space. (Sciama did not include G in his potential, choosing instead to calculate it from the assumption of gravitational induction of inertia. We do not replicate this aspect of his argument.)

The gravitomagnetic vector potential also resembles its electrodynamic counterpart:

$$\mathbf{A} = -(1/c) \int (G\rho\, \mathbf{v}/\mathbf{r})\, dV \tag{3.2}$$

where $\rho \mathbf{v}$ is the matter current density in the integration volume element. These potentials are attended by subtleties of choice of gauge that we ignore for the moment. The gravitational force produced by the matter in the universe on our test body is now easily computed. First, we note that if our test body moves translationally with respect to the rest of the universe, the curl of A vanishes by symmetry. Direct gravitomagnetic forces are therefore absent in our circumstances. The gravitoelectric field is given by

$$\mathbf{E} = -\nabla\phi - (1/c)\, \partial\mathbf{A}/\partial t \tag{3.3}$$

Since the density of matter is uniform and the test body, by assumption, does not contribute to the gravitational field, $\nabla\phi$ vanishes and

$$\mathbf{E} = -(1/c)\, \partial\mathbf{A}/\partial t \tag{3.4}$$

To find $\mathbf{E}$ for the test body all we need to do is compute $\mathbf{A}$ with Eq. (3.2) and take its time derivative. To do this Sciama notes that, from the point of view of the test body, every part of the universe appears to be moving rigidly with velocity $-\mathbf{v}$ at each instant, so $\mathbf{v}$ can be removed from the integral. The remaining expression, save for the factor $1/c$, is the integral for the scalar potential ϕ, so in this case,

$$\mathbf{A} = \phi\, \mathbf{v}/c \tag{3.5}$$

Removing $\mathbf{v}$ from the integral in Eq. (3.2) to get Eq. (3.5) amounts to a choice of gauge with serious consequences that we mention below. Substituting **A** from Eq. (3.5) into Eq. (3.4) yields

$$\mathbf{E} = -(\phi/c^2)\, \partial\mathbf{v}/\partial\mathbf{t} \tag{3.6}$$

It follows immediately from Eq. (3.6) that if our test body moves with constant $\mathbf{v}$ with respect to the universe, it experiences no gravitoelectric force. If it is accelerated, however, **E** is not zero. And if $\phi/c^2 = 1$, then the gravitoelectric force is the inertial reaction force.

What does the proposition $\phi/c^2 = 1$ entail? Sciama shows that, up to factors of the order of unity, ϕ must be equal to $-GM/R$, M and R are the mass and radius of the observable universe, and the density of matter is about 10^{-29} g/cm^3—that is, about “critical” (or closure) density. Although this is true in the linear approximation, is this generally true? This is an important matter, for if it is generally true, then problems might be expected if the observed matter density should turn out to be less than critical (as recent observations of distant supernovae now suggest). Moreover, we are also faced with the fact that R and perhaps M are functions of time, so ϕ may be a function of time too. If inertial reaction forces are gravitational, are they epoch dependent? We sketch the answers to these questions below. But first we deal with a more important issue: What is the relationship between general relativity theory (GRT)—the essentially universally accepted, now well-tested theory of gravity—and Mach’s principle?

3.2. GRT and Mach’s Principle

Several years ago a conference was convened in Tübingen to address Mach’s principle. Should you think that contentious problems no longer attend Mach’s principle, a quick reading of the conference proceedings, published as Volume 6 in the *Einstein Studies* Series,(2) will disabuse you of that notion. Profound problems remain to be solved. One of the deeper problems is the fact, first established by de Sitter within a year or so of the publication of GRT by Einstein, that the field equations of GRT admit solutions that are manifestly “anti-Machian.” In particular, de Sitter found a cosmological solution of the field equations corresponding to an empty, expanding universe. Expansion was not the problem. The emptiness was. Eventually, it led Einstein to assert that the universe must be spatially closed. The other exact solution of the field equations—the Schwarzschild solution for a spherical massive object—presented problems, too, from the Machian perspective: Spacetime is asymptotically flat far from the object. Flat, or Minkowskian, spacetime is replete with inertial structure. That is, test bodies in Minkowskian spacetime experience inertial reaction forces notwithstanding that the spacetime is empty. If the only substantial body producing gravity lies infinitely far away, and gravity is needed to induce inertia, how can local test bodies have inertia?

