International Journal of Theoretical Physics, Vol. 21, Nos. 10/11, 1982

Relativity and Quantum Mechanics[1]

Hüseyin Yilmaz

Applied and Basic Advanced Technology, Hamamatsu TV, 105 Church Street, Winchester, Massachusetts 01890

Received October 14, 1981

Conditions under which quantum mechanics can be made compatible with the curved space-time of gravitation theories is investigated. A postulate is imposed in the form $v = v_g$ where v is the kinematical Hamilton–Jacobi (geometric optic limit) velocity and v_g is the group velocity of the waves. This imposes a severe condition on the possible coordinates in which the Schrödinger form (the coordinate realization) of quantum mechanics can be set up for purposes of calculating observable effects. Some such effects are calculated for a class of theories and are compared with experiments.

1. INTRODUCTION

Today, as we prepare to celebrate the eightieth birthday of Professor Dirac, we also look back on fifty-five years of the glorious work he has done in diverse fields of physics: quantum mechanics, quantum statistics, theory of radiation, electron theory, quantum electrodynamics, cosmology, and gravitation. During a stay at the Institute for Advanced Study, Professor Dirac enjoyed cutting Oppenheimer's trees and taking long walks. I accompanied him on many of his excursions and we had frequent discussions on physics, biology, evolution, cosmology, and how to solve intricate puzzles. The following work is dedicated to Professor Dirac and the indelible memories of his kindness.

My subject is the relationship between quantum mechanics and gravity. The work is based on a revision of Einstein's theory such that the right-hand side of the field equations contains, beside the usual matter term, the stress-energy of the gravitational field itself. Interestingly, the presence of

[1]Presented at the Dirac Symposium, Loyola University, New Orleans, May 1981.

0020-7748/82/1000-0871$03.00/0

the field stress-energy is not motivated directly as a requirement. Rather, for entirely different reasons, the metric tensor $g_{\mu\nu}$ is required to satisfy certain conditions and these conditions determine the $g_{\mu\nu}$ *without* the help of any field equations. Then, when one asks what type of field equations are satisfied by such $g_{\mu\nu}$, it becomes possible to prove that a field stress-energy t_μ^ν is required which adds to the matter part τ_μ^ν and has exactly those properties that are expected from a gravitational field stress-energy. The motivations for constructing the $g_{\mu\nu}$ in the special form mentioned include the desire to formulate the theory so as to satisfy a strong principle of equivalence compatible with quantum mechanics.

Most interesting is the tightness and the inner consistency of the mathematical framework. In this theory the mathematical framework is simple and helps to clarify the physical interpretation of the theory. In this respect it is in accord with Professor Dirac's long-standing philosophy of physical theory in which he advocated the adoption of physical interpretation to the mathematical structure rather than the mathematical structure to the physical interpretation. His prime example for this philosophy was quantum mechanics. If we are right, the theory of gravitation provides a second example of the basic strength of this philosophy. Interestingly, the framework also meets Professor Dirac's demand for clear and simple mathematics in the formulation of a physical theory (Appendix A).

We start with extremely simple arguments having to do with the formulation of the principle of equivalence and, step by step, lead to the motivations of imposing certain conditions on $g_{\mu\nu}$. These conditions are *classically* desirable on grounds of dynamical aspects of the principle of equivalence, and also *quantum mechanically* on grounds of the equality of phase shifts for inertial and gravitational mass. They essentially determine the form of $g_{\mu\nu}$, and via $g_{\mu\nu}$, the form of the field equations. The field equations so obtained imply a *field-geometry equivalence* because the equations of motion are then obtainable from the t_μ^ν via the usual field-theory procedure that its divergence is the volume-force acting on a particle, and this leads exactly to the geodesic equations of motion of the $g_{\mu\nu}$. This demonstration completes the first part of the exposition.

In the second part it is demonstrated that there exists a parametric link between this theory and the usual theory of gravity such that if one writes $T_\mu^\nu = \tau_\mu^\nu + \lambda t_\mu^\nu$, where T_μ^ν is the total stress-energy, then the conventional theory and the present theory correspond to the special cases $\lambda = 0$ and $\lambda = 1$, respectively. The generalized (λ-dependent) equations are mathematically well defined and possess exact solutions for arbitrary λ. Since only one value of λ can correspond to a definite theory, we then propose to investigate the λ-dependent equations so as to determine the value of λ from arguments of a physical nature. Two kinds of arguments are used: Theoretical and experimental. They both lead to $\lambda = 1$. The theoretical arguments have to do with the *principle of equivalence* and the *uniqueness of the Hamiltonian*. They require the value of λ to be unity. The experimental arguments have to do with the empirical predictions of the λ-dependent equations and become clearest in the case of two *quantum gravity* experiments, namely, the Colella, Overhauser, and Werner experiment on the gravitational phase shift of a neutron, and the Hughes–Drever experiment on the absence of mass-anisotropy in a gravitational field. These two experiments are shown to independently confirm the value of λ to be unity to within their respective accuracy. The paper ends with a brief summary of the main results and some appendices to balance the brevity of the text.

2. THEORETICAL CONSIDERATIONS

2.1. Principle of Equivalence. The new theory formulates *dynamics* in such a way that everything depends only on *potential differences*. The argument leading to this construction is very simple: let the metric $g_{\mu\nu}$ be a function $g_{\mu\nu}(\tilde{\phi}, \phi)$ of a gravitational field tensor potential $\tilde{\phi} = (\phi_\mu^\nu)$ and its trace $\phi_\mu^\mu \equiv \phi$. Also, let $\tilde{\phi}$ and therefore ϕ be solutions of the covariant d'Alembert equations of the same space $g_{\mu\nu}$. We then do two things: (1) We interpret $\tilde{\phi}$ as covariant generalization of the usual Newtonian potentials. (2) Since the solutions will have integration constants, we write the metric as $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$, implying the $g_{\mu\nu}$ has a group property (a kind of gauge invariance) whereby *only* the potential differences are of physical significance.

How would one determine the nature of the group property? This is something which cannot be guessed ahead of time. We must study some empirical data and their interpretations in current theories so that we may use them as correspondence arguments. Consider the principle of equivalence as manifested in the gravitational red shift: This effect is always understood as $\nu' = \nu[g_{00}(r)/g_{00}(r')]^{1/2}$ and has recently been tested (Vessot & Levine, 1981), in this form, to a high accuracy of a few parts in 10^6. This suggests that g_{00} is *multiplicative* and may be an exponential, therefore. The question is, of what quantity? If now we look at the field equations of general relativity ($R_\mu^\nu = 0$), one of them, $R_0^0 = 0$, is equal to 2 times the general Laplacian (Eddington 1957) of $\phi = -\frac{1}{2}\log(g_{00})$. Therefore g_{00} can be written as $g_{00} = e^{-2\phi}$, where ϕ is interpretable as a potential. Theoretically the red shift formula takes the form $\nu' = \nu e^{-(\phi - \phi')}$. Using this form as a correspondence argument, the general form can then be viewed as exponential for all $g_{\mu\nu}$. Clearly the usual theory does not have this more general form.[2] However, it is interesting to note that Einstein himself might have fully approved such an interpretation. In a 1907 article, recently translated by Schwartz (1977), he states that time dilation must "in all strictness" be an exponential, $t' = te^{\gamma\xi/c^2}$, where $-\gamma\xi/c^2$ is a special case of $\phi - \phi'$. In his subsequent research Einstein seems to have forgotten this important requirement, which is here restated and carried to its logical conclusion.

2.2. Determination of the Metric. The next step of enforcing such a general form might appear virtually impossible (because of all the worries of nonlinear equations and arbitrary coordinates), but in fact it turns out to be quite simple: Let $g_{\mu\nu}$ be a matrix exponential having the form $g = e^{\Omega^T}\tilde{\eta}e^{\Omega}$, where $2\Omega = \tilde{A}\phi + B\tilde{\phi}$, $\tilde{\eta}$ is the special relativity limit $g_{\mu\nu} \to \eta_{\mu\nu}$, $\tilde{A} = \tilde{1}A$ is a scalar matrix, and A and B are constants to be determined. We interpret the static field as $\phi_\mu^\nu \to \phi_0^0$, all other components vanishing. Then $\phi = \phi_\mu^\mu = \phi_0^0$, so in this case $g_{00} = e^{(A+B)\phi}$, $-g_{ii} = e^{A\phi}$. Correspondence to the red shift already gives $A + B = -2$, and hence the only remaining task is to determine A or B. For this we need a new condition. This new condition can be an experimental fact or a theoretical argument of consistency. Both avenues lead to the same conclusion: (1) Experimentally we have the bending of light rays, which implies $A = 2$. (2) Theoretically the time dilation (red shift) has the counterpart of length contraction. Since the group property required by the red shift is basic, we must accord a similar group property to length contraction, because a *basic requirement* of relativistic theory is the symmetric treatment of space and time variables. This implies that $dx' = dx\, e^{\phi-\phi'}$ from which we get $A = 2$ as before. Therefore, the exponential (Yilmaz, 1974, 1979)

$$\tilde{g} = e^{(\phi-2\tilde{\phi})^T}\tilde{\eta}e^{(\phi-2\tilde{\phi})} \tag{1}$$

may be regarded as a solution to the problem of the metric. (Note that in first order this exponential gives the linearized Einstein metric.)[3] We already know the static limit by correspondence. So the static line element is

$$ds^2 = e^{-2\phi}dt^2 - e^{2\phi}(dx^2 + dy^2 + dz^2)$$

where $\phi = \Phi/c^2 = GM/c^2r$ is the solution of the Laplace equation of the

[2] For example in nonisotropic Schwarzchild metric where $-g_{rr} = e^{\eta}$ one of the field equations ($T_0^0 = 0$) is $e^{-\eta}(\eta'/r - 1/r^2) + 1/r^2 = 0$, which does not allow $\eta \to \eta - K$.

[3] Thus, the linearized Einstein metric could be used as a correspondence condition to determine A and B.

same line element. This line element is known to be in agreement with all experiments having to do with a static field. Likewise, when effects having to do with other components are considered (first and second order in $\tilde{\phi}$), the general metric is again found to be in complete agreement with experiments. Only two questions were ever raised about this theory (Will, 1974), and both are due to simple misunderstanding: In the first instance the scalar ϕ was understood as the whole field present, whereas it is only the scalar trace of the more general field $\tilde{\phi}$. In the second instance a conjecture concerning the most general form of the solution to the field equations was taken to be an essential part of the theory, whereas it only means that the most general solution of the field equations is not known. The conjecture itself is never proven or disproven because a prescription of how to expand the metric beyond second order (when certain terms do not commute) is missing. The latter is only a technical problem which, by the way, the usual theory also has. (See Parameter Extension, Section 3.1.)

