THE ASTROPHYSICAL JOURNAL, 286:7–14, 1984 November 1

DOES THE MISSING MASS PROBLEM SIGNAL THE BREAKDOWN OF NEWTONIAN GRAVITY?

JACOB BEKENSTEIN
Department of Physics, Ben Gurion University of the Negev, Beer-Sheva

AND

MORDEHAI MILGROM[1]
Department of Physics, Weizmann Institute of Science, Rehovot
Received 1984 March 28; accepted 1984 May 17

ABSTRACT

We consider a nonrelativistic potential theory for gravity which differs from the Newtonian theory. The theory is built on the basic assumptions of the modified dynamics, which were shown earlier to reproduce dynamical properties of galaxies and galaxy aggregates without having to assume the existence of hidden mass. The theory involves a modification of the Poisson equation and can be derived from a Lagrangian. The total momentum, angular momentum, and (properly defined) energy of an isolated system are conserved. The center-of-mass acceleration of an arbitrary bound system in a constant external gravitational field is independent of any property of the system. In other words, all isolated objects fall in exactly the same way in a constant external gravitational field (the weak equivalence principle is satisfied). However, the internal dynamics of a system in a constant external field is different from that of the same system in the absence of the external field, in violation of the strong principle of equivalence. These two results are consistent with the phenomenological requirements of the modified dynamics. We sketch a toy relativistic theory which has a nonrelativistic limit satisfying the requirements of the modified dynamics.

Subject headings: cosmology — galaxies: internal motions — gravitation

I. INTRODUCTION

The direct application of Newtonian dynamics to astrophysical objects on the scale of galaxies and galaxy aggregates produces a paradox. Galactic rotation curves do not show the expected Keplerian falloff, virial masses for clusters of galaxies much exceed mass estimates from luminosities, etc. The conventional reply to these problems is that there are large quantities of "dark mass" in galaxy systems. However, it is also possible to cast the blame for the aforementioned discrepancies on the Newtonian theory.

Over the years, deviations from Newtonian behavior in the nonrelativistic limit have been discussed in the literature on various occasions. Some recent references can be found in Will (1979). Particularly interesting in the present context is Newcomb's suggestion to modify Newtonian gravity to explain the excess perihelion shift of Mercury (as described in Weinberg 1972) as an alternative to effects of hidden matter. Such deviations were also brought up as possible explanations of the mass discrepancy in clusters and galaxies (e.g., Finzi 1963; Tohline 1983). As far as we know, all these suggested modifications can be described, in the nonrelativistic case, as modifications of the distance dependence of the gravitational field and are ruled out as the sole explanation of the mass discrepancy (see Paper II).

It has been suggested (Milgrom 1983*a*, *b*, *c*, 1984, hereafter Papers I, II, III, IV, respectively) that if the motion of objects within galaxies, and of galaxies within groups and clusters, are described by a certain modified form of nonrelativistic dynamics (MOND), there is no need to assume the existence of hidden mass in appreciable quantities in these systems. In addition, many of the observed properties of galaxies follow as unavoidable consequences of MOND.

The minimal set of assumptions, on which practically all of the results of Papers I–IV were based, is:

a) A breakdown of Newtonian dynamics (second law and/or gravity) occurs in the limit of small accelerations.

b) In this limit the acceleration, $\boldsymbol{a}$, of a test particle in a gravitating system is given by $\boldsymbol{a}(a/a_0) \approx \boldsymbol{g}_N$, where $\boldsymbol{g}_N$ is the conventional gravitational field and a_0 is a constant with the dimensions of acceleration.

c) The transition from the Newtonian regime to the small acceleration asymptotic region occurs within a range of order a_0 about a_0. The value of a_0 was determined (Paper II) to be approximately $2 \times 10^{-8}(H_0/50 \text{ km s}^{-1} \text{ Mpc}^{-1})^2$ cm s^{-2} (which approximately equals cH_0).

It is important to realize that more than one interpretation can be given to this set of assumptions, and that more than one detailed scheme which embodies these assumptions will work successfully at the present stage of the analysis.