More recently, other anti-Machian solutions of the field equations have been explored, most notably the Gödel solution, wherein the universe has a global rotation. (In other words, the "fixed stars" appear to be rotating for any observer at rest in a local inertial frame of reference.) We do not propose to deal with the problem of the existence of anti-Machian solutions of the field equations of GRT. Interesting though such universes may be, they are not directly relevant to our reality. Instead, we address a different, more practically oriented question: Does GRT say that inertial reaction forces are gravitational in a universe like ours? The answer to this question happens to be yes. But appreciation of that fact did not develop overnight.

Sciama, in 1953, thought that his arguments on the origin of inertia suggested that some theory other than GRT must be right. A vector theory of gravity is at best an approximation, and by 1964 Sciama had put his ideas into a tensor formalism equivalent to GRT.(25b) By the end of the 1960s, Sciama and co-workers had recast GRT in the form of integral equations to clarify further the relationship between Mach's principle and GRT.(26) The basic idea here was to express the local metric that determines dynamics (and thus inertial reaction forces) in terms of a volume and surface integral of the material sources within a specified "horizon." The horizon is defined as the most distant surface from which light signals presently arrive here. (It may be either an "event" or a "particle" horizon, depending on the cosmological model.) Since the surface integral represents the contribution to the metric from sources that lie beyond the horizon, demanding that the surface integral vanish is equivalent to asserting that inertial forces are produced solely by the matter that lies within the horizon. In 1970 Gilman(14) showed that the surface integral in this formalism for GRT does vanish in all Robertson–Walker cosmologies. This work eventually led to the investigations of Derek Raine,(22) who extended and generalized these results.

One of the peculiarities of field theory generally, and gravity in particular, illuminated by investigation of the integral formulation of GRT, is that objects beyond horizons can influence local events, notwithstanding that signals propagating at light speed may never connect them. As Ellis and Sciama(13) put it, we can "feel" the electrostatic force of an electrically charged object beyond an (expanding) event horizon even though we will never be able to see it via electromagnetic radiation. (This is a consequence of the fact that in the Coulomb gauge the electrostatic interaction effectively propagates instantaneously.) If, however, any electric charge present is isotropically distributed beyond the event horizon, then its influence (and thus the surface integral at the horizon) vanishes by symmetry. That is why Robertson–Walker cosmologies are Machian; their isotropy saves the day. While one may quibble with technicalities, it is known as a matter of fact that the distribution of matter at very large scales is remarkably isotropic. Accordingly, Robertson–Walker cosmologies arguably encompass our reality. It follows that GRT dictates that inertia is gravitationally induced irrespective of whether or not cosmic matter density is critical.

3.3. Linearized GRT and Inertial Reaction Forces

Although relativists concerned with Mach's principle may find the issues surrounding the existence of anti-Machian solutions of the field equations of GRT and precise mathematical formulations of the principle their central focus, those concerned chiefly with rapid spacetime transport may find such topics of peripheral interest at best. For them the question of greatest interest likely is: Does GRT lead to the same conclusion as Sciama's 1953 argument, namely, that inertial reaction forces are gravitational forces that arise from the existence of real gravitomagnetic interactions? As one would expect from the more general considerations of the preceding section, the answer to this question too is yes. Indeed, Einstein, before the completion of GRT late in 1915, had already created an approximate vector version of GRT equivalent to Sciama's formalism and understood its Machian implications as early as 1912 (see Refs. 1 and 17 on this matter). Einstein gave a series of lectures at Princeton University in 1921 that shortly thereafter were published as his book, *The Meaning of Relativity*.(12) In the last of those lectures Einstein addressed Mach's principle.

In light of de Sitter's work, Einstein could not claim that GRT fully encompassed Mach's principle; but he was at some pains to point out that, even in its linear approximation, GRT did suggest that inertia was at least partial gravitationally induced. The formal substance of his argument consisted of a solution of the GRT field equations in the linear approximation. He found scalar and vector potentials like those posited by Sciama 30-odd years later. And the equation of motion of a test particle turns out to be our (and Sciama's) equation for the gravitoelectric force, with two small differences. One is that Einstein included a term for rotation (that is, nonvanishing curl of the vector potential) that we assumed to be zero above. The other is an explicit contribution of gravitational potential energy to the mass of the test particle due to nearby "spectator" masses.