Completion of this theory into a mathematical framework is, again, very simple. The theory is given by just three equations: equation (1) and

$$\Box^2 \phi_\mu^\nu = 4\pi\sigma u_\mu u^\nu \tag{2}$$

$$\frac{du}{ds}\mu = \frac{1}{2}\partial_\mu g_{\alpha\beta} u^\alpha u^\beta \tag{3}$$

To compare with Einstein's theory, we compute

$$R_\mu^\nu - \tfrac{1}{2}\delta_\mu^\nu R = 2\left(\Box^2 \phi_\mu^\nu + t_\mu^\nu\right) \tag{4}$$

from the metric by computing its left-hand side. In all cases we have studied where the curvature is obtainable from the exponential in closed form, t_μ^ν is exactly the standard field stress-energy of $\tilde{\phi}$. Since such a t_μ^ν is interpretable as a *field stress-energy*, the remaining part, $\Box^2\phi_\mu^\nu$, the covariant d'Alembertian of $\tilde{\phi}$ with respect to $g_{\mu\nu}$, identifies the matter stress-energy, which is equation (2). This equation is as in other field theories of physics where the fields are related to the source. In this case the source is $\tau_\mu^\nu \rightarrow \sigma u_\mu u^\nu$. When $\tau_\mu^\nu = 0$ the field satisfies a d'Alembert equation, which is our original assumption. Equation (3) gives the geodesic equations of motion, which depend only on $g_{\mu\nu}$. Thus we can put (2) into (4), which become our (geometric) field equations. There is little need to solve them because we already have the solutions we need. Therefore the only thing we have to do is to form the geodesic equations and calculate the motions of particles. It is further found (Yilmaz, 1978)[4] (in all the cases we have computed in closed

[4]In Yilmaz (1978) and in Yilmaz (1979; 1980) p_μ' is to be corrected as $p_\mu' = p_\mu + K_\mu$ and $K_\mu = 0$ as in the present paper. For simplicity a 4π is sometimes omitted.

form) that one has $D_\nu t^\nu_\mu = \sigma\,\partial g_\mu g_{\alpha\beta} u^\alpha u^\beta/2$, which is σ times the right-hand side of the geodesic equations of motion. This completes the *field-theory interpretation*, because the equations of motion can now be written

$$\sigma \frac{du_\mu}{ds} = \frac{D_\nu t^\nu_\mu}{4\pi}$$

which is exactly as in field theory. As will be seen in Appendix B, the equations of motion then lead to the conservation laws of energy-momentum $\Sigma_k(m_{0k} u_{\mu k}) = C_\mu$. Note that, in the field-theory form above, masses must not be dropped out of the equations of motion (even when $m_i = m_g$), as their presence is needed for the total conservation laws and for the quantum mechanical phase shifts.

2.3. Basic Consequences of the Theory. The net result of this simple theory of gravitation is that the structure of space-time physics is turned into a standard field theory of spin-2 particles, although in a space curved partly by its own stress-energy. Several of its desirable features are worthy of special notice. We mention a few below that are easily provable on the simple static metric:

(1) A particle with zero rest mass has a unique signal velocity $v = V = v_g = c(x) = ce^{-2\phi}$ for *all wavelengths* (Yilmaz, 1980; 1977); v, V, v_g being the particle, phase, and group velocities. This allows a unique operational procedure of space-time measurements for both waves and particles.

(2) With $\phi \to \phi - \phi'$ the metric is scaled and the scaled metric is still a solution of the original field equations. Thus $x \to x'$, where x' is the point of observation, leads to a "local" Lorentz metric $g_{\mu\nu} \to \eta_{\mu\nu}$, $c(x) \to c$, although the frame is still noninertial (not freely falling). Thus in this theory the *observed value* of the velocity of light is always c as in special relativity.

(3) For a particle with nonzero rest mass one has $v = v_g$, $vV = c^2(x) \to c^2$, also as in special relativity. This permits a generalization of quantum mechanics of particles from flat to curved space-times and thereby allows a detailed analysis of some *quantum gravity* experiments (Yilmaz, 1980; 1977).

(4) With gauge $c(x) \to c$ the t^ν_μ turns out to be a pure tensor, not eliminable with any coordinate transformation consistent with that gauge, and reduces *exactly* to the Newtonian field stress-energy of a static gravitational field.

(5) In the strong-field limit the theory does not possess an event horizon (Yilmaz, 1975; 1979) and does *not* lead to a black hole behavior as can also be seen from the refractive index analogy $n = e^{2\phi}$. Radially directed light will always escape (red shifted).

Thus the theory is compatible with a strong principle of equivalence, the wave–particle duality of quantum mechanics ($v = v_g \leftrightarrow$ probability postulate), the symmetry of space-time variables, gauge, theory,[5] operational procedure of space-time measurements, and a local field-theory interpretation of space-time geometry. Because of these, a synthesis of space-time with quantum mechanics seems to be possible. Although the emphasis here is on the relation between gravity and ordinary quantum mechanics only, we believe that the theory will also permit a reasonable field quantization and provide a more general connection to the rest of physics than presently available.

3. EXPERIMENTAL CONSIDERATIONS

3.1. Parametric Extension. The above completes the first part of our discussions where we have formulated the theory on purely theoretical grounds. In the second part we will direct our efforts toward an experimental assessment of this theory vis-à-vis the usual theory of gravitation. Since the two theories are very different, yet possess essential similarities, a reasonable way to compare them would be to establish a parametric link covering both, and test the value of the connecting parameter. We will show that there indeed exists an interesting extension in terms of a parameter λ such that the two theories correspond to two special values of λ, namely, $\lambda = 0$ for the usual theory, and $\lambda = 1$ for the new theory. Furthermore, the parameter λ is so introduced that, for any value of λ, the geodesic equations of motion are a consequence of the field equations as in the usual theory of gravitation. In other words, in a *kinematical* sense any value of λ satisfies the principle of equivalence, that is, mass drops out of both sides of the equations of motion. However, when one investigates the *dynamical* aspects, for example, the quantum mechanical phase shifts, the value of λ makes a crucial difference. For, although *kinematically*, any mass drops out of the geodesic equations of motion $Du_\mu = 0$, the magnitude (and the structure) of the Hamiltonian nevertheless *dynamically* depends on the value of λ. It is therefore possible to calculate those effects which depend on λ and compare them with various experiments. It is shown that experiments available at the time of this writing constrain λ to be close to unity. This demonstration constitutes the main thrust of the second part of the paper. It is emphasized, further, that $\lambda = 1$ also follows from fundamental theoretical arguments based on the principle of equivalence and the uniqueness of the Hamilto-

[5] The existence of a local group $g_{\mu\nu}$ as an extension of the global $\eta_{\mu\nu}$ of special relativity is, of course, in the spirit of gauge theory.

nian. The discussions that follow should therefore be considered not as an empirical determination of λ but as an observational confirmation of $\lambda = 1$.

We introduce the parametric extension by the following three equations:

$$R_\mu^\nu - \tfrac{1}{2}\delta_\mu^\nu R = 8\pi\left(\tau_\mu^\nu + \lambda t_\mu^\nu/4\pi\right) \tag{1'}$$

$$\Box^2\phi_\mu^\nu = 4\pi\tau_\mu^\nu + (\lambda - 1)t_\mu^\nu \tag{2'}$$

$$\frac{du_\mu}{ds} = \tfrac{1}{2}\partial_\mu g_{\alpha\beta}u^\alpha u^\beta \tag{3'}$$

When $\lambda = 1$, these equations are equivalent to (1)–(3) of the new theory. When $\lambda = 0$, equations (1′) and (3′) are those of the conventional theory (Tupper, 1974).[6] In this case equation (2′) can either be ignored, or kept as a field-theory transcription of Einstein's theory.[7] In fact equation (2′) with $\lambda = 0$ reduces to the Gupta, the Feynman, Thirring, and Weinberg interpretation of Einstein's theory.[8] For arbitrary λ it is again found that

$$D_\nu t_\mu^\nu/4\pi = \tfrac{1}{2}\partial_\mu g_{\alpha\beta}\tau^{\alpha\beta}$$

hence the geodesic equations of motion are consequences of the field equations. In other words the λ-dependent extension is so constructed that for $\lambda = 0$ and $\lambda = 1$ it reduces to the usual and present theories, respectively, and for any value of λ the geodesic equations of motion follow from the extended field equations under a current conservation $\partial_\nu[(-g)^{1/2}\sigma(u_\mu)^\lambda u^\nu] = 0$.

We now show that, in the extended form, λ is the ratio of "active gravitational" mass m_0' to the "inertial" mass m_0

$$\lambda = m_0'/m_0$$

To show this we first define the inertial energy-momentum of a particle. Let the inertial stress-energy be $\tau_\mu^\nu = \sigma u_\mu u^\nu$. Integrating over a small volume containing the particle one has

$$p_\mu = \int(-g)^{1/2}\tau_\mu^\nu\, dV_\nu = m_0 u_\mu$$

[6]Tupper (1974) noted that the "final arbitrariness" in the field equations resides only in the value of λ.

[7]Some authors introduce a $\bar{t}_\mu^\nu = -t_\mu^\nu$ and interpret Einstein's theory in the form $\Box^2\phi_\mu^\nu = 4\pi\tau_\mu^\nu + \bar{t}_\mu^\nu$ $(\lambda = 0)$. In our theory $(\lambda = 1)$ the nonlinear $\bar{t}_\mu^\nu$ does not appear in the equation (2).

[8]Ohanian (1976) gives a systematic development based on Gupta, Feynman, Thirring, and Weinberg papers. See Introduction and p. 103.

