PHYSICAL REVIEW D VOLUME 18, NUMBER 12 15 DECEMBER 1978

General relativity without general relativity

Jacob D. Bekenstein and Amnon Meisels
Physics Department, Ben Gurion University of the Negev, Beer-Sheva, Israel
(Received 27 June 1978)

Recently one of us proposed a general theory of variable rest masses (VMT) compatible with post-Newtonian solar-system experiments for a wide range of its two parameters r and q, provided the asymptotic value of its fundamental field f is in a certain narrow range. Here we show that the stationary matter-free black-hole solutions of the VMT are identical to those of general relativity. In addition, for $r<0$ and $q>0$ (part of the range mentioned), relativistic neutron-star models in the VMT are very similar to their general-relativistic counterparts. Thus experimental discrimination between the two theories in the strong-field limit seems unfeasible. We show that in all isotropic cosmological models of the VMT capable of describing the present epoch, the Newtonian gravitational constant G_N is positive throughout the cosmological expansion. There exist nonsingular VMT cosmological solutions; this is an advantage the VMT has over general relativity. For $r<0$ and $q>0$ all VMT cosmological models converge to their general-relativistic analogs at late times. As a consequence the asymptotic f attains just the required values to guarantee agreement of the VMT with post-Newtonian experiments. The VMT with $r<0$ and $q>0$ predicts a positive Nordtvedt-effect coefficient. It also predicts that G_N is currently decreasing on a time scale which could be long compared to the Hubble time. Verification of these predictions would rule out general relativity; its most natural replacement would be the VMT with $r<0$ and $q>0$, and not a generic scalar-tensor theory. The success of general relativity in most respects could then be understood because the VMT with $r<0$ and $q>0$ mimics it. Because of this, general relativity could still be used, for most purposes, as a good approximation to the correct gravitational theory.

I. INTRODUCTION

General relativity (GR) assumes the validity of the strong equivalence principle, and in particular of the assumption that particle rest masses are spacetime constants in units for which the gravitational "constant" G is indeed constant. This assumption is a strong one; the experimental evidence relating to it is fragmentary, and is far from establishing it on a firm empirical basis. Let us recall that variability of a rest mass m in units for which G is constant would be equivalent to variability of the dimensionless quantity $\gamma = Gm^2/\hbar c$. Measurements of the orbital motion of Mercury have set an upper limit on the possible time variability of the γ for the electron γ_e, but this limit is not incompatible with variability on a time scale of the Hubble time t_H.[1] By contrast, studies of the moon's orbital motion have been interpreted as showing variability of γ_e on the time scale of t_H.[2] Thus the evidence for the constancy of rest masses is ambiguous; there clearly is a need for a general theoretical framework for testing the possibility that rest masses vary in a wider class of phenomena.

Such a framework was proposed by one of us in the form of a general classical field theory of variable rest masses (VMT) within the framework of Einstein's gravitational field equations.[3] The VMT has just two free parameters, r and q. In units for which rest masses are *constant* the VMT turns out to be identical to a small subset of the scalar-tensor theories of gravitation (STT's).[4-6] In scalar-tensor form the VMT is easily compared with the results of solar-system post-Newtonian (PN) experiments. In this way a large fraction of the rq plane is experimentally ruled out.[3] But there is still a wide range of rq values for which the theory will be compatible with the experiments if the asymptotic value of a scalar field f appearing in it falls in a certain narrow range. In the present paper we shall show that for most of this range of rq values, the predictions of the VMT for solar-system PN experiments, black holes, neutron stars, and late phases of cosmology are so close to those of GR that there is little hope that observations in the foreseeable future will discriminate between the theories (barring confirmation of the time variability of γ which would rule out GR). Thus GR is bound to be a good approximation to the correct theory even if the constant masses postulate is violated.

The plan of the paper is as follows. In Sec. II we briefly review the VMT and show that in its STT form it is also the most general variable-G theory possible under fairly general assumptions. In Sec. III we show that all stationary black-hole (BH) solutions of the VMT are identical to the corresponding ones in GR. Thus, for the strongest gravitational fields the predictions of the VMT and GR merge. In Sec. IV we show that relativistic neutron-star (NS) models in the VMT with $r<0$ and $q>0$ (range compatible with PN experiments) approximate the corresponding ones in GR so closely as to make observational discrimination between the two theories unfeasible. Thus in the strong-field limit the VMT with $r<0$ and $q>0$ is virtually indistinguishable from GR.

The asymptotic value of f which decides whether the VMT with appropriate r and q can fit the solar-system experiments is determined by the cosmological model chosen to describe the universe at large.[3] Thus in Sec. V we turn to cosmological models in the VMT. We write the equations for isotropic models, and show that unlike cosmological models in a generic STT, VMT models have a positive Newtonian gravitational constant G_N throughout the expansion. Further, unlike GR models, VMT models can be nonsingular. Thus rejection of the experimentally unverified constant-masses postulate of GR offers a way out of the cosmological singularity dilemma. In Sec. VI we show that, again for $r<0$ and $q>0$, VMT cosmological models spontaneously converge to those of GR at late times. Further, the asymptotic value of f determined by the model at late times is just such as to secure agreement of the VMT predictions with the results of PN experiments. The time scale for f to approach this value can be much shorter than the time scale for changes of G_N. Thus a model starting with nearly arbitrary initial conditions can reach the required f by the present epoch without this leading to large variability of G_N which is not observed.

Our conclusions are summarized in Sec. VII.