Since Einstein had recovered Sciama's expression for the gravitoelectric force early on, why did he not claim that inertia was completely, rather than partially, gravitational induced? Perhaps because in the early twenties the universe was thought to be but a minute fraction of its now known size. As a result, the value of the scalar potential could not reasonably be assumed to approximate the square of the speed of light, and the gravitational force computed with Eq. (3.6) would be but a minute fraction of actual inertial reaction forces. This, together with the anti-Machian character of Minkowski spacetime, evidently forced Einstein to claim that spatial closure was required to implement Mach's principle. One cannot help but wonder how Einstein would have argued had he known the true extent of the universe when he was developing GRT.

3.4. Modern Simple Machian Ideas and GRT

A more recent examination of Mach's principle in the context of the linear approximation field equations of GRT was carried out by Nordtvedt.(20) Nordtvedt's motivation was some unfortunate remarks by a panel of the U.S. National Academy of Sciences about gravitomagnetic effects being so small that they had not yet been observed. He did not consider gravitomagnetic forces that arise from cosmic matter. Rather, his point was that even in solar system-scale phenomena, gravitomagnetism plays an important role. (Nordtvedt's paper is mandatory reading for anyone who wants to investigate schemes involving inertia.) For example, as is well known, GRT gives the orbit of a test particle (planet) around a massive central body (the Sun) as an ellipse (plus a minute perihelion advance) when coordinates where the central body is at rest are employed. If one considers this situation in a moving frame of reference, however, one does not recover a Lorentz-boosted ellipse unless one takes into account the gravitomagnetic vector potential. Thus, while direct experimental observation of frame dragging by the Earth's rotation (see Ref. 8 in this connection) remains a worthy endeavor, the existence of the gravitomagnetic field is borne out by the fact that planetary orbits are elliptical for all observers irrespective of their state of motion.

Nordtvedt notes that the gravitomagnetic vector potential leads to linear accelerational frame dragging, as well as the better-known rotational effects. Using the linear-order (in the mass) parameterized post-newtonian (PPN) formalism for GRT, he shows that local inertial frames generally are dragged when nearby moving matter produces a nonvanishing vector potential. For instance, a celestial body that is translationally accelerated by some external force drags the space within it as

$$\delta \mathbf{a} = (4\phi/c^2)\,\mathbf{a} \tag{3.7}$$

[This is the last (unnumbered) equation in Nordtvedt's paper assuming GRT values of PPN coefficients and written in the notation we are using.] **a** is the acceleration of the body and $\delta\mathbf{a}$ is the accelerative dragging of the local inertial frame of reference at some point inside the body.

Nordtvedt does not ask the question, suggested by Sciama's argument, What is the condition for the accelerative dragging of local space to be complete? That is, when does an accelerating body drag everything within it in such a way that there is no relative acceleration of the body and any of its parts. While this notion may be foreign to nonrelativists, it is quite common in GRT. For example, near the event horizon of a black hole, this type of dragging of spacetime takes place. Here, however, we are not interested in local dragging of spacetime of this sort. Rather, we want to know the condition for global dragging of space should the universe, viewed as a three-dimensional spatial hypersurface of (arbitrarily chosen) simultaneity, be subjected to a transient acceleration—a "jerk" that is—by some "external" agent. Evidently, we are talking about a thought experiment, for, put colloquially, this amounts to the question: Should God jerk the universe at some instant, what condition must be satisfied so that such jerks are undetectable by us inhabitants? Given Eq. (3.7), the answer to our question is plain. If $4\phi/c^2 = 1$, then space everywhere on the hypersurface is rigidly dragged. And up to the factor 4, it is Sciama's condition for inertial reaction forces to be gravitomagnetic in origin. So should we try to keep some element of the universe from being jerked by God, we would have to apply a compensating force. And the relatively accelerating universe would produce what we interpret to be an inertial reaction force. Inertial forces *are* gravitational forces.[4]

So what? A determined, dyed-in-the-wool devotee of the electromagnetic ZPF conjecture on the cause of inertia might argue (as has a commentator on this paper in the review process) that all of the material in this section (and Section 4 too) is nothing more than an uninteresting, empty tautology. After all, the equivalence of inertia and gravity is implicit in the equivalence principle, so it is to be expected that in GRT, inertial forces are gravitational forces. On this view, saying that inertial reaction forces are caused by the gravitational interaction with cosmic matter is no explanation of the origin of inertia because the cause of gravity and inertia (by a ZPF mechanism or whatever) is not provided. Should such claims be taken seriously? Well, as we have seen, the electromagnetic ZPF is not the