Here σ is assumed to be a number of mass concentrations $\sigma \to \Sigma(-g)^{-1/2} m_{0j} \delta(x - x_j)$. By looking at equation (2′) the gravitational energy-momentum p'_μ would be equal to p_μ if $\lambda = 1$. When $\lambda \neq 1$, p'_μ is of the form

$$p'_\mu = p_\mu + (\lambda - 1) \int (-g)^{1/2} t^\nu_\mu \, dV_\nu$$

Using the Gaussian substitution $dV_\nu = \partial_\nu \, d\Omega$, and the relation $\partial_\nu[(-g)^{1/2} t^\nu_\mu] = (-g)^{1/2} D_\nu t^\nu_\mu = (-g)^{1/2} \sigma \, du_\mu / ds$ (Appendix A), the integral is

$$\int \partial_\nu \left[(-g)^{1/2} t^\nu_\mu \right] d\Omega = \int \int (-g)^{1/2} \sigma \, du_\mu \, d\Omega / ds$$

Setting $d\Omega / ds = u^\nu \, dV_\nu$ and integrating over du_μ we have

$$\int (-g)^{1/2} t^\nu_\mu \, dV_\nu = \int (-g)^{1/2} \tau^\nu_\mu \, dV_\nu = p_\mu$$

The relation between p'_μ and p_μ therefore is

$$p'_\mu = \lambda p_\mu$$

There can in principle be a numerical integration constant K_μ causing p'_μ to become $p'_\mu \to p'_\mu - K_\mu$ due to dV integral, but this is zero since p_μ and p'_μ refer to the same particle, and since, given the initial values x^μ, u_μ they must both trace the same geodesic $Du_\mu = 0$. Thus setting $p_\mu = m_0 u_\mu$, $p'_\mu = m'_0 u_\mu$ the relation $\lambda = m'_0 / m_0$ immediately follows. This result may appear a little unexpected but it is correct. (It can be rederived by using other methods as shown in Appendix B.) It is just a consequence of the condition that the field equations lead to the geodesic equations of motion for any value of λ. Then for $\lambda = 1$ the *inertial mass* and *active gravitational mass* reduce to each other. The existence of a field-geometry equivalence was first noticed by Feynman (1971), who studied it in the context of the conventional theory. It turns out, however, that unless also $\lambda = 1$, the equivalence does not extend to dynamics ($p'_\mu = p_\mu$), that is, it does *not* guarantee the equality of active and passive gravitational mass.

3.2. Experimental Test of λ. The relation $\lambda = m'_0 / m_0$ shows that, although for any λ the principle of equivalence is satisfied in the sense of "*passive*" gravitational mass equals "inertial" mass (both $Dp_\mu = 0$ and $Dp'_\mu = 0$ imply the same geodesic equations of motion $Du_\mu = 0$), it is nevertheless not satisfied in the sense of "*active*" gravitational mass equals "inertial" mass. Consequently, the two energy momenta p_μ and p'_μ and therefore the two Hamiltonians $H = p_0$ and $H' = \lambda p_0$ differ by a factor λ. Furthermore, each Hamiltonian depends, through $u_0 = u_0(\lambda)$, on λ. One can therefore compute various λ-dependent effects and (if there are any that are measurably sensitive to λ), test the value of λ by experiment.

There are two quantum gravity experiments which are sufficiently accurate to serve this purpose: The gravitational phase shift of a neutron and the isotropy of mass in a static gravitational field. Before we go into detail we must here make some comments as to the nature of the term in the Hamiltonian referring to these experiments. The quantity t^ν_μ is second order in $\phi = GM/rc^2$, hence the metric $g_{\mu\nu}$ and therefore $u_0(\lambda)$ is not influenced by λ in first order of ϕ. Therefore, in the first-order Newtonian limit the two Hamiltonians are given by (constants m_0c^2, $\lambda m_0 c^2$ are omitted)

$$H = p^2/2m_0 - m_0\phi, \qquad H' = \lambda(p^2/2m_0 - m_0\phi)$$

The neutron experiment is first order in ϕ and will be analyzed in view of these two expressions.

The isotropy experiment is sensitive to a second-order term in H. In this case the relevant part of the Hamiltonians H and H' is a perturbative anisotropic term A

$$\Delta H = (\lambda - 1)\phi^2 A, \qquad \Delta H' = \lambda(\lambda - 1)\phi^2 A$$

where $A = (2/9)(p^2/2m_0)P_2(\cos\alpha)$. The isotropy experiment will be analyzed in terms of these expressions. Here it is important to clarify how the anisotropic term is extracted and how it relates to quantum mechanics:

(a) The term is obtained consistently with the quantum mechanical demand that the kinematical velocity v of a particle is equal to the group velocity v_g of the associated wave. Consequently the above expressions are consistent with the probability postulate of quantum mechanics.

(b) When $\lambda \neq 1$ the metric cannot be written as[9] $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$ although it can as $g_{\mu\nu}(\tilde{\phi}, \phi)$. This is so because, when $\lambda \neq 1$, the form $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$ no longer satisfies the field equations with appropriate boundary conditions, so it cannot be used as a metric. In other words, for $\lambda \neq 1$ the potentials $\tilde{\phi}, \phi$ become locally observable quantities, that is, they would be detectable by a sufficiently accurate experiment if $\lambda \neq 1$ (Appendix C).

We now consider the neutron phase shift experiment (Colella et al., 1975). This experiment tests the applicability of quantum mechanics in a

[9]See for example footnote 1 above.

gravitational field in the case of phase shift and, via the phase shift, the consistency of quantum mechanics with the strong principle of equivalence. The experiment yields a phase shift

$$\delta = -m_0 \Delta\phi t/h$$

where t is the time of flight of neutron between two coincidences in a coherent beam. The accuracy of the experiment is about 1% and to this accuracy it is consistent with H and H' only if $|\lambda - 1| < 1\%$. Note that this test of the principle of equivalence is fundamentally different from other tests which are concerned with the mass independence of the equations of motion only. Before the result of the phase shift experiment was known there was a worrisome uncertainty because the weak principle alone could not lead to a definite prediction (Greenberger & Overhauser, 1980; Greenberger, 1968).

The above interpretation of the neutron experiment assumes that both H and H' are equally legitimate (principle of equivalence). Some colleagues seem to argue that if H were regarded as more legitimate than H' (I would not know how to justify such an assumption), then the neutron experiment does not necessarily imply $\lambda = 1$. Although we do not agree with such an interpretation (see Appendix B), we nevertheless point out that, even if one is willing to so ignore H', the isotropy experiment can be used to infer $\lambda = 1$ via H alone.

The isotropy experiment is accurate to about 5×10^{-23} and to this accuracy yields a *null* result. Calculations show that the experiment would have yielded a positive effect if in the anisotropic term (Appendix C) of H

$$\Delta H = \frac{2}{9}(\lambda - 1)\phi^2 \left(\frac{p^2}{2m_0} \right) P_2(\cos\alpha)$$

the factor $\lambda - 1$ were greater than 5×10^{-3}. The experiment was performed by two independent groups and two different settings and no detectable violation of the isotropy of inertia (energy) was observed (Hughes, 1964; Drever, 1961). The $j = 3/2$ Zeemann terms in H were, however, found implying $\lambda \neq 0$. We conclude that the two experiments here considered independently imply $|\lambda - 1| < \%1$ confirming the strong principle of equivalence.[10]

4. THEORETICAL CONSIDERATIONS ON λ

From a theoretical point of view the problem of λ is, of course, quite clear cut. Unless $\lambda = 1$ the Hamiltonian is not well defined, and therefore, to

[10] Note that $H' = \lambda H$ applies also in the classical case, hence any classical experiment having to do with the value of λ can also be used for test purposes.

achieve a *unique* Hamiltonian (energy), the value of λ must be set to unity. This is a basic physical requirement, both in classical and in quantum mechanics. As we have repeatedly emphasized, the principle of equivalence provides additional motivation for setting λ to unity. In a previous discussion we have seen that m_0 is the inertial mass, which is equal to "passive" gravitational mass on account of the geodesic equations of motion. The mass m'_0 is the "active" gravitational mass. The uniqueness of the Hamiltonian and the equality of "active" to "passive" gravitational mass are therefore related in the sense that one implies the other. What is surprising here is that this stronger form of the principle of equivalence is not satisfied unless $\lambda = 1$ although it was always presumed that the ordinary form of the principle of equivalence would guarantee the equality of "passive" and "active" gravitational mass.

One could carry this discussion a step deeper by noting that, in the Newtonian theory, the equality of passive to active gravitational mass is a consequence of the *momentum conservation laws*. In order to appreciate the nature of this fundamental connection let us study the *Newtonian limit* of a two-body system in our field theory form. We have

$$\sigma \frac{du_\mu}{dt} = \frac{\partial_\nu t_\mu^\nu}{4\pi}$$

where $\sigma = \sum_j m_j \delta(x - x_j)$ and

$$t_\mu^\nu = -\partial_\mu \phi \, \partial^\nu \phi + \tfrac{1}{2} \delta_\mu^\nu \partial^\lambda \phi \, \partial_\lambda \phi$$

$$\nabla^2 \phi = -\sigma'$$

with $\sigma' = \sum_j m'_j \delta(x - x_j)$. The total field is

$$\phi = \sum_j \left| \frac{m'_j}{x - x_j} \right|$$

The field-theory equations of motion are (after taking the divergence of t_μ^ν)

$$\sigma \frac{du}{ds} = \nabla^2 \phi \, \partial_\mu \phi$$

and these are equivalent to two *coupled equations* written once at m_1 and once at m_2:

$$m_1 \left(\frac{du_\mu}{dt} \right)_1 = + m'_1 \frac{m'_2}{|x_1 - x_2|^2}$$

$$m_2 \left(\frac{du_\mu}{dt} \right)_2 = - m'_2 \frac{m'_1}{|x_2 - x_1|^2}$$

This is so because at m_1, $\nabla^2\phi$ is equal to m'_1 and at m_2 it is equal to m'_2. From $\partial_\mu\phi$ one gets $m'_1/|x_1-x_2|^2$ at m_2. This is because the self terms $\partial_\mu(m'i/|x_i-x_i|)$ are zero in both cases (a particle is not under a force due to its own field). Now we can see two interesting things: One is the equality of "inertial" and "active" gravitational mass (due to $\nabla^2\phi=-\sigma'$ in $\partial_\nu t_\mu^\nu=\nabla^2\phi\,\partial_\mu\phi$ and the imposition of $\sigma'=\sigma$). The other is that if we add the two equations we get the conservation of momentum

$$\frac{d}{dt}(m_1u_{\mu 1}+m_2u_{\mu 2})=0$$

This example also clarifies how the concept of a *test particle* must be understood in a field theory such as gravitation. If m_1 is much smaller than m_2 one would think that its reaction on m_2 would be negligible. But the above analysis shows that this is not so. *No matter how small*, m_1 attracts m_2 with equal (and opposite) force as m_2 attracts m_1 (earth attracts the apple and the apple attracts the earth with equal and opposite force), and this is a consequence of total conservation of momentum. In this way we can see that the uniqueness of the Hamiltonian, the principle of equivalence, the field-theory representation of the equations of motion in terms of t_μ^ν, and the conservation laws are all related to the simple requirement that the "passive" gravitational mass is equal to the "active" gravitational mass, which is, in turn, equal to the "inertial" mass. Although this relation is here pointed out in the Newtonian limit, the relativistic theory must obviously reproduce this case, at least as a correspondence limit. Thus everything we know about gravitation, both experimental and theoretical, points to the principle of equivalence in the form $\lambda=1$, $p'_\mu=p_\mu$, hence also to the uniqueness of energy, $E_{ag}=E_{pg}=E_i=E$, as stated here.