II. THEORY OF VARIABLE REST MASSES

On the basis of a number of experiments all rest masses are regarded as proportional to a universal scalar mass field χ.[3] The dynamical action for χ is determined by simple assumptions: general covariance, second-order equation for χ, and the absence of a constant scale of length in the theory. The gravitational action is taken as Einstein's action so that the VMT retains Einstein's field equations. Of the resulting general VMT[3] a special case is Dicke's[7] reformulation of Brans-Dicke theory[8] as a variable-masses theory. For the general case $\chi \propto \psi^r$ where r is a real parameter, and the dynamical action for the real field ψ is

$$S=-\tfrac{1}{2}\int(\psi_{,\alpha}\psi^{,\alpha}+q\tilde{R}\psi^2)(-g)^{1/2}d^4x\,. \tag{1}$$

Here q is the second (real) parameter of the theory, and $\tilde{R}$ and $\tilde{g}$ are the scalar curvature and determinant of the metric, respectively. Up to now the units are such that the gravitational "constant" is constant. It is convenient to transform to units for which rest masses are constant. Let

$$f\equiv 8\pi G c^{-4}\psi^2>0\,, \tag{2}$$

$$\phi=(1-qf)f^{-r}\,, \tag{3}$$

and define the implicit function of ϕ

$$\omega(\phi)\equiv-\tfrac{3}{2}-\tfrac{1}{4}f[(1-6q)qf-1][r+(1-r)qf]^{-2}\,. \tag{4}$$

Then in the new units the combined action for gravitation and the field χ takes the STT form[3]

$$S=c^4(16\pi G_0)^{-1}\int[\phi R-\omega(\phi)\phi^{-1}\phi_{,\alpha}\phi^{,\alpha}](-g)^{1/2}d^4x\,. \tag{5}$$

Here G_0 is a constant coupling constant and R and g refer to the metric in the new units. (From now on we shall set $c=1$.) Thus the general VMT is equivalent to a very small subset of the STT's — those defined by constant ω (Brans-Dicke theory), and those defined by $\omega(\phi)$ given by (3) and (4). The STT's are commonly regarded as theories of a variable gravitational constant G. Here we shall argue that this interpretation is meaningful only for Brans-Dicke theory, or for STT with $\omega(\phi)$ defined by (3) and (4). As stressed by Dicke[7] one can talk about absolute variability only in reference to a dimensionless quantity, such as the γ_e of Sec. I. If γ_e varies, one can describe this as variability of mass and constancy of G in certain units, or as variability of G and constancy of mass in some other units. It follows that any good theory of variable G (in units for which masses are constant-particle units) must also be expressible as a theory of variable masses (in units for which G is constant-Planck-Wheeler units). The Brans-Dicke theory meets this criterion; Dicke[7] has given the transformation to a variable-masses theory. The STT's defined by (3)–(5) also meet the criterion; their variable masses version is just the general case of VMT.[3] However, for a STT with more general $\omega(\phi)$ there is no way to effect the transformation to a variable-masses–constant-G version.

Suppose such a transformation were possible. It would necessarily give a variable-masses theory more general than the VMT of Ref. 3. In this case one or more of the postulates on which the VMT is predicated would be violated. The postulate that the VMT is generally covariant cannot be violated since every STT is generally covariant, and the scaling factor of the transformation must be a scalar (it is just a ratio of two different units for the same quantity). The postulate that the equation for the mass field is of second order cannot be violated since the equations for all STT's [see (6) and (8) below] are of second order and a transformation of units will not raise the order (the scaling factor cannot depend on the derivatives as indicated below). Finally, the postulate that the variable-masses theory does not contain a constant scale of length could be violated only if the scaling factor of the transformation contains this length since the STT's do not. Now the (scalar) scaling factor is a ratio of two different units for the same quantity, and is therefore dimensionless. It can thus contain a constant with dimensions of length only in conjunction with some scalar having dimensions of length. The only quantities we have available to construct such a scalar are derivatives of ϕ or the curvature. If the transformation in question exists, it maps a variable G into a constant G, and this means in general that the variable G depended on derivatives of ϕ or on the curvature. Yet in the STT's G depends only on ϕ [see Eq. (24)]. Thus the scaling factor containing a constant scale of length cannot be found. We conclude that the VMT of Ref. 3 in STT form is also the most general theory of variable G under very general assumptions.

The gravitational field equations of the VMT in STT form are obtained by varying the complete action with respect to the metric[4]:

$$G_\mu{}^\nu = 8\pi G_0\phi^{-1}T_\mu{}^\nu + \omega\phi^{-2}(\phi_{,\mu}\phi^{,\nu} - \tfrac{1}{2}\phi_{,\gamma}\phi^{,\gamma}\delta_\mu{}^\nu) + \phi^{-1}(\phi_{,\mu}{}^{;\nu} - \phi_{,\gamma}{}^{;\gamma}\delta_\mu{}^\nu)\,, \tag{6}$$

where $G_\mu{}^\nu$ is the Einstein tensor and $T_\mu{}^\nu$ is the stress-energy tensor obtained by varying the matter action. In particle units the matter action does not depend on ϕ; thus variation of the total action with respect to ϕ gives

$$R + \phi^{-1}(\omega' - \omega\phi^{-1})\phi_{,\alpha}\phi^{,\alpha} + 2\omega\phi^{-1}\phi_{,\alpha}{}^{;\alpha} = 0\,, \tag{7}$$

where $\omega' \equiv d\omega/d\phi$. Replacing the scalar curvature R by the expression obtained from the trace of (6) we get an equation[4] which may be put in the form

$$\Omega^{1/2}(\Omega^{1/2}\phi^{,\alpha})_{;\alpha} = 8\pi G_0 T\,, \tag{8}$$

where $T = T_\mu{}^\mu$ and $\Omega = 3 + 2\omega$. Equations (2)–(4), (6), and (8) are the basic equations we shall use.