---

[4] We note here that the issues addressed in this and the preceding section can also be dealt with in the context of elaborations of GRT, for example, an Einstein–Cartan type of theory with teleparallelism. (See Chap. 3 of Ref. 27 in this connection.) But it is important to note that it is not necessary to invoke elaborations of GRT to explain inertial reaction forces. They are already accounted for by GRT.

cause of inertia, and it follows from the equivalence principle that gravity is also not so explained. So the electromagnetic ZPF is not going be the sought explanatory mechanism. Maybe something else along these lines can be cooked up. But is it required to "explain" inertial reaction forces? Only if you are prepared to accept the proposition that the phenomena of, say, electromagnetism are not explained by constructing an inferential path to them from Maxwell's equations and the Lorentz force (allegedly because they are already implicit in those equations and the principles upon which they rest). Understanding and explanation, however, consist precisely in showing how specific phenomena follow as a consequence of general principles and fundamental formalisms. In this sense, inertial forces are explained as gravitational in origin by relativistic gravity—a proposition that evidently is not yet universally understood.

4. A SUMMARY OF SOME SIGNIFICANT SUBTLETIES

Were the arguments in the preceding section without subtleties, no doubt they would be widely known and understood, and we would not have written this paper. But they are attended by several issues that can becloud the basic argument. To illustrate the sort of problems one can get into, we start by considering the issue of "tensor mass."

4.1. "Tensor Mass" and the Origin of Inertia

RH, attempting to establish that Sciama's argument about the origin of inertia was wrong, thereby creating a need for their ZPF explanation of it, asserted that were Sciama's argument right, then mass would have a tensorial character. That is, the masses of objects generally should depend on the spatial direction in which they are measured. In fact, very accurate experiments carried out around 1960—the so-called Hughes–Drever experiments—showed that mass is not tensorial. The idea that the measure of mass might depend on the direction of measurement, assuming that inertial reaction forces are gravitational forces, has a simple, intuitive origin. In addition to the force exerted by the smooth, large-scale distribution of matter, forces are also exerted by local concentrations of matter: the Earth, the Sun, the Galaxy, and so on. Given the anisotropy of the forces due to local matter concentrations, perhaps those anisotropies cause corresponding anisotropies in the measure of mass. That is, should we measure the mass of an object by the extension of a spring balance when it is accelerated with a fixed electromagnetic force, perhaps the extension will depend on the direction of the acceleration should an anisotropic gravitational field be present.

The idea that mass might be tensorial, first broached in the context of Mach's principle by Cocconi and Salpeter in 1958,(9) was exhaustively investigated by Dicke and others in the early 1960s. Compelling reasons exist to believe that mass is not tensorial. That this should be true can already be appreciated by considering Eq. (3.3):

$$\mathbf{E} = -\nabla\phi - (1/c)\,\partial\mathbf{A}/\partial t \tag{3.3}$$

The effects of local, anisotropically distributed matter enter this equation through the $\nabla\phi$ term, that is, the usual gravitational force term. Since it is the spatial variation of ϕ that causes this term to be nonvanishing, and as the spatial variation of ϕ is not an explicit function of the *acceleration* of its matter sources, this term makes no contribution to inertial reaction forces, even when it is nonzero. ϕ also enters the $(1/c)\,\partial\mathbf{A}/\partial t$ term in Eq. (3.3) since $\mathbf{A} = (1/c)\,\phi\,\mathbf{v}$, as we have seen above. Unlike $\nabla\phi$, ϕ is a scalar, without directionality. So $(1/c)\,\partial\mathbf{A}/\partial t$ cannot have different values in different spatial directions arising from gravity, as would have to be the case to make inertial reaction forces, and thus mass, direction dependent. This direction independent term is the origin of inertial reaction forces, as we have seen above.