Finally, the single most important consequence of the $\lambda=1$ theory might be its ability to generalize the mass–energy relation $E=mc^2$ from flat to curved space-times. This can be seen from the above discussion (e.g., $E_g=E_i=E$ when $\lambda=1$) and also from a very simple example as follows: Let a stationary particle with a gravitating mass m_g be displaced adiabatically in a gravitational field. The energy expended is $dE=-m_g\,d\Phi$. The inertial relation, $g^{\mu\nu}p_\mu p_\nu=m_0^2c^2$, on the other hand, gives $g^{00}(E/c)^2=m_0^2c^2$, hence by using the metric above $dE=-E\,d\Phi=-(E/c^2)\,d\Phi$. A comparison then shows $m_g=E/c^2$ which is an *exact* result (Yilmaz, 1973).[11] It is also unique to this theory, because in no other space-time theory has it been possible to generalize $E=mc^2$ as a rigorous basis of a strong principle of

[11] The same result follows from the red shift $\nu=\nu_0e^{-\phi}$, $dE=-(h\nu/c^2)\,d\phi$ since by the principle of equivalence $dE=-m_g\,d\Phi$, hence $m_g=h\nu/c^2=E/c^2$.

equivalence and tie it securely to quantum mechanics. That the new theory so establishes the general validity of $E = mc^2$ may be regarded one of its compelling advantages.

5. SUMMARY OF THE RESULTS

This paper addresses a basic and very complex issue, namely, the proper formulation of a space-time theory of gravitation vis-à-vis the principles of quantum mechanics. To achieve a comprehensible survey we have given, in the text, the basic arguments leading to the construction of the theory and its experimental confirmation. Where the need is perceived for further clarification we have referred to the appendices at the end of the paper or to some literature where such detail may be found. Still, because of the complexity and many-sidedness of the issue, the need exists for a brief recapitulation of the crucial points. The purpose of this section is to provide such a synopsis.

(1) If we are to pursue a theory of gravitation in the sense of a generalized principle of relativity, the absolute values of the gravitational potentials must be *locally unobservable*, that is, in any given frame of reference the physical effects must depend only on *potential differences* between the observer and the objects he observes. In combination with the red shift as a *correspondence* argument, and the relativistic requirement of *symmetry between space and time variables*, this commits us to an exponential line element with a multiplicative group property in terms of potentials. It is not absolutely essential to independently introduce $\tilde{\phi}$, ϕ because they can be eliminated by the substitutions $\phi = \frac{1}{4}\ln(-g)$, $\tilde{\phi} = \frac{1}{4}\ln[(-g)^{1/2}\tilde{g}^{-1}\tilde{\eta}]$. What is most important is that if the field equations are not chosen so as to satisfy the group property, then the absolute values of the potentials will become locally observable and will *explicitly* appear in the expressions of some classical or quantum mechanical effects.

(2) Having introduced $g_{\mu\nu}$ as a multiplicative group $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$, the integration constants $\tilde{\phi}', \phi'$ are interpretable as the potentials at the point of observation. This leads to $g_{\mu\nu} \to \eta_{\mu\nu}$ in the vicinity of the observer, that is, to a kinematics which is locally (observer not necessarily freely falling) special relativistic. For example the velocity of light, as defined by $ds^2 = 0$ would then be locally c as in special relativity. This allows the definition of an *operational procedure* of space-time measurements similar to special relativity. This alone is not, however, sufficient because $ds^2 = 0$ gives the velocity of light only in the sense of a particle. In order for the operational procedure to be applicable both for *waves* and *particles* (unity of waves and particles), the wave equation $\Box^2\chi = 0$ must also yield the same

velocity c.[12] This means that for a particle of zero rest mass the kinematical velocity v, the phase velocity V, and the group velocity v_g must be equal and locally reduce to c. The new theory does indeed satisfy this requirement and *preserves* the operational procedure implicit in special relativity *both for waves and particles* and for *all* wavelengths.

(3) In order to be able to set up the quantum mechanics in curved space one further needs to consider particles with nonzero rest mass and secure that the kinematical velocity v of the particle is equal to the group velocity v_g of the associated wave. The latter is necessary for the *probability interpretation* of quantum mechanics. Again the new theory meets this requirement since in this theory one has $v = v_g$, $v \cdot V = c^2(x) \to c^2$. With this condition satisfied it is possible to set up quantum mechanics of atomic and nuclear systems in a gravitational field and compare the predictions with experiments. We have seen that the experiments are in agreement with the theory.

(4) What happens when one or more of these postulates are not satisfied? We have examined a λ-dependent generalization of the theory so as to recover the present theory when $\lambda = 1$ and lead to the usual Einstein theory when $\lambda = 0$. The purpose of this exercise was to mathematically study how the predictions change as a function of λ and what basic principles of physics are violated when λ is not equal to unity. To do this as concretely as possible we have selected two quantum gravity experiments, one being the Colella–Overhouser–Werner experiment on neutron phase shift, and the other being the Hughes–Drever experiment on the isotropy of inertia. The importance of this choice is that both experiments are *quantum mechanical* in nature and both are *accurate enough* to imply a significant upper limit on $|\lambda - 1|$.[13]

(5) In the case of neutron phase shift what is violated when $\lambda \neq 1$ is the strong principle of equivalence. The inertial energy-momentum is $p_\mu = \int (-g)^{1/2} \sigma u_\mu \, dV$ but the gravitational energy momentum turns out to be $p'_\mu = \lambda p_\mu$. The two Hamiltonians $H = p_0$ and $H' = \lambda p_0$ differ by a factor λ. These both give the same geodesics (two Hamiltonians differing by a constant factor give the same equations of motion), yet because of λ they lead to different phase shifts. Thus, unless $\lambda = 1$ the Hamiltonian is nonunique. The uniqueness is restored by imposing the strong principle of equivalence, namely, the equality of the *passive* and *active gravitational mass*. The experiment confirms this result by showing that there is no Hamiltonian in which λ can differ significantly from unity.

[12] The equality of v_g and v is required by quantum mechanics in the coordinate (Schrödinger) representation to give $|\Psi|^2$ a probability interpretation.

[13] Actually there are physical reasons to believe that λ is a two-valued ($\lambda = 0, 1, \lambda^2 = \lambda$) variable. In this case the experiments imply λ exactly with a high level of confidence.

(6) The Hughes–Drever experiment is studied because one might wish to satisfy the principle of equivalence in its weak form only, and still try to cover all the experiments. In this case one would keep the inertial Hamiltonian H and ignore the gravitational Hamiltonian H'. Although such a procedure is highly unsatisfactory (see Appendix B) it nevertheless frustrates the $\lambda = 1$ implication of the neutron experiment. The Hughes–Drever experiment, however, shows that we still cannot escape the $\lambda = 1$ imposition because this experiment is sensitive to a second-order term in $H = m_0 u_0(\lambda)$ and, via that term, confirms the theoretical value $\lambda = 1$ on the basis of H alone.

ACKNOWLEDGMENTS

The author is grateful to Professors Eugene P. Wigner, Robert H. Dicke, Felix H. M. Villars, Arthur K. Kerman, Daniel Greenberger, Michael A. Horne, Mehmet Rona, and Alan Guth for discussions, and to Alexander P. Doohovskoy for verifying the solutions on MACSYMA (M.I.T. Project MAC Symbolic Manipulation Computer) and for a critical reading of the paper.

APPENDIX A: MATHEMATICAL STRUCTURE (OR PRETTY MATHEMATICS[14])

This appendix presents a mathematical framework which can be studied independent of any physical association and has a simple and elegant structure. Since the new theory is describable with this structure, by giving suitable interpretations to its equations, and since many of the statements of the text are easily provable through it, we devote this appendix to an exposition of this simple mathematics.

Consider a metric $g_{\mu\nu}$. We can conceive of a functional substitution, $g_{\mu\nu}(\tilde{\phi}, \phi)$ where $\tilde{\phi} = \phi^{\alpha}_{\beta}$ is a symmetric matrix and $\phi = \phi^{\mu}_{\mu}$ is its trace. Such a relation may be thought of as expressing $g_{\mu\nu}$ in terms of $\tilde{\phi} = \tilde{\phi}(x)$ instead of directly by x. Now consider the form

$$g_{\mu\nu} = \left(\tilde{\eta} e^{2(\phi - 2\tilde{\phi})} \right)_{\mu\nu}$$

where $\tilde{\eta}$ is the special relativistic limit $g_{\mu\nu} \to \eta_{\mu\nu}$ and ϕ is the diagonal matrix $\tilde{1}\phi$. One can construct special cases where the matrix expansions are known functions of ϕ^{ν}_{μ}, ϕ. In such cases the curvature quantities can easily be calculated. It turns out that all such cases follow a regular pattern in that

[14] This appendix is written in the spirit of Professor Dirac's opening lecture, "Pretty Mathematics."

they satisfy the equations

$$R_{\mu}^{\nu} - \tfrac{1}{2}\delta_{\mu}^{\nu}R = 2\left(\Box^{2}\phi_{\mu}^{\nu} + t_{\mu}^{\nu}\right) \tag{A.1}$$

$$\Box^{2}\phi_{\mu}^{\nu} = \tau_{\mu}^{\nu} \tag{A.2}$$

$$D_{\nu}t_{\mu}^{\nu} = \tfrac{1}{2}\partial_{\mu}g_{\alpha\beta}\tau^{\alpha\beta} \tag{A.3}$$

where $\Box^2$ is the general d'Alembertian and τ_{μ}^{ν} is introduced (for convenience) to represent $\Box^{2}\phi_{\mu}^{\nu}$.