III. IDENTITY OF BLACK HOLES IN THE VMT AND IN GR

The viability of the VMT for weak fields was considered in Ref. 3. Let us now enquire how it fares in the limit of the strongest gravitational fields possible, those of black holes. For this purpose we focus on an exterior matter-free BH solution of the VMT endowed with stationary and axial symmetry. Let $\tilde\phi = \phi - \phi_0$, where ϕ_0 is the asymptotic value of ϕ. We shall show that necessarily $\tilde\phi = 0$ throughout the BH exterior so that the metric satisfies Einstein's equations with gravitational constant $G_0\phi_0{}^{-1}$ [see (6)]. Thus our VMT BH solution is also a solution of GR. That every GR BH solution is also a VMT BH solution (with ϕ constant) is trivially true. Thus the Kerr black holes constitute the totality of stationary matter-free black holes in the VMT.

The procedure we follow is adapted from those of Hawking[9] and one of us,[10] used to establish the identity of Brans-Dicke and GR black holes. It follows from (8) that

$$\tilde\phi(\Omega^{1/2}\tilde\phi^{,\alpha})_{;\alpha} = 0 \tag{9}$$

since $\tilde\phi_{,\alpha} = \phi_{,\alpha}$. Let us integrate (9) over the whole BH exterior between two spacelike hypersurfaces related by a translation along an asymptotically timelike Killing vector of the geometry. After integration by parts we get

$$\int \Omega^{1/2}\tilde\phi_{,\alpha}\tilde\phi^{,\alpha}(-g)^{1/2}d^4x - \oint_H \Omega^{1/2}\tilde\phi\tilde\phi^{,\alpha}d\Sigma_\alpha = 0\,, \tag{10}$$

where the surface integral is taken over the horizon H between the spacelike hypersurfaces. The integrals over the hypersurfaces have canceled by symmetry and that over infinity has vanished because $\tilde\phi$ and $\tilde\phi_{,\alpha}$ must vanish as $1/r$ and $1/r^2$, respectively. The vector $\tilde\phi_{,\alpha}$ has no components in the time and axial directions by symmetry. Thus on H it is orthogonal to $d\Sigma_\alpha$ which, being parallel to H's generator, is entirely in a Killing direction.[9] Therefore, the surface integral in (10) will vanish provided the norm of the vector $\Omega^{1/2}\tilde\phi\tilde\phi_{,\alpha}$ is bounded on H. If it is not bounded we proceed as follows.

We first write the surface integral as

$$I_s = \oint_H \Omega^{1/2}\tilde\phi(d\tilde\phi/df)f^{,\alpha}d\Sigma_\alpha\,. \tag{11}$$

Now f must not vanish or blow up on H. Such behavior would lead either to χ becoming unbounded (masses infinite at the horizon) or vanishing there (with consequent vanishing of the metric in particle units or blowing up in Planck-Wheeler units since χ^2 is the ratio between the two metrics[3]). All these kinds of behavior are incompatible with the regular character of a BH horizon. A direct calculation based on (3) and (4) shows that $\Omega^{1/2}\tilde\phi(d\tilde\phi/df)$ will be bounded on H, the pole of $\Omega^{1/2}$ being canceled by $d\tilde\phi/df$. It only remains to show that $f^{,\alpha}d\Sigma_\alpha$ vanishes on H.

Our assumption that the norm of $\Omega^{1/2}\tilde\phi(d\tilde\phi/df)f_{,\alpha}$

diverges on H clearly means that $f_{,\alpha}f^{,\alpha}\to\infty$ there. By eliminating $\phi_{,\alpha}{}^{;\alpha}$ between (7) and (8) with $T=0$ we have

$$f_{,\alpha}f^{,\alpha}=R\phi^2[(d\phi/df)^2(\omega-3\omega'\phi\Omega^{-1})]^{-1}. \quad (12)$$

Clearly, the scalar curvature R must be bounded on H for it to qualify as a BH horizon. Also f must be bounded and nonzero on H. Thus $f_{,\alpha}f^{,\alpha}$ can become unbounded as assumed only if f takes on H a constant value f_H which is one of the zeros of the function in square brackets in (12). [There are at most four such zeros; see (3) and (4).] Let us exploit this fact to define in the neighborhood of H a radial-like coordinate which is taken to be a function of f only [also $f=f(\xi)$ only] with $\xi=0$ at $f=f_H$. Clearly there is an enormous freedom in defining ξ; we shall use up part of it to arrange that the metric component $g^{\xi\xi}$ behaves as an integral power (positive, negative, or zero) of ξ near $\xi=0$.