Dicke's arguments are more elegant and profound than this simple examination of Eq. (3.3). But the conclusion is the same. Briefly, Dicke shows that any initially assumed tensorial mass can be absorbed into the spacetime metric by a conformal transformation if the mass anisotropy is universal. That is, if all types of matter respond the same way to the anisotropy, then

$$m_{\mathrm{ij}} = m f_{\mathrm{ij}} \tag{4.1}$$

where m_{ij} is the tensor mass, m the scalar mass, and f_{ij} a universal, dimensionless metric tensor field. Though the tensor field may have anisotropies at extended scales, it is locally Minkowskian in sufficiently small regions of spacetime. It follows that inertial reaction forces are not direction dependent, as is observed. Dicke's point is that the Hughes–Drever experiments are actually tests of the local isotropy of spacetime (rather than tests of supposed tensorial mass), and the results confirm to exquisite accuracy the assumptions of GRT. (Dicke's papers on tensor mass are reprinted in Ref. 11.)

4.2. The Locally Measured Value of ϕ

In Sections 3.1 and 3.3 we mentioned that the value of ϕ might depend on the cosmological epoch or the presence of nearby "spectator" masses. Were either of these possibilities true, inertial reaction forces would be time or space dependent were Sciama's simple vector potential model of gravity correct. In the case of epoch dependence, it would appear that the value of the gravitational constant would change in cosmic time. Owing largely to Dirac's speculations on the "large numbers hypothesis," the time dependence of G has been extensively investigated over the years. No substantial evidence exists that suggests that any variation has taken place. Likewise, no evidence has ever been adduced that supports that Einstein's suggestion that the masses of things should depend on the presence of nearby matter is right. The question then is: Naive intuitions notwithstanding, should we really expect inertial reaction forces to be epoch or location dependent? The principle of relativity dictates that the answer to this question is no. The total gravitational potential, like the speed of light in vacuum, therefore must be a *locally measured* invariant.

The first person to argue, persuasively, that "spectator" matter should not contribute to the masses of objects was Carl Brans, in 1962.(5) The essential point of Brans' argument is that in GRT, in sufficiently small regions of spacetime, by construction, spacetime is sensibly flat. That is, in the limit as the region of spacetime under consideration becomes arbitrarily small, the local metric differs inappreciably from the Minkowski metric of flat spacetime. Therefore, the masses of small objects, electrically charged or neutral, placed in a small, shielded laboratory must be unaffected by the presence of spectator masses outside, but nearby, the laboratory. For, since the metric in the laboratory is sensibly flat, the effects of the spectator masses can be removed by a coordinate transformation to a local, freely falling, inertial frame of reference. In that local inertial frame the charge to mass ratios of the electrically charged objects must be insensitive to the presence of the spectator masses if the principle of relativity is to be preserved. (Were the charge to mass ratios dependent on the presence of the spectator masses, then gravitational fields could always be distinguished from accelerated frames of reference, a violation of the equivalence principle that is the cornerstone of GRT.) Since electric charge is a locally measured invariant quantity, it follows that mass must be too. So if gravitational potential energy contributes to the masses of objects (as in fact it does), then the total gravitational potential must also be a locally measured invariant in GRT. (This is already suggested by the fact that the dimensions of gravitational potential are velocity squared, those of c^2 to which, up to factors of order unity, it is evidently equal.)

By invoking local flatness of spacetime and the Minkowski metric to establish the locally measured invariance of the total gravitational potential, it may seem that we have created a paradox. We have used the notoriously anti-Machian Minkowski metric to confirm the local invariance of the gravitational potential needed to ensure that gravitationally induced inertial reaction forces are epoch and location independent, as observed. It seems that Mach's principle rests ultimately on an anti-Machian foundation. In this connection, however, it is worth remembering that there are *two* spatially flat metrics in GRT, not one. In addition to the Minkowski metric, there is the metric of Robertson–Walker cosmology for "critical" cosmic matter density. So the local flatness needed to make it possible to eliminate the effects of spectator matter by a coordinate transformation can be achieved either way. When the critical matter density route is taken, the locally measured value of the total gravitational potential turns out to be that required to make inertial reaction forces gravitational in the simple linear-order approximation of Sciama's argument and Nordtvedt's PPN formalism.