Note that $g_{\mu\nu}(\tilde{\phi}, \phi)$ can be inverted as in

$$\phi = \tfrac{1}{4}\ln(-g), \qquad \tilde{\phi} = \tfrac{1}{4}\ln\left[(-g)^{1/2}\tilde{g}^{-1}\tilde{\eta}\right]$$

where g is the determinant of $g_{\mu\nu}$. This is easy to see from the original expression of $g_{\mu\nu}$ since $g = -\exp[\operatorname{tr} 2(\phi - 2\tilde{\phi})] = -e^{4\phi}$. The inversion above shows that the equations could also be written directly in terms of $g_{\mu\nu}$.[15]

Below we present two interesting solutions which can serve as existence proofs (by example) and as guides to physical interpretation. They can also give an idea of the elegant nature of the underlying mathematics:

Example 1: The Static Field. Let $\phi_{\mu}^{\nu} = \phi_{0}^{0}(x, y, z)$, all other components being zero. Then $\phi = \operatorname{tr}(\phi_{\mu}^{\nu}) = \phi_{0}^{0}$ so that the general exponential yields

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \exp 2\phi \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \exp 4\phi \begin{pmatrix} -1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}$$

$$g_{00} = e^{-2\phi}, \qquad -g_{ii} = e^{2\phi}$$

The line element therefore is

$$ds^2 = e^{-2\phi}dt^2 - e^{2\phi}(dx^2 + dy^2 + dz^2)$$

computing the d'Alembertian $\Box^{2}\phi = (-g)^{-1/2}\partial_{\nu}[(-g)^{1/2}g^{\mu\nu}\partial_{\nu}]\,\phi$ one has

$$\Box^{2}\phi = -e^{-2\phi}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi$$

[15] Since, for $\tau_{\mu}^{\nu} = 0$, the $\tilde{\phi}$, ϕ satisfy the d'Alembert equation of the same space, the $g_{\mu\nu}$ may be considered a function of its own representations.

Let $\tau_\mu^\nu = \tau_0^0 = \tau$ be $\tau = \Sigma_j (-g)^{-1/2} m_{0j} \delta(x - x_j)$. Since $(-g)^{1/2} = e^{2\phi}$ the solution is then of the form

$$\phi = \sum_j \frac{m_j}{|x - x_j|}$$

Introducing this into the metric and computing the left-hand side of equation (A.1) we get[16]

$$t_\mu^\nu = -\partial_\mu \phi \, \partial^\nu \phi + \tfrac{1}{2} \delta_\mu^\nu \, \partial^\lambda \phi \, \partial_\lambda \phi$$

For the assumed forms here considered

$$D_\nu t_\mu^\nu = \Box^2 \phi \partial_\mu \phi = \tfrac{1}{2} \partial_\mu g_{\alpha\beta} \tau^{\alpha\beta}$$

namely, it satisfies (A.3).

Example 2: The Gravity Waves. Let $\phi_\mu^\nu = \phi_1^1(t - z) = -\phi_2^2(t - z)$ all other components being zero. Let ξ denote ϕ_1^1 and note that $\phi = \phi_\mu^\mu = 0$. The general exponential now gives

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \exp 4\xi \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \end{pmatrix}$$

$$-g_{11} = e^{4\xi}, \qquad -g_{22} = e^{-4\xi}$$

all other components of $g_{\mu\nu}$ being as in $\eta_{\mu\nu}$. Similarly, if $\phi_\mu^\nu = \phi_1^2(t - z) = \phi_2^1(t - z)$, $\phi_1^2 = \zeta$, $\phi = 0$ then one gets

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \exp 4\zeta \begin{pmatrix} 0 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 0 \end{pmatrix}$$

$$g_{11} = g_{22} = -\cosh 4\zeta, \qquad g_{12} = g_{21} = -\sinh 4\zeta$$

[16] This is exactly the Newtonian field stress-energy which is here recovered as a second-order correspondence limit. The conventional theory uses only a first-order correspondence, hence misses the t_μ^ν. It also misses the field Lagrangian because the field Lagrangian, $L_\phi = -t/2$, that is, $-1/2$ times the trace of t_μ^ν.

These two solutions are not independent. They form a single transverse time-dependent solution

$$\Box^2\phi_\alpha^\beta = \left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2}\right)\phi_\alpha^\beta = 0$$

where ϕ_α^β is given by

$$\tilde{\phi} = \phi_\beta^\alpha = \begin{pmatrix} 0 & & & \\ & \phi_1^1 & \phi_1^2 & \\ & \phi_2^1 & \phi_2^2 & \\ & & & 0 \end{pmatrix}$$

Under a rotation operator $Ra_1 = a_2$, $Ra_2 = -a_1$ this wave behaves as $R^2\tilde{\phi} = -4\tilde{\phi}$ so that the eigenvalues of iR are ± 2. This shows that the wave corresponds to a spin-2 field.

The wave solutions ϕ_β^α above give

$$t_\mu^\nu = -2\left(\partial_\mu\phi_\beta^\alpha\,\partial^\nu\phi_\alpha^\beta - \tfrac{1}{2}\delta_\mu^\nu\,\partial\lambda\phi_\beta^\alpha\,\partial_\lambda\phi_\alpha^\beta\right)$$

$$D_\nu t_\mu^\nu = -2\Box^2\phi_\beta^\alpha\,\partial_\mu\phi_\alpha^\beta = \tfrac{1}{2}\partial_{\mu\alpha\beta}\tau^{\alpha\beta}$$

again satisfying all the equations.

General Case. Similar calculations, using all fields so constructable, give the *more general* relations[17]

$$t_\mu^\nu = -2\left(\partial_\mu\phi_\beta^\alpha\,\partial^\nu\phi_\alpha^\beta - \tfrac{1}{2}\delta_\mu^\nu\,\partial^\lambda\phi_\beta^\alpha\,\partial_\lambda\phi_\alpha^\beta\right) + \partial_\mu\phi\,\partial^\nu\phi - \tfrac{1}{2}\delta_\mu^\nu\,\partial^\lambda\phi\,\partial_\lambda\phi$$

$$D_\nu t_\mu^\nu = -2\Box^2\phi_\beta^\alpha\,\partial_\mu\phi_\alpha^\beta + \Box^2\phi\,\partial_\mu\phi = \tfrac{1}{2}\partial_\mu g_{\alpha\beta}\tau^{\alpha\beta}$$

showing that there are many solutions satisfying our equations. It has also been proven that, at least for all the exact solutions now known, the t_μ^ν satisfies the relation

$$D_\nu t_\mu^\nu = (-g)^{-1/2}\partial_\nu\left[(-g)^{1/2}t_\mu^\nu\right] = \tfrac{1}{2}\partial_\mu g_{\alpha\beta}\tau^{\alpha\beta}$$

This follows from the form of t_μ^ν as given above, because an extra term, $\Gamma_{\mu\beta}^\alpha t_\alpha^\beta$, vanishes.

[17]The solutions possess integration constants $\tilde{\phi}', \phi'$ such that $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$ is also a solution. In the most general case these may be conceived as infinite number of infinitesimal transformations.

This purely mathematical structure is suggestive of a physical interpretation as follows:

(1) The matter stress-energy is τ_μ^ν. It generates the gravitational field according to $\Box^2\phi_\mu^\nu = \tau_\mu^\nu$.

(2) The gravitational stress-energy is t_μ^ν. It provides the equations of motion $\sigma\, du_\mu/ds = D_\nu t_\mu^\nu$.

With this interpretation in mind one immediately identifies[18] the total stress-energy $T_\mu^\nu = \tau_\mu^\nu + t_\mu^\nu$, $D_\nu T_\mu^\nu \equiv 0$ and the geodesic equations of motion $du_\mu/ds = \frac{1}{2}\partial_\nu g_{\alpha\beta}u^\alpha u^\beta$.

Alternatively, the general expression of the curvature invariant is (rectangular coordinates will be used)

$$-R = \Box^2 \ln(-g)^{1/2} + \tfrac{1}{2}\Gamma^\mu_{\alpha\beta}\,\partial_\mu g^{\alpha\beta} - (-g)^{1/2}$$
$$\times\left[\partial_\mu \ln(-g)^{1/2} - \partial_\mu\right]\partial_\nu\left[(-g)^{1/2} g^{\mu\nu}\right]$$

Comparing with the result of our metric for the same quantity, namely, $-R = 2(\Box^2\phi + t)$, up to an irrelevant divergence, we have

$$\Box^2\phi = \sigma, \qquad (-g)^{1/2} = e^{2\phi}$$

The Lagrangian $L_\phi = -t/2$,

$$L_\phi = -\partial^\lambda\phi^\alpha_\beta\,\partial_\lambda\phi^\beta_\alpha + \tfrac{1}{2}\partial^\lambda\phi\,\partial_\lambda\phi$$

is represented by the second term, $-\frac{1}{8}\Gamma^\mu_{\alpha\beta}\,\partial_\mu g^{\alpha\beta}$, of $R/4$. The other terms (in the general expression of R) vanish if one imposes the constraints (Yilmaz, 1975; 1979)

$$\partial_\nu\left[(-g)^{1/2} g^{\mu\nu}\right] = 0$$

which can serve as subsidiary gauge conditions. They are satisfied by all of our exact solutions above.

It may be noted that the exponential line elements were originally obtained by us by using the combination of the principle of equivalence and local Lorentz covariance.[19] Again no field equations were used, indicating the preeminence of the principle of equivalence over arbitrarily motivated field equations.

[18]Note the way test particles are to be interpreted (Appendix B.)

[19]See Yilmaz (1976), Appendix B. In that paper the general metric is obtained in three different ways. When Fourier components, satisfying the gauge condition $k_\nu\phi_\mu^\nu = 0$, are used, the solutions may be regarded exact.

APPENDIX B: ALTERNATIVE DERIVATION OF $\lambda = m'_0/m_0$

In the text we have shown that the parametric extension covering the usual and the present theories of gravitation satisfies $\lambda = m'_0/m_0$. Here we provide an alternative proof based directly on the fundamental field equations (1′). The proof will also show how the divergence identities are related to the conservation laws of energy-momentum.

Let the field equations be written as

$$R^{\nu}_{\mu} - \tfrac{1}{2}\delta^{\nu}_{\mu}R = 8\pi T^{\nu}_{\mu}$$

Obviously T^{ν}_{μ} is used here in the sense of "active" gravitational stress-energy since it is the source of the geometric curvatures. On the other hand the condition $D_{\nu}T^{\nu}_{\mu} \equiv 0$ must be assumed as the basis of conservation laws of energy-momentum. Our problem is how to relate the active (gravitational) energy-momentum p'_{μ} to the total stress-energy T^{ν}_{μ}.