Let us define the hypersurface element on surfaces $\xi=$const:

$$d\Sigma_\alpha=\tfrac{1}{6}(-g)^{1/2}\epsilon_{\alpha\beta\gamma\delta}dx^\beta\wedge dx^\gamma\wedge dx^\delta. \quad (13)$$

The dx^β are in all directions normal to the ξ direction. Thus the only nonvanishing component of $d\Sigma_\alpha$ is $d\Sigma_\xi$. In the limit $\xi\to 0$, $d\Sigma_\alpha$ must become the product of the horizon's (normal) generator n_α, the two-surface element on H, and an increment in affine parameter along n_α.[11] Thus $d\Sigma_\xi$ cannot vanish as $\xi\to 0$. However, by the null character of n_α, $d\Sigma^\alpha d\Sigma_\alpha=g^{\xi\xi}(d\Sigma_\xi)^2\to 0$ as $\xi\to 0$. Clearly, $g^{\xi\xi}$ vanishes in the limit, and since it behaves as an integral power of ξ, it must vanish at least as ξ. Thus, $f^{,\alpha}d\Sigma_\alpha=f_{,\xi}g^{\xi\xi}d\Sigma_\xi$ must be bounded by $\xi f_{,\xi}$ in the limit. Now, in order for f to be bounded at $\xi=0$ the integral $\int f_{,\xi}d\xi$ must converge there, and this implies that $\xi f_{,\xi}\to 0$ as $\xi\to 0$. Thus, $f^{,\alpha}d\Sigma_\alpha\to 0$ and the surface integral in (11) must vanish even if the norm of $\Omega^{1/2}\tilde\phi\phi_{,\alpha}$ blows up on H.

From (10) it follows that the four-dimensional integral vanishes. Since $\tilde\phi_{,\alpha}$ has no component in the timelike directions it is spacelike: $\tilde\phi_{,\alpha}\tilde\phi^{,\alpha}\geq 0$. Suppose that Ω is strictly positive or strictly negative in the BH exterior. Then the integral can vanish only if $\tilde\phi$ is constant throughout. If Ω is of one sign throughout, but vanishes on a closed two-surface, $\Omega^{1/2}$ can be opposite in sign inside and outside it. In this case we carry out the procedure leading to (10) twice, once in each region. The two-surface does not contribute a boundary term since $\Omega^{1/2}$ vanishes on it. The proof that $\tilde\phi$ is constant for the interior region exactly parallels that given earlier, while that for the exterior region is trivial since there is no boundary term. If Ω is positive in one region and negative in another, the requirement that real and imaginary parts of the four-dimensional integral vanish separately produces the desired result $\tilde\phi=$const. Finally, if Ω vanishes throughout the exterior, $\tilde\phi$ is constant trivially. Since $\tilde\phi=0$ asymptotically we conclude that $\phi=\phi_0=$constant throughout the BH exterior. Our earlier conclusion that the BH solutions in VMT and GR coincide follows.

Our assumption of axisymmetry is not really required. For GR BH solutions Hawking has shown[12] that axisymmetry follows from stationarity, and little modification is needed to extend the proof to the general STT. Our proof also applies for a black hole endowed with an electromagnetic field for which $T=0$ also. The assumption $T=0$ is essential and matter in the BH exterior will split the degeneracy between VMT and GR. However, in all astrophysically interesting situations, matter will be a small perturbation and the difference between the predictions of the VMT and GR should be minute. Thus, observations of x-ray sources which may be black holes[13] will not allow one to discriminate between GR and VMT in the foreseeable future in view of the large uncertainties in the nongravitational physics of such sources.[13]

We must stress that our proof relies heavily on the specific form of ω for the VMT. It is not clear that black holes in an arbitrary STT will be identical to those of GR. The role played by the variable-masses interpretation of the theory must also be underlined. One of us has given a solution of GR representing a BH endowed with a conformal scalar field.[14] The equations are also, formally, those of the VMT with $r=1$, $q=\frac{1}{6}$. But a black-hole interpretation in the VMT is ruled out because χ blows up on the would-be horizon.

IV. RELATIVISTIC NEUTRON STARS IN THE VMT

Next to black holes, neutron stars are the systems with the strongest gravitational fields. Whereas for BH's study of matter-free solutions suffices to elucidate the main features of the problem, for NS's, consideration of the matter content is crucial for the analysis. We saw that matter can split the degeneracy between VMT and GR BH's. To what extent do VMT and GR models of relativistic neutron stars differ? For simplicity we shall consider only spherically symmetric stationary models.

The metric can be chosen of the form

$$ds^2=-Bdt^2+Dd\zeta^2+\zeta^2(d\theta^2+\sin^2\theta\, d\Phi^2), \quad (14)$$

where $B=B(\zeta)$ and $D=D(\zeta)$ are positive, well behaved at $\zeta=0$, and reduce to unity as $\zeta\to\infty$. Since $\phi=\phi(\zeta)$ and $T=T(\zeta)$, the scalar equation (8) reduces to

$$[(B\Omega/D)^{1/2}\xi^2(d\phi/df)f']' = 8\pi G_0(BD/\Omega)^{1/2}\xi^2 T\,, \tag{15}$$

where a prime denotes derivative with respect to ξ. Integrating (15) between the center of the star $\xi=0$ and a general ξ, we have, using (3),

$$f' = -8\pi G_0(D/B\Omega)^{1/2}\xi^{-2}f^{r+1}[q(1-r)(f-f_c)]^{-1} \times \int_0^{\xi}(BD/\Omega)^{1/2}T\xi^2 d\xi\,, \tag{16}$$

where

$$f_c \equiv r(r-1)^{-1}q^{-1}\,. \tag{17}$$

For all known forms of matter, $T<0$. Thus f' has the same sign as $q(1-r)(f-f_c)$.