A point that might cause confusion here merits mention. Since the locally measured value of ϕ must be an invariant to preserve the principle of relativity, one might think that the gradient of the gravitational potential must vanish everywhere. Accordingly, it would seem that no local gravitational fields should exist. But the gradient of a locally measured invariant need not vanish if it is not a *global* invariant. The total gravitational potential is not a global invariant. As a result, the "coordinate" value of the gravitational potential in some frame of reference may vary from point to point, notwithstanding that the numerical value measured at each point is the same everywhere. And the gradient of the potential in these coordinates may be nonvanishing. As a familiar example of this sort of behavior we point to the vacuum speed of light—a locally measured invariant—in the presence of a gravitational field. As is well known, the speed of light in intense gravitational fields measured by *nonlocal* observers (that is, the "coordinate" speed of light) is often markedly different from the locally measured value. And for those nonlocal observers, the speed of light in general will have a nonvanishing gradient in their coordinates. (Given the dimensional identity of gravitational potential with velocity squared, it is perhaps not so surprising that gravitational potential should have this curious property along with the vacuum speed of light.)

4.3. Mach's Principle and the Vacuum

As mentioned above, recent measurements of distant supernovae suggest that the vacuum is not completely devoid of any energy content whatsoever. The observed energy density, however, is not of the scale implied by relativistic quantum field theory. It seems to be of the order of magnitude of the density of ordinary matter. That is, of the order of 10^{-29} g/cm^3. So, while this vacuum energy will not prove the vast reservoir waiting to be tapped by some "free energy" scheme, it is nonetheless very curious. In the words of one commentator, "Nobody really understands the connection between the vacuum of quantum theory and general relativity. All you can really say is that the vacuum is weird and the vacuum sucks."(7) The vacuum "sucks" because the energy density is characterized by a negative pressure; that is, the energy density in the vacuum, being in a state of tension, is "exotic." The exoticity of the vacuum is a consequence of assuming that the energy density of the vacuum is positive which makes the pressure negative. (This is succinctly explained in Ref. 3, pp. 162–167.) Pressure, like energy density, gravitates, and if its exotic influence outstrips the influence of the energy density, then space is an exotic source of gravity. Globally the exotic influence of the vacuum seems to be about twice the positive energy density of normal matter (and the vacuum). This is not the electromagnetic ZPF that HRP would have be the origin of inertia. But it may be related to inertia.

To this point, we have been discussing a "weak" version of Mach's principle, namely, the version that merely asserts that inertial reaction forces are caused by the gravitational action of chiefly the most distant matter in the cosmos when external forces accelerate local objects. A stronger version of Mach's principle exists.(28) In the stronger version, as Berkeley insisted in Newton's day, a single object in an otherwise empty universe experiences no inertial reaction forces when external forces are impressed on it, and therefore it has no mass. That is, in the stronger version of Mach's principle *mass itself* arises from the gravitational action of the distant matter in the universe on local objects—mass is just the total gravitational potential energy a body possesses. [Recall that in nonlinear theories (like GRT) potentials cannot be arbitrarily scaled by an additive constant, so this assertion has real physical content.]

If we take this proposition seriously and assume that cosmic matter density is "critical" and positive, we immediately encounter a small problem: gravitational potential energies produced by positive matter are *negative* (because the force is attractive). One may reasonably ask how negative gravitational potential energy can be responsible for positive masses. If, however, space is "exotic" with an effective negative energy density that significantly exceeds the positive energy density of normal matter, then the total gravitational potential energy of local objects turns out positive, as observed. Indeed, the stronger version of Mach's principle *requires* that this be the case. How does this affect Sciama's argument regarding inertial reaction forces? Well, ϕ/c^2 must still be roughly equal to one, irrespective of the sign of ϕ. Note that if net cosmic matter density is positive as hitherto thought, then ϕ is negative, and our explanation of inertial reaction forces entails that $\phi \approx -c^2$. The observed exoticity of space, in addition to according with the stronger version of Mach's principle, also yields the result that $\phi \approx c^2$. That is, the effective gravitational potential of negative vacuum energy is positive, so the potential automatically has the same sign as the square of the speed of light, a positive definite quantity. This result is arguably at least as reasonable as $\phi \approx -c^2$.

4.4 Mach's Principle and Radical Timelessness

Today it is a commonplace that in creating special relativity theory in 1905 Einstein radically altered our understanding of the nature of time from the absolute notion of Newton to the local notion of the "twin paradox"—how time is affected by motion being the culminating point of his classic 1905 paper. But when Minkowski carried through the geometrical consequences of relativity theory a few years later, even Einstein at first rebelled against the concept of spacetime. The sense of the uniqueness of the present and the flow of time, belied by relativity theory, is so immediate that even long experienced relativists do not behave as if the totality of reality—past, present, and future—simply exists, the "present" being merely the event taken to be "now" at each point along a worldline. The "spatialization" of time—"radical timelessness" in the turn of phrase of Julian Barbour—demanded by relativity theory is so counterintuitive that almost no one really takes it seriously.