Here an analogy may help: Suppose we were interested in the electric charge. We have the divergence analog $D_{\nu}j^{\nu} \equiv 0$. Since for a vector j^{ν} the divergence is equal to $(-g)^{-1/2}\partial_{\nu}[(-g)^{1/2}j^{\nu}]$ one can write

$$\Delta Q = \int (-g)^{1/2} D_{\nu}j^{\nu}\,d\Omega = \int \partial_{\nu}\left[(-g)^{1/2}j^{\nu}\right]d\Omega = 0$$

$$Q = \int (-g)^{1/2} j^{\nu}\,dV_{\nu} = C$$

Since the second part of the first equation is in the form of ordinary divergence, the theorem of Gauss applies and, by a well-known method (Landau & Lifshitz, 1962; Dirac, 1975),[20] the conservation of charge $Q = \Sigma_k q_k = C$ is obtained from the condition that $D_{\nu}j^{\nu} \equiv 0$. This procedure is perfectly valid for $D_{\nu}T^{\nu}_{\mu} \equiv 0$, but one must take care of a technical matter, namely, Gauss's theorem applies only to *ordinary divergence* whereas $D_{\nu}T^{\nu}_{\mu}$ contains an extra term besides the ordinary divergence. To overcome this (mathematical) hurdle we let $T^{\nu}_{\mu} = \tau^{\nu}_{\mu} + \lambda t^{\nu}_{\mu}$ and write the covariant divergence of T^{ν}_{μ} as

$$(-g)^{1/2}D_{\nu}T^{\nu}_{\mu} = \partial_{\nu}\left[(-g)^{1/2}\tau^{\nu}_{\mu}\right] + (-g)^{1/2}\left(\lambda D_{\nu}t^{\nu}_{\mu} - \tfrac{1}{2}\partial_{\mu}g_{\alpha\beta}\tau^{\alpha\beta}\right) = 0$$

In analogy to the case of charge, we can now apply Gauss's theorem as

$$\Delta P'_{\mu} = \int (-g)^{1/2} D_{\nu}T^{\nu}_{\mu}\,d\Omega = 0$$

$$P'_{\mu} = \int \partial_{\nu}\left[(-g)^{1/2}\tau^{\nu}_{\mu}\right]d\Omega + (\lambda - 1)\int\int (-g)^{1/2}\sigma\,du_{\mu}\,d\Omega/ds = C'_{\mu}$$

[20] The method of deriving the conservation of charge from $D_{\nu}j^{\nu} \equiv 0$ appears in many places, but for $D_{\nu}T^{\nu}_{\mu} \equiv 0$ no detailed proof seems to be given. L. D. Landau did, however, point out how Gauss's theorem is to be applied. (See Note Added in Proof).

Using the substitution $d\Omega\,\partial_\nu = dV_\nu$ and setting $\int(-g)^{1/2}\tau_\mu^\nu\,dV_\nu = P_\mu$ the expression analogous to the conserved charge is

$$P'_\mu = P_\mu + (\lambda-1)\int\int(-g)^{1/2}\sigma\,du_\mu\,u^\nu\,dV_\nu$$

where in the second integral we have set $d\Omega/ds = u^\nu\,dV_\nu$. Integrating over du_μ the last integral is again P_μ. Thus we have

$$P'_\mu = \lambda P_\mu$$

where $P'_\mu = \Sigma_k m'_{0k} u_{\mu k}$ and is conserved by virtue of $D_\nu T_\mu^\nu \equiv 0$. In this derivation we have used $\tau_\mu^\nu = \sigma u_\mu u^\nu$, $\sigma = \Sigma(-g)^{-1/2} m_{0j}\delta(x-x_j)$. One also finds of course $P_\mu = \Sigma m_{0k} u_{\mu k}$ to be conserved. This is due to the fact that for all values of λ the geodesic equations of motion are valid as a consequence of the field equations so the inertial $p_\mu = m_0 u_\mu$ and gravitational $p'_\mu = m'_0 u_\mu$ obey the same equations. Conservation laws take the form

$$P'_\mu = \sum(m'_0 u_\mu) = C'_\mu = \lambda C_\mu$$

as an analog of charge conservation $Q = \Sigma q = C$. From the covariance of the procedure it is then evident that the energy-momentum of each particle is of the form $p'_\mu = \lambda p_\mu$, hence one has $\lambda = m'_0/m_0$ for all particles. The result is the same as in the text which used the integrations only over small volumes containing individual particles.

The present method reveals that, besides the relation $\lambda = m'_0/m_0$, the λ-dependent extension has the property of conserving the total energy-momentum $P'_\mu = \lambda P_\mu$ for any value of λ. Here we wish to emphasize two interesting points: One is that the appropriate energy-momentum tested by the neutron experiment is P'_μ and not necessarily P_μ. If we were to predict an experimental fact, and if we insisted the Hamiltonian to come out of the fundamental gravitational field equations, we would choose P'_μ because P_μ is not directly implied by the field equations unless $\lambda = 1$. This can be explicitly demonstrated as in section 3, where the $(\lambda - 1)t_{\mu\nu}$ in equation (2′) causes the Hamiltonian to become λH. The second is that the conservation laws are consistent with the equations of motion because, when all the particles are included, the field is the sum total of the fields of all the particles. It is easy to see this in the case of a static field where $\phi = \Sigma_j m_{0j}/|x-x_j|$. When the equations of motion are evaluated one has

$$m_{0k}\frac{du_{\mu k}}{ds} = m_{0k}\sum_{j\neq k}\frac{\partial}{\partial x_k^\mu}\frac{m_{0j}}{(x_k - x_j)}$$

and there are as many such equations as there are particles. Summing over all particles one has the conservation laws[21]

$$P_\mu = \sum_k m_{0k} u_{\mu k} = C_\mu$$

because the sum of the right-hand side of the equations of motion vanishes. In the new theory $\lambda = 1$ so that $D_\nu T_\mu^\nu \equiv 0 \to (-g)^{-1/2}\partial_\nu[(-g)^{1/2}\tau_\mu^\nu] = 0$ may be considered as the expression of total conservation laws. This is evident from $\int(-g)^{1/2} D_\nu T_\mu^\nu \, d\Omega \to \int \partial_\nu[(-g)^{1/2}\tau_\mu^\nu]\, d\Omega$ which is directly transformable by Gauss's theorem. That the equation $\partial_\nu[(-g)^{1/2}\tau_\mu^\nu] = 0$ represents an expression of the conservation laws was pointed out a long time ago by Landau and Lifshitz (1962).

In the full relativistic case there is of course also the energy-momentum radiation. Our present emphasis is, however, the intrinsic symmetry between "*test*" and "*source*" particles, namely, the concept of a test particle where the reaction (of test particle) on the rest of the system is not negligible but is equal and opposite to the action of the rest of system on the particle. The relativistic form of this concept and the phenomenon of radiation of particles with zero rest mass will be treated in a different communication.

APPENDIX C: THE ISOTROPY EXPERIMENT

In treating the isotropy experiment of Hughes and Drever two crucial points are the following: (1) Quantum mechanics requires $v_g = v$ and when this requirement is satisfied the Hamiltonian H possesses an anisotropic term which vanishes only when $\lambda = 1$. If $\lambda \neq 1$ it does not vanish. (2) It is not possible to argue that the anisotropic term would depend on potential differences $\tilde{\phi} - \tilde{\phi}'$, $\phi - \phi'$, because unless $\lambda = 1$ the metric does not exhibit a group property $g_{\mu\nu}(\tilde{\phi} - \tilde{\phi}', \phi - \phi')$. (When $\lambda \neq 1$ this form does not satisfy the field equations with appropriate boundary conditions and cannot be used as a metric.) Therefore when $\lambda \neq 1$ the absolute values of $\tilde{\phi}$ and ϕ become locally observable, leading to an experimental test of λ by virtue of the high precision of the Hughes–Drever experiment. The purpose of the present appendix is to explicitly show the above properties of the λ-dependent equations.

[21]Notice that masses cannot drop out of the equations even when $m_i = m_p = m_a$ as their presence is necessary for the global conservation laws. Quantum mechanics also tells us that mass cannot drop out; otherwise one cannot infer the phases uniquely.

The most effective way to deal with this problem is to solve the λ-dependent equations and explicitly display the Hamiltonian. The unique λ-dependent solution that satisfies the $v_g = v$ condition in spherical coordinates is (Yilmaz, 1980; 1977)

$$ds^2 = e^{-2(\phi-k)}\,dt - e^{2(\phi-k)}$$
$$\times\left[\left(\frac{\varepsilon\phi}{\sinh\varepsilon\phi}\right)^2(r^2\,d\theta^2 + r^2\sin^2\theta\,d\psi^2) + \left(\frac{\varepsilon\phi}{\sinh\varepsilon\phi}\right)^4 dr^2\right]$$
$$\phi = M/r, \qquad \lambda = 1-\varepsilon^2$$

(Note that for $\lambda = 0$ this is a transform of the Schwarzchild metric.) where $\phi = M/r$ is a solution of the equation $\Box^2\phi = 4\pi M\,\delta(r)$. One of the important points to note here is that this is the most general form satisfying $v = v_g$ and the boundary conditions at $r = \infty$.[22] The integration constant k appears in the exponents but *not* in the spatial relations inside the bracket. Consequently the quantity

$$\Gamma = \frac{c_r^2}{c_\theta^2} = \left(\frac{\sinh\varepsilon\phi}{\varepsilon\phi}\right)^2 \simeq 1 + \frac{\varepsilon^2}{3}\phi^2$$

depends on $\varepsilon\phi$ and *not* on $\phi - k = \phi - \phi'$. One can say that the reference frame is determined by the condition $v = v_g$ and that the physical Anisotropy as probed by the quantum mechanic is $\Gamma - 1$. The second important point is that letting $(\Box^2 + m_0^2)\chi = 0$, $\chi = r^{-1}f(t, r)$ one finds $v = v_g$. The latter is required by the probability interpretation of quantum mechanics. One can therefore set up quantum mechanics in this line element. The Hamiltonian $H = i\hbar D_0$ is given by (for $\lambda = 1$ see the last page)

$$H = e^{-(\phi-k)}\left[m_0^2 + e^{-2(\phi-k)}\Gamma(p^2\sin^2\alpha - \Gamma p^2\cos^2\alpha)\right]^{1/2}$$

where p is the instantaneous momentum (for detail see Yilmaz, 1980; 1977; Hughes, 1964; Drever, 1961). Locally, that is, when $k \simeq \phi$, this Hamiltonian gives (Yilmaz, 1975; 1977; 1979; 1980)

$$H = H_0 + \frac{2}{9}(\lambda-1)\phi^2\left(\frac{p^2}{2m_0}\right)P_2(\cos\alpha)$$

where H_0 is isotropic. The extra term is an anisotropic perturbation. This extra term depends on $\lambda - 1$ and ϕ^2. Although on the surface of the earth

[22] In general the metric does not exhibit a group property in the coefficient of $\eta_{\mu\nu}$. See the example of footnote 1 above.