Consider first the case $r<0$ and $q>0$. We see that as ξ decreases, f approaches f_c monotonically whatever the sign of $f-f_c$. As noted in Ref. 3, for the VMT to agree with the PN solar-system experiments, the (universal) asymptotic value of f, f_0, must lie sufficiently near f_c, the double pole of ω, in order that $\omega(f_0)>60$ [the residue of the pole is necessarily positive, so the possibility[3] $\omega(f_0)<-33$ need not be considered]. Thus, far from the star, f is near f_c, and it steadily draws nearer to it as ξ decreases. Therefore, ω steadily grows from its large asymptotic value $\omega(f_0)>60$ as one approaches and enters the star. Now Eq. (8) shows that for a given matter distribution $\phi_{,\alpha}$ is $O(1/\omega)$ (recall that $\Omega\approx\omega$ here). The same is true of $\phi_{,\alpha,\beta}$. Thus the derivative terms in the gravitational-field equations (6) are $O(1/\omega)$ and become negligible compared to $8\pi G_0\phi^{-1}T_\mu{}^\nu$ for large ω. In this same limit ϕ becomes constant. Thus the larger ω, the closer STT is to GR locally. In view of our previous conclusions we can plausibly conclude that a NS model in the VMT with $r<0$ and $q>0$ will resemble its GR counterpart closer than the corresponding Brans-Dicke model with $\omega=60$.

Now Hillerbrandt and Heintzmann[15] and Saenz[16] have shown that Brans-Dicke NS models with $\omega=6$ are already very close to the GR models. The difference in critical mass being only 5–10%, for instance. For models with $\omega=60$ the agreement should be considerably closer, and for the VMT models even more so. Therefore, in view of the uncertainties in the nongravitational physics of pulsars[17] and neutron-star x-ray sources,[13] one can state with confidence that observations of neutron stars will not be able to discriminate between the VMT, with $r<0$ and $q>0$, and GR in the foreseeable future.

We cannot draw similar conclusions for other viable regions of the rq plane, especially $0<r<1$ and $q<0$ [for which the pole of ω (17) is also in the physical range $f>0$], because in that case f tends to recede from f_c as ξ decreases. This, however, is of little consequence since only for $r<0$ and $q>0$ can one understand why f_0 is near f_c in a cosmological context (see Sec. VI). Let us thus turn our attention to cosmology.

V. VMT ISOTROPIC COSMOLOGICAL MODELS

As a first step in the demonstration that for $r<0$ and $q>0$ one expects f_0 to be near f_c as required in the previous section and in the considerations of Ref. 3, we study some general properties of VMT isotropic cosmological models. We take the metric in Robertson-Walker form

$$ds^2 = -dt^2 + a(t)^2[(1-k\xi^2)^{-1}d\xi^2 + \xi^2(d\theta^2+\sin^2\theta\, d\Phi^2)]\,, \tag{18}$$

where $k=0$, ±1 is the curvature index. The scalar equation (8) takes the form

$$(a^3\Omega^{1/2}\dot\phi)^{\cdot} = -8\pi G_0 a^3\Omega^{-1/2}T\,, \tag{19}$$

where a dot signifies a time derivative. We shall find it useful to write this equation in integrated form:

$$\dot\phi = -8\pi G_0 a^{-3}\Omega^{-1/2}\left(\int_0^t Ta^3\Omega^{-1/2}dt + C\right), \tag{20}$$

where $t=0$ represents the beginning of the expansion, and C is an integration constant. The field equations (6) with $\mu=\nu=0$ and $\mu=\nu=\xi$ give, respectively,

$$\left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G_0\rho}{3\phi} + \frac{\omega}{6}\left(\frac{\dot\phi}{\phi}\right)^2 - \frac{\dot a\dot\phi}{a\phi}\,, \tag{21}$$

$$2\frac{\ddot a}{a} + \left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = -\frac{8\pi G_0 p}{\phi} - \frac{\omega}{2}\left(\frac{\dot\phi}{\phi}\right)^2 - \frac{\ddot\phi}{\phi} - 2\frac{\dot a\dot\phi}{a\phi} \tag{22}$$

where $\rho = -T_t{}^t$ and $p = T_\xi{}^\xi$ are the average density and pressure of the fluid filling the universe. We assume both are positive. Equations (3)–(4) and (20)–(22) are the equations of VMT isotropic cosmology. In view of their complexity we shall be content to establish some general results without delving into the exact form of the solutions.

In astrophysics an important role is played by the local Newtonian gravitational constant G_N. For the STT the standard PN analysis gives[4]

$$G_N = G_0\phi^{-1}(1+\Omega^{-1})\,, \tag{23}$$

where in the present context ϕ and Ω are just the cosmological quantities determined by Eq. (20). It is not immediately clear that G_N will always be positive. A negative G_N would have unpleasant consequences: galaxies might not form and clusters of galaxies might become unbound due to the effectively repulsive character of gravity. It is thus satisfying that one can show that for a VMT model capable of describing the present epoch, $G_N > 0$ throughout the expanding era of the universe.

First, it is clear that today $G_N > 0$ (gravitation is locally attractive). From PN solar-system experiments we know[3] that either $\Omega > 123$ or $\Omega < -63$. Thus $G_N \approx G_0\phi^{-1}$, and since G_0 is positive from first principles,[3] it follows that $\phi > 0$ today. Suppose ϕ is negative at some other time. Then it must pass through zero at some $t = t_1$ [a change in sign through infinity is ruled out by the form of ϕ, Eq. (3), and the condition $f > 0$]. We must have $\dot{\phi}(t_1) \neq 0$. Suppose that $\dot{\phi}(t_1)$ were to vanish. Then (20) could be written as

$$\dot{\phi} = -8\pi G_0 a^{-3}\Omega^{-1/2}\int_{t_1}^{t} Ta^3\Omega^{-1/2}dt\,. \tag{24}$$

Since $T = -\rho + 3p < 0$ we see that $\dot{\phi}$ would change from negative to positive at $t = t_1$, i.e., ϕ would have a minimum at $t = t_1$ and thus could not change sign there as assumed. Thus $\dot{\phi}(t_1) \neq 0$. Finally, we note that as a function of f, ω changes sign through zero at $f = q^{-1}$, precisely the point where ϕ changes sign. (The only exception to this statement occurs for $12rq = 1$; this case of the VMT was ruled out in Ref. 3 and is therefore, uninteresting.)