If relativity theory were the only thing that suggested that time does not flow, we might safely ignore this troubling feature of time. But quantum mechanics too leads to the conclusion that time does not flow. This is already indicated by now well-known "nonlocal" interactions inherent in quantum theory. Moreover, in so-called "delayed choice" experiments, now almost routinely done, actions that occur long after given events affect the way in which the events transpire. One can engage in subtle handwaving, but the message of these experiments is clear: events are influenced by both the past *and the future*, which, consequently, must already have some "objective" existence, just as it does in relativity theory. This peculiar aspect of quantum theory is made transparent in John Cramer's(10) "transactional interpretation" of quantum mechanics, where Wheeler–Feynman "absorber" theory is invoked to demystify quantum weirdness. The price paid in this interpretation is that the future, via retarded and advanced waves, has an objective existence and affects the present and past.

In a very limited sense, Mach's principle bridges the gulf that separates GRT and quantum mechanics. Inertial reaction forces in GRT are produced by a radiative interaction with chiefly the most distant matter in the universe. But they are instantaneous. So, it would seem, a "nonlocal" interaction equivalent to those encountered in quantum measurement must occur. As might be expected, RH argue that this feature of inertial interactions in GRT is a compelling reason to reject gravitational explanations of inertia and adopt their local vacuum fluctuations approach. Quite apart from issues relating to gravity, this argument looses much of its force when one takes into consideration that, as shown by Milonni (Ref. 19, especially the last section of Chap. 7), the fluctuation–dissipation theorem reveals that electromagnetic zero-point fluctuations can equally well be regarded as the action of a radiation reaction "field." This field can then be interpreted as the consequence of a Wheeler–Feynman "action-at-a-distance" retarded/advanced interaction with the distant matter in the universe. *That is, any local fluctuational explanation can be reinterpreted as a nonlocal, retarded/advanced interaction with distant matter.* One might think that since we are talking about "equivalent interpretations" here, the radiation reaction aspect of processes usually viewed as due to vacuum fluctuations can simply be ignored. Not so. As Milonni points out, strict calculation of vacuum fluctuation contributions to, say, the Lamb shift or Casimir effect yields prediction of only one-half of the observed effects. The other half is supplied by the radiation reaction "field." Milonni argues that at least this much of these effects must be attributed to the radiation reaction "field" in any event, for if this is not done, then it follows that quantum vacuum fluctuations *must* induce stimulated excitations of normal matter that are not seen. This is a matter of fact, not "interpretation." So adopting a local, zero-point fluctuation scheme (likely of any sort) to try to evade the radical timelessness implicit in classical field theories buys us nothing. We are stuck with radiation reaction (and its action-at-a-distance character) whether we like it or not. We can talk about local vacuum fluctuations if that makes us feel good, but there is no physical reason to believe they necessarily are real. The phenomena cited as evidence for their reality (the Lamb shift and Casimir effects for example) are explicable with a truly empty vacuum as long as the distant matter needed for the action-at-a-distance explanation is present out there.

Since one of us (J.F.W., 1996) has recently addressed radical timelessness at some length, we do not pursue this matter farther here. But we do note that correctly understanding the origin of inertia has some very strange consequences indeed.

5. CONCLUSION

To sum up, we have seen that local zero-point fields—other than quantum gravity perhaps (should it ever be invented)—cannot account for the origin of inertia. They do not display the universal coupling to mass-energy, possessed by gravity, that is required of any candidate field. GRT, however, already predicts the existence of forces that correspond to the inertial forces we experience in fact, so no new explanation of the origin of inertia is required in the first place. Objections raised against the gravitational origin of inertia, in particular, the putative tensor mass argument, we note were shown by Dicke nearly 40 years ago to be without merit. And the fact that the locally measured value of the total gravitational potential must be an invariant, like the vacuum speed of light, suffices to dispose of objections involving epoch, location, and environmental effects that might be present in a simple linear theory. This, too, through the work of Brans, has been known since the early 1960s.