$\phi^2 = 5\times 10^{-19}$, the anisotropic term (if it exists) can be detected in certain atoms or nuclei (^{7}Li with $j = 3/2$ was used) by applying a nuclear magnetic resonance technique. Such a technique is known to be accurate to a ramarkable precision of 5×10^{-23} (ratio of $\langle|\Delta H|\rangle$ to $T = p^2/2m_0 = 10$ MeV). Experiments show *no such anisotropy*, thus setting an upper limit $|\lambda - 1| < 5\times 10^{-3}$.

The theoretical treatment of the Hughes–Drever experiment shows that quantum mechanical requirement $v_g = v$ (probability postulate), in some sense, leads one to use selected systems of coordinates, namely, the ones in which the kinematical velocity v of a particle is equal to the group velocity v_g of its associated wave. To what extent this commits us to the concept of a privileged system of coordinates is not clear. Here we may assume that some sort of a commitment is involved, especially if we insist on the Schrödinger representation. The Schrödinger representation (that is, the representation in which position operators are ordinary numbers) is, after all, the *coordinate representation* of quantum mechanics and it would not be surprising if this representation required special coordinates in order to be consistent with the probability postulate. Such a requirement could in fact be regarded as a *selection process* as quantum mechanics plays a selective role over the classically possible dynamical states. On the problem of quantum mechanics and coordinates Professor Dirac once made the following remark (Dirac, 1976, footnote p. 114):

> This assumption is found in practice to be successful only when applied with the dynamical coordinates and momenta referring to a Cartesian system of axes and not to more general curvilinear coordinates.

The remark refers to the important problem of how to make a transition from classical to quantum mechanics (Yilmaz, 1981).[23] If we have a classical Hamiltonian $H(p, q)$, and if q's are Cartesian coordinates, then the substitution $p = i\hbar\,\partial/\partial q$ yields a valid quantum mechanical Hamiltonian, but if q's are curvilinear coordinates we do not know how to set up correctly the quantum mechanics.

Note that the problem already occurs in the simplest possible situations, namely, the nonrelativistic quantum mechanics of a particle, and even when the forces acting on a particle are trivially simple. The relativistic gravitational counterpart of making a valid transition from classical to quantum mechanics in curved space must therefore be regarded as correspondingly more difficult and, up to present, unresolved.

[23] This means the signal velocity has no dispersion which can be observationally tested. Another test of the theory could be a direct comparison of the velocity of light in vertical and horizontal directions in the laboratory. The present theory favors local Lorentz covariance as the proper generalization of special relativity and relegates general covariance to a mathematical requirement of computational consistency.

Note also that the problem is not necessarily how to make a transition to curvilinear coordinates when one has the correct Hamiltonian in Cartesian coordinates. This can presumably be done in some cases by coordinate transformations. The problem is how to set up quantum mechanics when one did not start out with a Cartesian system or when there are *no* Cartesian systems of axes, as in the case of a curved space.

Our proposed solution to this problem of how to set up quantum mechanics in curvilinear coordinates in general, and in curved spaces in particular, is as follows: Quantum mechanics requires the *kinematical velocity* v of a particle and the *group velocity* v_g of its associated wave to be equal (probability postulate). In the limit of zero rest mass, or infinitely large momenta, the phase velocity V and the group velocity v_g are also equal. Thus for zero rest mass one has a *unique* signal velocity $v = v_g = V = c(x)$. This unique signal velocity can serve as the basis of an operational procedure of measurements valid both for waves and particles. This is necessary for the *unity* of *wave* and *particle* points of view, beginning at the level of their operational contents. Note that the signal velocity must be the same for *all wavelengths* (Yilmaz, 1981), for we cannot have different operational procedures for different wavelengths. When these conditions are satisfied one can show that a particle of nonzero rest mass satisfies $v_g = v$, $v \cdot V = c^2(x)$ for all frequencies including that of infinite frequency $v_\infty = v$. In other words quantum mechanics can then be set up in analogy to Lorentz space (in Minkowskian coordinates) and to Galilean space (in Cartesian coordinates). From our point of view the virtue of the latter two systems is that they satisfy the above conditions automatically.

Now why does the choice of coordinates make such an unexpected difference? The reason is that the coordinate velocity of a particle with zero rest mass is found from the line element $ds^2 = 0$, whereas the group velocity of a wave with zero rest mass is found from $\Box^2\chi = 0$. These two (coordinate) velocities are in general different unless the coordinate system is of a selected kind.[24] This intuitively nonobvious fact can be appreciated only by an explicit example. Take the line element to be diagonal and isotropic

$$ds^2 = A\,dt^2 - B(dx^2 + dy^2 + dz^2)$$

where, A, B are functions of x, y, z. There exists a solution of the form

$$q = q_0 e^{i(\omega t - kx)}$$

[24] H. Weyl often stressed the need for selected systems of coordinates both for wave–particle duality ($v_g = v$) and for purposes of representation of spinor particles (local orthogonal group).

where q_0 is constant.

$$\omega^2 = AB^{-1}(k^2 + \xi \cdot k)$$

where $\xi = i(AB)^{-1/2}\partial(AB)^{1/2}$. The three velocities are

$$v = \left(\frac{dx^2 + dy^2 + dz^2}{dt^2}\right)^{1/2} = (AB^{-1})^{1/2}$$

$$V = \frac{\omega}{k} = (AB^{-1})^{1/2}\left(1 + \frac{\xi}{k}\right)^{1/2}$$

$$v_g = \frac{\partial\omega}{\partial k} = AB^{-1}V^{-1}\left(1 + \frac{\xi}{2k}\right)$$

We can see that only when $\xi = 0$ are the three velocities equal, and this requires $AB = \text{const.}$ (If $\xi \neq 0$ then only $k \to \infty$ limit satisfies $v_g = v_\infty$ which is, in turn, equal to v.) A special relativistic boundary condition then implies $A = 1/B$, so in order to be able to set up quantum mechanics in isotropic coordinates the metric must also satisfy $A = 1/B$. According to this result one cannot set up quantum mechanics, for example, in the isotropic coordinates of Einstein's theory, since its isotropic solution does not satisfy this condition.[25] One can, however, set it up in the nonisotropic coordinates given in the earlier part of this appendix (set $\lambda = 0$ or $\varepsilon = 1$). But then the Hamiltonian (and also the signal velocity) will have a small anisotropic term which can be calculated and experimentally tested as to its existence. We have seen that the experiment gives $\lambda = 1$, which in turn implies $A = 1/B$. This result may be considered empirical evidence that quantum mechanics is consistent with the strong principle of equivalence which also requires $\lambda = 1$.

We conclude the discussion with a slight elaboration. Let the wave equation be of the form

$$(D^\nu D_\nu + m_0^2/\hbar^2)q = 0$$

$$q = q_0 e^{iS}, \qquad S = (1/\hbar)\int p_\mu \, dx^\mu = \int (\omega \, dt - k \cdot dx)$$

Then there exists a solution which renders $H(p, x)$ and $\omega = H(\hbar k, x)$

[25]Note that the condition $\xi = 0$ is none other than the gauge condition $\partial_\nu[(-g)^{1/2}g^{\mu\nu}] = 0$ of Appendix A. The wave $e^{i(\omega t - k \cdot x)} = e^{(i/\hbar)p_\mu x^\mu}$ is not plane since p_μ is constant under covariant and not under ordinary differentiation.

functionally the same if

$$\Box^2 q_0 = 0, \qquad \Box^2 S + 2q_0^{-1} g^{\mu\nu} D_\mu q_0 D_\nu S = 0$$

In rectangular coordinates $q_0 = \text{const}$ and in spherical coordinates $q_0 = C/r$.[26] Such coordinates allow the Schrödinger equation to be set up directly, or by transformation from another satisfying the same conditions. In rectangular coordinates these are equivalent to the "Harmonic conditions" $\partial_\nu[(-g)^{1/2} g^{\mu\nu}]k_\mu = 0$ which are satisfied by our solutions. Such coordinates form a "selected" group within all classically allowable ones and may here be called "Schrödinger coordinates" since they are restricted to a group in which quantum mechanics can be set up in the Schrödinger representation.[27] Solutions of our theory automatically meet this quantum mechanical condition and, in addition, satisfy the strong principle of equivalence. It is further found that the strong principle of equivalence and the quantum mechanical superposition of states are consistent and this avoids certain ambiguities which would arise if the principle of equivalence were assumed only in its weak form. In the static limit the $m_0 \neq 0$ and $m_0 = 0$ Hamiltonians are

$$H = e^{-\phi}\left(m_0^2 + e^{-2\phi} p^2\right)^{1/2}$$

$$H = e^{-2\phi} |p|$$

and these reproduce, without ambiguity, all known experiments having to do with a static field of gravitation (Yilmaz, 1980; 1977).

HISTORICAL NOTE ON FIELD STRESS-ENERGY

As is well known, Einstein presented a field stress-energy (let us call it $\bar{t}_\mu^\nu$) on the basis of his linear approximation to $g_{\mu\nu}$. This $\bar{t}_\mu^\nu$ was obtained by computing the product terms of Christoffel symbols in $R_\mu^\nu - \frac{1}{2}\delta_\mu^\nu R$. Less well known is the fact that some 20 years later Einstein found this was a mistake. In a letter to Max Born (Born & Born, 1971) he wrote:

> Together with a young collaborator, I arrived at the interesting result that the gravitational waves do not exist, though they had been assumed a certainty to the first approximation. This shows that the nonlinear general relativistic field equations can tell us more, or rather, limit us more than we have believed up to now.

[26] The equations lead to $q_0 q_0'' - 2q_0'^2 = 0$, hence $q_0 = C, C/r$. Thus $\Box^2 q_0 = 0$ gives (for the previous metric) $A = 1/B$ both in rectangular and in spherical coordinates.