Now we focus attention on (21). For the expanding era, $\dot{a} \neq 0$ and the right-hand side changes sign when ϕ changes sign. For $k = 0$, +1 the left-hand side is strictly positive. Thus ϕ cannot change sign throughout the expansion era. For $k = -1$ it would appear that such a change in sign is possible provided $\dot{a}$ changes from less than unity to more than unity, or vice versa, at $t = t_1$. But this would mean that $\ddot{a}$ does not change sign at $t = t_1$. Now let us subtract (21) from (22) and replace $\ddot{\phi}/\phi$ by its value determined by (19):

$$2\frac{\ddot{a}}{a} = -\frac{8\pi G_0}{\phi}\left[\frac{\rho - 3p}{\Omega} + \frac{\rho + 3p}{3}\right] - \left(\frac{\omega}{6}\right)\left[4 - \frac{6\omega'\phi}{\Omega\omega}\right]\left(\frac{\dot{\phi}}{\phi}\right)^2 + 2\left(\frac{\dot{a}}{a}\right)\left(\frac{\dot{\phi}}{\phi}\right) \tag{25}$$

We recall that ω changes sign with ϕ. A check of (4) shows that $\omega'\phi$ does likewise. By contrast Ω does not. Thus the right-hand side of (25) changes sign if ϕ changes sign while the left-hand side cannot change sign by our previous comment. The contradiction shows that even for $k = -1$ ϕ keeps its (positive) sign throughout the expansion era. (We note that our proof is inapplicable to Brans-Dicke theory for which ω cannot change sign.)

Since $\phi > 0$, it follows that $qf < 1$. A look at (4) shows that $\Omega > 0$, since $f > 0$ always. Then it follows from (23) that $G_N > 0$. Thus, in any VMT cosmological model which describes correctly the present epoch (G_N positive today), G_N will be positive throughout the expansion era. This is an attractive feature of the VMT. (Incidently, our conclusion that $\Omega > 0$ rules out the possibility[3] that $\Omega < -63$ today.)

A further attractive property of the VMT is that it possesses nonsingular cosmological solutions in contrast to GR which has none. For the metric (18) the cosmological GR equations can be formally obtained by setting $\phi = 1$ in (21) and (22). Subtracting one equation from the other we get

$$2a^{-1}\ddot{a} = -\tfrac{8}{3}\pi G_0(\rho + 3p)\,. \tag{26}$$

We see that $\ddot{a} < 0$ always. Thus a cannot have a minimum which would imply $\ddot{a} > 0$. Neither can a be bounded from below without having a minimum; this would imply that $\ddot{a}$ becomes positive asymptotically at early or late time. We conclude that a must pass through zero at some finite t giving rise to a singularity. Thus, in GR, nonsingular solutions do not exist.

In the VMT the story is different. The analog of Eq. (26), Eq. (25), shows that $\ddot{a}$ could have either sign. Thus nonsingular cosmological solutions exist.

For example, suppose it is desired to construct a model which expands from a minimum at $t = 0$ [$\dot{a}(0) = 0$], at which time ρ and p have predetermined values appropriate to very hot matter (hot big bang). It is well-known that for such matter $0 < \rho - 3p \ll \rho + 3p \leq 2\rho$. According to (21) we have, at $t = 0$,

$$\frac{\omega}{6}\left(\frac{\dot{\phi}}{\phi}\right)^2 = -\frac{8\pi G_0\rho}{3\phi} + \frac{k}{a^2}\,. \tag{27}$$

Since we want $\phi > 0$ we must arrange for the term in (25) proportional to $\dot{\phi}^2$ to be positive and dominate the other terms at $t = 0$ so that $\ddot{a} > 0$. For $0 < f < q^{-1}$ we have $\Omega > 0$; in addition $\omega' > 0$ for the interesting cases $q > 0$ $r < 0$. For a wide range of values of $f(0)$ near $f = 0$ one finds, either analytically or numerically, that $\omega < 0$ while Ω is not small. One can then use (27) to determine $\dot{\phi}(0)$ for a choice of f which makes $\omega < 0$ [in the case $k = +1$ one chooses $a(0)$ sufficiently large]. Comparison of (25) with (27) shows that the positive term in (25) will definitely dominate the negative matter dependent terms. Thus one can set initial conditions for our nonsingular model.

It is worthwhile noting that even though a nonsingular model can be constructed for the given initial conditions, it is not clear *a priori* that it will be capable of describing the real universe. This will become clear only after the model is integrated numerically. We intend nothing so complicated here. However, there is one aspect of the fit between model and reality which can be investigated readily. To this we turn now.