The scheme used in Papers I–IV can be described in either of the following two ways. A modification of the law of inertia

$$m\mu(a/a_0)\boldsymbol{a} = \boldsymbol{F} , \tag{1a}$$

where $\boldsymbol{F}$ is an arbitrary static force field assumed to depend on its sources in the conventional way and m is the gravitational mass of the accelerated test particle. For gravity $\boldsymbol{F} = m\boldsymbol{g}_N$, where $\boldsymbol{g}_N = -\nabla\varphi_N$ and φ_N is the gravitational potential deduced in the usual manner from the Poisson equation. Alternatively, the MOND can be described as a modification of gravity leaving the law of inertia ($m\boldsymbol{a} = \boldsymbol{F}$) intact but such that for gravity, $\boldsymbol{F} = m\boldsymbol{g}$ and $\boldsymbol{g}$ is a modified gravitational field derived from $\boldsymbol{g}_N$ using the relation

$$\mu(g/a_0)\boldsymbol{g} = \boldsymbol{g}_N . \tag{1b}$$

[1] Supported in part by a grant from the MINERVA Foundation.

The formulation expressed by equation (1a) is equivalent to that of (1b) when only gravitational forces are present, but the two are, of course, not the same in general (see discussion in Paper I). The function μ satisfies $\mu(x \gg 1) \approx 1$, so that Newtonian dynamics is recovered in the limit of large accelerations, and $\mu(x \ll 1) \approx x$. Otherwise μ remained unspecified except that it was assumed to be monotonic in its argument. Methods for determining μ observationally were discussed in Paper II.

In cases of high symmetry (spherical, plane, or cylindrical), $\boldsymbol{g}$ as given by equation (1b) is derivable from a potential $\varphi(\boldsymbol{g} = -\nabla\varphi)$; however, in general, it is not.

As explained in Paper I, the formulation given by equation (1) cannot be considered a theory, but only a successful phenomenological scheme for which an underlying theory is needed. To avoid problems with momentum conservation, the use of equation (1) was restricted to the description of the motion of light objects in the static mean field of a massive body. If one tries to apply this equation to describe an arbitrary N-body system, one encounters various problems. In particular, momentum is not conserved. This fact has been brought up by many colleagues as an argument against MOND (see, e.g., Felten 1984).

In this paper we present a model theory (derivable from a Lagrangian) which satisfies the basic requirements of MOND listed above. In this theory a_0 and μ are put in by hand; the theory thus does not make use of the near equality of a_0 and cH_0. The theory demonstrates that MOND can be formulated in a fashion which is internally consistent, and which is consistent with the usual conservation laws. We hope that it will be a useful step in the development of a more complete theory in which perhaps μ and a_0 will be expressible in terms of more fundamental entities.

Even in its present form, the theory makes important steps beyond the phenomenological scheme presented in Papers I–II. Besides the usual advantages which a Lagrangain theory offers (e.g., conservation laws), it enables one to calculate (at least in principle) the dynamics of an arbitrary, nonrelativistic gravitating system. In particular, this theory can be shown to justify two important working assumptions which were made in Papers I–IV on a phenomenological basis (see §§ IV, V below):

1. A composite particle (say a star or a cluster of stars) moving in an external field, say of a galaxy, moves *like a test particle* according to the MOND rules, even if within it the relative accelerations are large. This is the case, provided the mass of the particle is small compared with that of the galaxy.

2. When a system is accelerated as a whole in an external field, the internal dynamics of the system is affected by the presence of the external field (even when this is constant so that no tidal effects are present). In particular, in the limit when the external (center-of-mass) acceleration of the system becomes much larger than a_0, the internal dynamics approaches exact Newtonian behavior even when accelerations within the system are much smaller than a_0.

Newtonian gravity is recovered at the nonrelativistic limit of general relativity, and of a number of other relativistic theories of gravity. We, however, need a new relativistic theory of gravity which yields various aspects of MOND in the appropriate limit. This is no trivial task, for there is no lore on the construction of theories with the specific non-Newtonian limit we seek. Such a theory is essential for two main reasons: (*a*) to help incorporate the principles of MOND into the framework of modern theoretical physics; (*b*) to provide tools for investigating cosmology in light of MOND. This last is particularly pressing since some of the arguments adduced in support of the MOND from the empirical viewpoint (in Paper III) have a cosmological aspect which cannot be treated self-consistently without a relativistic cosmology.

In Appendix B, we describe briefly a relativistic theory the nonrelativistic limit of which (in a static universe) is basically the same as the theory which is the subject of the present paper. So far we have studied some of the properties of this theory superficially, and we present it here only to demonstrate that the requirements of the MOND can form the basis of a relativistically invariant theory.

In § II we present the nonrelativistic theory. In § III we discuss conservation laws. In § IV we consider the center of mass motion of composite systems in an external field. Section V describes some of the possible influences of an external acceleration on the internal dynamics of a system. In § VI we discuss miscellaneous points and summarize our conclusions.