[27] There are theoretical reasons to believe that the only Schrödinger coordinates with a relativistic covariance group are those which at the same time satisfy $\lambda = 1$. Note that $\Box^2 q_0 = 0$ has solutions more general than $q_0 = C, C/r$ e.g. $R_{nl} Y_l^m(\cos\theta)$ making up the amplitude q_0.

With hindsight the origin of the mistake is clear: $R_\mu^\nu - \frac{1}{2}\delta_\mu^\nu R$ contains terms $\partial_\nu \Gamma$ and $\Gamma \cdot \Gamma$ where Γ are Christoffel symbols. A first-order approximation to $g_{\mu\nu}$ can be meaningful *only* for the $\partial_\nu \Gamma$ part since $\Gamma \cdot \Gamma$ is second order. To calculate t_μ^ν meaningfully (it is second order), one has to carry $g_{\mu\nu}$ to second order. When this is done the second-order contributions to $\partial_\nu \Gamma$ combine with $\Gamma \cdot \Gamma$ giving a mathematically legitimate quantity. But in Einstein's theory this combination must *necessarily vanish* in free space because the free space equations are $R_{\mu\nu} = 0$, $R = 0$, requiring everything to vanish *order by order*. Unfortunately Einstein's result seems to be forgotten, creating confusion even to this day. Some authors transfer a t_μ^ν to the right-hand side of the equations and claim it a legitimate stress-energy. This is equivalent to transferring the second-order contributions of $\partial_\nu \Gamma$ to the left, which destroys the *integrity* of the divergence-free tensor $R_\mu^\nu - \frac{1}{2}\delta_\mu^\nu R$. These matters were pointed out by many authors, including C. Möller (1958), L. Infeld (1959), A. E. Scheidegger (1956), N. Rosen (1956), A. A. Logunov and co-workers (Logunov et al., 1977; Denisov and Logunov, 1980), the present writer (Yilmaz, 1975, 1979), and, as it turns out, clearly noted by Einstein himself. It is a curious phenomenon that despite simple mathematical proof, the confusion lingers, leading to elaborate calculations with conflicting results. In the new theory a legitimate field stress-energy (t_μ^ν) exists, as we have seen. It is the canonical stress-energy of the Lagrangian of Appendix A.

NOTE ON FIELD QUANTIZATION

A chief ingredient of field quantization in any field theory is to bring the field energy-momentum

$$P_\mu = \int (-g)^{1/2} t_\mu^\nu \, dV_\nu$$

into a form $P_\mu = \Sigma_j |c_j|^2 n_j \hbar k_{\mu j}$ where $n_j = 0, 1, 2, \ldots, \infty$ or $n_j = 0, 1$. This is achieved by interpreting the field quantities appearing in the quadratic expression of t_μ^ν as operators and subjecting them to commutation relations. The commutation relations are also quadratic in field quantities, so the quantization can be achieved by choosing the commutation properties appropriately. It is obvious, however, that unless one has a legitimate nonzero t_μ^ν to start with, commutation relations cannot serve any purpose. When one has a legitimate nonzero t_μ^ν one may facilitate the appropriate formulation of quantization by obtaining the t_μ^ν as the canonical stress-energy of a field Lagrangian. One can check that our t_μ^ν is the canonical stress-energy of our field Lagrangian in Appendix A.

The above can also be done directly on $g_{\mu\nu}$ if we rewrite the Lagrangian as

$$L = -\tfrac{1}{16} g^{\mu\nu} g^{\alpha\beta} g^{pq} \left(\partial_\mu g_{\alpha p} \partial_\nu g_{\beta q} - \tfrac{1}{2} \partial_\mu g_{\alpha\beta} \partial_\nu g_{pq} \right)$$

and vary it under the gauge condition $\partial_\nu[(-g)^{1/2} g^{\mu\nu}] = 0$ (rectangular coordinates). The gauge condition is equivalent to imposing $\partial^\nu = g^{\mu\nu}\partial_\mu$, that is, to absorbing the $g^{\mu\nu}$ into ∂^ν and thereby omitting the variation of the corresponding $g^{\mu\nu}$ from the variation of the total action. (In other words the gauge here acts as a kinematical constraint). A lengthy but straightforward calculation leads to the Euler–Lagrange equations

$$\partial^\nu \left(g^{\alpha\lambda} \partial_\nu g_{\beta\lambda} \right) = 0, \qquad \partial^\nu \left(g^{\alpha\beta} \partial_\nu g_{\alpha\beta} \right) = 0$$

which are equivalent to our field equations $\Box^2 \phi^\alpha_\beta = 0$, $\Box^2 \phi = 0$. When the matter action is added, these equations become $\Box^2 \phi^\alpha_\beta = \tau^\alpha_\beta$, $\Box^2 \phi = \tau$, where τ^α_β is the matter current and τ is its trace.

An important consequence of obtaining both the field stress-energy t^ν_μ and the term $\Box^2 \phi^\nu_\mu$ identifying the matter stress-energy from the same Lagrangian is that both t^ν_μ and τ^ν_μ must have the same numerical coefficient (same in all choices of units) when they are added to form the total stress-energy $T^\nu_\mu = \tau^\nu_\mu + t^\nu_\mu$. In other words τ^ν_μ alone cannot be present in the geometric field equations; it must always be accompanied by the t^ν_μ on equal footing.

A quantization of the gravitational field along the lines of the new theory is carried out for the simple case of prescribed c-number sources with a coordinate measure $(-g)^{1/2} d^4x$. No fundamental difficulty seems to be encountered. The more difficult case of Dirac-type q-number sources is being attempted as a local gauge theory. We hope to report on this at a future occasion.

NOTE ADDED IN PROOF

The crucial discussion in Appendix B may be made clearer if the basic conserved quantities P_μ, $\Delta P_\mu = 0$ are worked out ahead of time. Start with $D_\nu T^\nu_\mu = 0$ and form the increment $\Delta P_\mu = \int \sqrt{-g} D_\nu T^\nu_\mu \, d\Omega = 0$. To obtain the conserved quantity P_μ we need to use Gauss' theorem $d\Omega\, \partial_\nu = dV_\nu$ but for this we must have the ordinary divergence ∂_μ, because Gauss' theorem works with the ordinary divergence only. We therefore split T^ν_μ as $T^\nu_\mu = \tau^\nu_\mu + t^\nu_\mu$ and impose the "integrability conditions" $D_\nu t_\mu^\nu = \frac{1}{2}\partial_\mu g_{\alpha\beta}\tau^{\alpha\beta}$. We then get

$$\partial_\mu(\sqrt{-g}\tau_\mu^\nu) = 0$$

$$P_\mu = \int \sqrt{-g}\tau_\mu^\nu \, dV_\nu = C_\mu$$

which is the mathematically correct form of the conservation laws for a symmetric tensor $T_{\mu\nu} = T_{\nu\mu}$, $D_\nu T_\mu^\nu = 0$. For a vector j^ν, $D_\nu j^\nu = 0$, or for an antisymmetric tensor $K_{\alpha\beta}^\nu = -K_{\beta\alpha}^\nu$, $D_\nu K_{\alpha\beta}^\nu = 0$, such a splitting is not necessary because then the divergence is already in a form suitable for Gauss' theorem. The integrability conditions turn out to be the basis of the field-geometry equivalence mentioned in the text.

REFERENCES

Born, M., and Born, H. (1971). The *Born–Einstein Letters*, p. 125. Walker and Co., New York.

Colella, R., Overhouser, A. W., and Werner, S. A. (1975). *Physical Review Letters*, **34**, 1472.

Denisov, V. I., and Logunov, A. A. (1980). Institute of Nuclear Research, Academy of Sciences USSR, P-0159.

Dirac, P. A. M. (1975). *General Theory of Relativity*, p. 38. Wiley, New York.

Dirac, P. A. M. (1976). *Quantum Mechanics*, p. 114, fourth revised edition. Oxford University Press.

Drever, R. W. P. (1961). *Philosophical Magazine*, **6**, 683.

Eddington, A. S. (1957). *Theory of Relativity*, p. 94, Cambridge University Press.

Feynman, R. P. (1971). *Lectures on Gravitation*, p. 110. California Institute of Technology, Pasadena.

Greenberger, D. M., and Overhauser, W. A. (1980). *Scientific American*, May, 66.

Greenberger, D. M. (1968). *Annals of Physics*, **47**, 116.

Hughes, V. W. (1964). *Gravitation and Relativity*, H. Y. Chiu and W. F. Hoffman, eds., p. 106. W. A. Benjamin, New York.

Infeld, L. (1959). *Annals of Physics*, **6**, 341.

Landau, L. D., and Lifshitz, H. M. (1962). *Classical Theory of Fields*, footnote, p. 341. Addison-Wesley, Reading, Massachusetts.

Logunov, A. A., Folomeskin, N. V., and Vlasov, A. A. (1977). *Theoretical and Mathematical Physics (USSR)*, **33**, 174.

Möller, C. (1958). *Max Planck Festschrift*, p. 139. Berlin.

Ohanian, H. C. (1976). *Gravitation and Spacetime*, W. W. Norton, New York.

Rosen, N. (1956). *Jubilee of Relativity Theory*, Basel.

Schwartz, H. M. (1977). *American Journal of Physics*, **45**, 899.

Scheidegger, A. E. (1956). *Reviews of Modern Physics*, **25**, 451.

Tupper, B. O. J. (1974a). *Nuovo Cimento*, **19B**, 135.

Tupper, B. O. J. (1974b). *Nuovo Cimento Letters*, **10**, 627.

Vessot, R. F. C., and Levine, M. W. (1981). *Physical Review Letters*, **45**, 26, 2082.

Weinberg, S. (1965). *Physical Review*, **138**, 988 (p. 999).

Will, C. M. (1974). *Experimental Gravitation*, B. Bertotti, ed., pp. 44, 10. Academic Press, New York.

Yilmaz, H. (1973). *Nuovo Cimento Letters*, **7**, 337.
Yilmaz, H. (1975). *Nuovo Cimento*, **26B**, 577.
Yilmaz, H. (1976). *Annals of Physics*, **104**, 413.
Yilmaz, H. (1977). *Nuovo Cimento Letters*, **20**, 681.
Yilmaz, H. (1978). *Nuovo Cimento Letters*, **22**, 647.
Yilmaz, H. (1979). *Hadronic Journal* **2**, 1196.
Yilmaz, H. (1980). *Hadronic Journal*, **3**, 1478.
Yilmaz, H. (1981). *Physics Today*, October.