VI. THE CONVERGENCE TO GR

The PN predictions of the VMT are determined by two parameters which are calculated by the standard procedure[4]

$$\gamma=(\Omega-1)(\Omega+1)^{-1}, \tag{28}$$

$$\beta=1+(\Omega+1)^{-2}\Omega^{-1}\omega'\phi, \tag{29}$$

where ϕ and $\omega(\phi)$ refer to the cosmological values (asymptotically far from the solar system). Agreement with solar-system experiments can be obtained for a large part of the rq plane, principally for $0<r<1$, $q<0$ and for $r<0$, $q>0$.[3] However, it is essential that the cosmological value of f be in a narrow range about the point $f_c=r(r-1)^{-1}q^{-1}$ which marks the pole of Ω.[3] The question before us is, how likely is a cosmological model with arbitrary (and unknown) initial conditions to produce today an f in the required range? To answer this question we focus on (20) which we rewrite as an equation for f:

$$\dot{f}=\frac{8\pi G_0\sqrt{2}f^{r+1/2}}{a^3(1-qf+6q^2f)^{1/2}}\left[\frac{|r+(1-r)qf|}{r+(1-r)qf}\right] \times\left(\int_0^t Ta^3\Omega^{-1/2}dt+C\right). \tag{30}$$

The term in square brackets is just the sign of $(1-r)q(f-f_c)$. Now we enquire into the sign of C.

Let us define a time t_C by

$$\int_{t_C}^0 Ta^3\Omega^{-1/2}dt=C. \tag{31}$$

Then $\dot{f}(t)$ is proportional to an integral over t extending from t_C to t. On grounds of causality we expect that $t_C<t$ over the whole evolution of the universe (otherwise the dynamics of f at t will depend on the behavior of the matter for times later than t). This means $t_C\leq 0$. The value $t_C=0$ corresponding to $C=0$ seems appropriate to a model which begins from a singular state at $t=0$. A value $t_C<0$ corresponding to $C<0$ is more appropriate to a model having a minimum at $t=0$. Thus on physical grounds we conclude that $C\leq 0$.

It follows from all we have said that the sign of $\dot{f}$ is opposite that of $(1-r)q(f-f_c)$. For $0<r<1$ and $q<0$, $\dot{f}$ has the same sign as $f-f_c$. Thus, whether $f>f_c$ or $f<f_c$, f always recedes from f_c. For generic initial conditions f should be far from f_c by the present time. Thus the VMT with $0<r<1$ and $q<0$ fails to provide a self-consistent framework for understanding the results of the solar-system experiments. However, for $r<0$ and $q>0$, $\dot{f}$ has a sign opposite that of $f-f_c$. Thus, whether $f>f_c$ or $f<f_c$, f approaches f_c monotonically: for generic initial conditions f converges to f_c. If the time scale for convergence is sufficently short, f will be near f_c at present. Thus the VMT with $r<0$ and $q>0$ provides a self-consistent explanation of the results of the PN solar-system experiments.

The sucess of the VMT with $r<0$ and $q>0$ is not limited to this. We note that as $f\to f_c$, $\omega\to\infty$. Recalling the discussion in Sec. IV we see that for $r<0$ and $q>0$ the VMT cosmological models themselves (the run of a with time, for instance) spontaneously converge to the corresponding GR models at late times. This is one more example of the merging of the predictions of VMT and GR that we noticed in Secs. III and IV. From now on we shall assume $r<0$ and $q>0$.

We return now to the question of the time scale for f to approach f_c. Without integrating (30) and (21) numerically we cannot be certain that the approach takes place fast enough to be relevant. But supposing there are models for which it does, we face a potential difficulty. If the time scale for approach is short compared to the Hubble time t_H (expansion time scale of the universe), we might expect that the time scale for significant change in G_N is equally short compared to t_H in disagreement with experiment.[1] To show that this need not be the case, we differentiate (23) with respect to time to get

$$\dot{G}_N/G_N=-[1+2\Omega^{-1}(\Omega+1)^{-1}\omega'\phi]\dot{\phi}/\phi. \tag{32}$$

The experiments[1] place constraints on $\dot{G}_N/G_N$ or, equivalently, on the time scale for variations of G_N, $t_G\equiv G_N/\dot{G}_N$ (t_G can be positive or negative). Let us eliminate $\omega'\phi$ from (32) by means of (29) and replace $\dot{\phi}$ by the expression in terms of $\dot{f}$. We get

$$t_G=-\frac{(1-r)qf(1-qf)}{[r+(1-r)qf]^2\Omega[\Omega^{-1}+2(\beta-1)(1+\Omega^{-1})]}t_f, \tag{33}$$

where $t_f\equiv(f_c-f)/\dot{f}$ is the time scale for f to approach f_c. It is a positive quantity according to our previous discussion. Let us now substitute for Ω the expression following from (4), and for Ω^{-1} the expression $(1+\gamma)(1-\gamma)^{-1}$ which follows from (28). Our final result is

$$t_G=-\frac{2(1-r)q(1-qf)(1+\gamma)}{(1-qf+6q^2f)(4\beta-\gamma-3)}t_f. \tag{34}$$

The beauty of (34) lies in the fact that it relates t_G to $4\beta-\gamma-3$, the coefficient which is a measure of the Nordtvedt effect.[4] This is the "polarization" of the moon's orbit in the sun's direction which is expected in any STT. (For GR, $4\beta-\gamma-3=0$.) A search of this effect by laser ranging has been analyzed by two groups[1,18] leading to the results $4\beta-\gamma-3=-0.001\pm0.015$ and $4\beta-\gamma-3=0.00\pm0.03$, respectively. Since $\gamma\approx1$ [see (28)], we see that

$$2(1+\gamma)\,|4\beta-\gamma-3|^{-1}>133\,. \tag{35}$$

Now, for $|r|$ and q of order unity (which means f_C and f are of order unity), the other factors in (34) are also of order unity. Thus $|t_G|$ can well be two orders of magnitude larger than t_f. Hence, the agreement of the VMT with the PN experiments at the present epoch does not clash with the absence of variations of G_N on time scales shorter than t_H.[1]