So inertial forces are gravitational forces, as the principle of relativity and their universal coupling to mass-energy demand. Any theory that proposes that this is not the case should be regarded with deep suspicion, for it almost certainly violates the principle of relativity. And should you be interested in investigating rapid spacetime transport schemes, deep suspicion is indicated for any proposal that violates the principle of relativity. The likelihood that the principle of relativity is wrong is vanishingly small.

ACKNOWLEDGMENTS

Discussion with several colleagues and friends have been helpful. J.F.W. notes a stimulating conversation with A. Rueda at the Breakthrough Propulsion Physics Workshop at the NASA Lewis Research Center in August 1997. We are also indebted to two referees whose comments led us to strengthen the arguments presented in Section 2.

REFERENCES

1. J. B. Barbour, in *Modern Cosmology in Retrospect*, B. Bertotti *et al.*, eds. (Cambridge University Press, Cambridge, 1990), pp. 47–66.
2. J. B. Barbour and H. Pfister, *Mach's Principle: From Newton's Bucket to Quantum Gravity*, Vol. 6 of Einstein Studies (Birkhäuser, Boston, 1995).
3. J. Bernstein, *An Introduction to Cosmology* (Prentice–Hall, Englewood Cliffs, NJ, 1995).
4. T. H. Boyer, *Phys. Rev. D* **29**, 1089-1098 (1984).
5. C. H. Brans, *Phys. Rev.* **125**, 388–396 (1962).
6. S. Chen, A. Maksimchuk, and D. Umstadter, *Nature* **396**, 653–655 (1998).
7. M. Chown, *New Sci.* **3**, 29–32 (1999).
8. I. Ciufolini, E. Pavlis, F. Chieppa, E. Fernandes-Vieira, and J. Perez-Mercader, *Science* **279**, 2100–2103 (1998).
9. G. Cocconi and E. E. Salpeter, *Nuovo Cimento* **10**, 646–651 (1958).
10. J. G. Cramer, *Rev. Mod. Phys.* **58**, 647–687 (1986).

11. R. H. Dicke, *The Theoretical Significance of Experimental Relativity*, Documents on Modern Physics Series (Gordon & Breach, New York, 1964), esp. pp. 14–22, 31–34, 72–77.
12. A. Einstein, *The Meaning of Relativity*, 5th edn. (Princeton University Press, Princeton, NJ, 1956), pp. 99–108.
13. G. F. R. Ellis and D. W. Sciama, in *General Relativity: Papers in Honour of J. L. Synge*, O'Raifeartaigh, ed. (Clarendon, Oxford), pp. 35–59.
14. R. C. Gilman, *Phys. Rev. D* **2**, 1400–1410 (1970).
15. B. Haisch and A. Rueda, *J. Sci. Explor.* **11**, 473–485 (1997).
16. a. B. Haisch, A. Rueda, and H. Puthoff, *Phys. Rev. A* **49**, 678–694 (1994). b. Presentation at the 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, AIAA paper 98-3143 (1998).
17. C. Hoefer, in Ref. 2, pp. 67–90.
18. R. Matthews, *Science* **263**, 612–613 (1994).
19. P. Milonni, *The Quantum Vacuum* (Academic Press, New York, 1993), *passim*.
20. K. Nordtvedt, *Int. J. Theor. Phys.* **27**, 1395–1404 (1988).
21. P. C. Peters, *Am. J. Phys.* **49**, 564–569 (1981).
22. D. J. Raine, *Rep. Prog. Phys.* **44**, 1151–1195 (1981).
23. W. Rindler, *Introduction to Special Relativity*, 2nd edn. (Clarendon, Oxford, 1991).
24. a. A. Rueda and B. Haisch, *Phys. Lett. A* **240**, 115–126 (1998). b. *Found. Phys.* **28**, 1057 (1998).
25. a. D. W. Sciama, *Mon. Not. Roy. Astron. Soc.* **113**, 34–42 (1953). b. *Rev. Mod. Phys.* **36**, 463–469 (1964).
26. D. W. Sciama, P. C. Waylen, and R. C. Gilman, *Phys. Rev.* **187**, 1762–1766 (1969).
27. H.-J. Treder *et al.*, *Fundamental Principles of General Relativity Theories* (Plenum, New York, 1980).
28. J. F. Woodward, *Found. Phys. Lett.* **8**, 1–39 (1995); **9**, 1–23 (1996).