The sign of $\dot{G}_N$ can also be deduced. According to (20) the sign of $\beta-1$ is the same as that of $\omega'\phi$. A direct calculation gives

$$\omega'\phi=\tfrac{1}{4}(1-qf)[(1+r-12rq)qf-r][r+(1-r)qf]^{-4}\,. \tag{36}$$

It is easily verified that for $r<0$ and $q>0$, $\omega'\phi>0$ for $f\approx f_c$. Returning to (29) we see that according to the VMT β is slightly greater than unity [recall that $r+(1-r)qf$ is small for $f\approx f_c$]. Also, according to (28) γ is slightly smaller than unity. Thus $4\beta-\gamma-3$ must be small and positive: the VMT, like Brans-Dicke theory, predicts that there is a Nordtvedt effect, and that its coefficient is positive. This is an experimentally testable prediction. Since $\phi>0$, $1-qf>0$. Consulting (34) we see that for $r<0$ and $q>0$, $t_G<0$. Thus, the VMT, just like Brans-Dicke theory, predicts that G_N is presently decreasing. Further, the VMT reveals a relation between the magnitude of the Nordtvedt effect and t_G which may have deep significance.

Van Flandern and Muller[2] have presented evidence suggesting that G_N is decreasing on a time scale t_H. It is too early to draw conclusions from these preliminary investigations. However, if this result is confirmed by future work, GR will be ruled out, and the VMT with $r<0$ and $q>0$ would be its natural replacement. It is also clear that if future investigations reveal that G_N is increasing, the VMT itself will prove unviable.

VII. SUMMARY AND CONCLUSIONS

The VMT presented in Ref. 3 is also the most general genuine theory of variable gravitational constant (Brans-Dicke theory can be regarded as a special case). The VMT retains all elements of GR except for the experimentally unverified assumption that the gravitational coupling constant $Gm^2/\hbar c$ is a space time constant. Because of this the VMT seems more appropriate than GR as a working theory by "Occam's razor." The matter-free black-hole solutions of both theories are identical. Neutron-star models in the VMT with $r<0$ and $q>0$ are very similar to their GR counterparts. For the same range of r and q all VMT cosmological models converge to the corresponding GR ones at late times. Related to this is the convergence of the PN predictions of the VMT to those of GR. Because of all this, the VMT with $r<0$ and $q>0$ is today as experimentally viable as GR.

Like GR the VMT predicts that gravitation has been attractive throughout the expansion of the universe. Unlike GR the VMT possesses nonsingular cosmological solutions. Conditional on a demonstration that some of these solutions are realistic, this property makes the VMT all the more attractive than GR. By contrast to GR, the VMT with $r<0$ and $q>0$ predicts that the Nordtvedt effect coefficient is positive, and that G_N is presently decreasing on a time scale which could be long compared to t_H. According to the VMT the two effects are connected. Verification of these predictions would rule out GR and would point to the VMT with $r<0$ and $q>0$ as its most natural replacement. One would understand the success of GR in other respects by virtue of the tendency of the VMT to converge to GR. By virtue of this one could continue to use GR for most purposes despite its being incorrect as a theory of gravitation.

Let us also contrast the VMT with Brans-Dicke theory. In the latter agreement with the PN experiments is secured by setting its *parameter* ω large by hand. In the VMT with $r<0$ and $q>0$, ω is a dynamical quantity which, whatever its initial value, spontaneously grows large as the universe expands leading to agreement with the PN experiments.

---

[1] I. I. Shapiro, C. C. Counselman III, and R. W. King, Phys. Rev. Lett. 36, 555 (1976).
[2] T. C. van Flandern, Mon. Not. Astron. Soc. 170, 33 (1975); P. M. Muller, Jet Propulsion Laboratory Internal Report, 1976 (unpublished).
[3] J. D. Bekenstein, Phys. Rev. D 15, 1458 (1977).
[4] C. M. Will, in *Experimental Gravitation*, edited by B. Bertotti (Academic, New York, 1974). There is a wrong sign in Eq. (5.34) of this reference.
[5] P. G. Bergmann, Int. J. Theor. Phys. 1, 25 (1968).
[6] R. V. Wagoner, Phys. Rev. D 1, 3209 (1970).
[7] R. H. Dicke, Phys. Rev. 125, 2163 (1962).

[8]C. Brans and R. H. Dicke, Phys. Rev. 124, 925 (1961).
[9]S. W. Hawking, Commun. Math. Phys. 25, 167 (1972).
[10]J. D. Bekenstein, Phys. Rev. D 5, 2403 (1972).
[11]S. W. Hawking and J. B. Hartle, Commun. Math. Phys. 27, 283 (1972).
[12]S. W. Hawking, Commun. Math. Phys. 25, 152 (1972).
[13]R. P. Kraft, in *Neutron Stars, Black Holes and Binary X-Ray Sources*, edited by H. Gursky and R. Ruffini (Reidel, Dordrecht, 1975).
[14]J. D. Bekenstein, Ann. Phys. (N.Y.) 91, 75 (1975).
[15]W. H. Hillerbrandt and H. Heintzmann, Gen. Relativ. Gravit. 5, 663 (1974).
[16]R. A. Saenz, Astrophs. J. 212, 816 (1977).
[17]G. Baym and C. Pethick, Annu. Rev. Nucl. Sci. 25, 27 (1975).
[18]J. G. Williams *et al.*, Phys. Rev. Lett. 36, 551 (1976).