II. THE FIELD EQUATIONS

In Newtonian gravity test bodies move with an acceleration equal to $\boldsymbol{g}_{\mathrm{N}} = -\nabla\varphi_{\mathrm{N}}$, where φ_{N} is the Newtonian gravitational potential. It is determined by the Poisson equation $\nabla^2\varphi_{\mathrm{N}} = 4\pi G\rho$, where ρ is the mass density which produces φ_{N}. The Poisson equation may be derived from the Lagrangian

$$L_{\mathrm{N}} = -\int d^3r\{\rho\varphi_{\mathrm{N}} + (8\pi G)^{-1}(\nabla\varphi_{\mathrm{N}})^2\} . \tag{2a}$$

In searching for a modification of this theory we will want to retain the notion of a *single* potential φ from which acceleration derives. And, as in Newtonian gravity, it is desirable that φ be arbitrary up to an additive constant. The most general modification of L_{N} which will yield these features is

$$L = -\int d^3r\left\{\rho\varphi + (8\pi G)^{-1}a_0^2\mathscr{F}\left[\frac{(\nabla\varphi)^2}{a_0^2}\right]\right\} , \tag{2b}$$

where $\mathscr{F}(x^2)$ is an arbitrary function. Note that a scale of acceleration is necessary unless we are in the Newtonian case.

Variation of L with respect to φ with variation of φ vanishing on the boundary yields

$$\nabla \cdot [\mu(|\nabla\varphi|/a_0)\nabla\varphi] = 4\pi G\rho , \tag{3}$$

with $\mu(x) = \mathscr{F}'(x^2)$, as the equation determining the modified potential. A test particle is assumed to have acceleration $\boldsymbol{g} = -\nabla\varphi$. We supplement equation (3) by the boundary condition $|\nabla\varphi| \to 0$ as $r \to \infty$.

It is useful to write the field equation in terms of the unmodified Newtonian field $\boldsymbol{g}_{\mathrm{N}} = -\nabla\varphi_{\mathrm{N}}$, for the same mass distribution, which satisfies the Poisson equation. By eliminating ρ we get

$$\nabla \cdot [\mu(\nabla\varphi/a_0)\nabla\varphi - \nabla\varphi_{\mathrm{N}}] = 0 . \tag{4}$$

The expression in parentheses in equation (4) is thus a curl field, and we may write

$$\mu(g/a_0)\boldsymbol{g} = \boldsymbol{g}_{\mathrm{N}} + \nabla \times \boldsymbol{h} . \tag{5}$$

We are now ready to make contact with equation (1) and show that the present theory satisfies the basic assumptions of the MOND. If the curl term in equation (5) can be neglected in a certain region of space, the acceleration of test particles as given by the field equation is the same as that given by equation (1). We shall show in Appendix A that at large distances from a bound object of total mass M the curl term in equation (5) decreases faster with r than the other two terms. We thus have in this limit

$$\mu(g/a_0)\boldsymbol{g} = -G\boldsymbol{r}/r^3 + O(r^{-3}) = \boldsymbol{g}_{\mathrm{N}} + O(r^{-3}) , \quad (6)$$

as in equation (1).

For our theory to satisfy the assumptions of MOND we identify μ in the field equation with that of equation (1). In particular, as we require $\mu(x) \approx x$ for $x \ll 1$; we find for an arbitrary bound system of mass M that

$$\boldsymbol{g} \xrightarrow{r\to\infty} -(MGa_0)^{1/2}\boldsymbol{r}/r^2 + O(r^{-2}) , \quad (7)$$

and thus, in this limit,

$$\varphi \to (MGa_0)^{1/2} \ln (r/r_0) + O(r^{-1}) , \quad (8)$$

where r_0 is an arbitrary radius. This potential leads to an asymptotically constant circular velocity $V_\infty = (MGa_0)^{1/4}$ as observed in the outskirts of spiral galaxies (see Paper II).

The field equation (3) is nonlinear and difficult to solve in general. However, in all cases of high symmetry (i.e., spherical, plane, or cylindrical symmetry) the curl term in equation (5) vanishes identically and we have the exact result $\mu(g/a_0)\boldsymbol{g} = \boldsymbol{g}_{\mathrm{N}}$ identical to equation (1b). Solving for φ in such cases is straightforward. This result can be readily obtained for example by applying Gauss's theorem to equation (4) on a surface of symmetry. Some of the important properties of the Newtonian field for systems of such symmetry can be carried over to the present theory. For example, the acceleration field at distance r from the center in a spherical system depends only on the total mass, $M(r)$, interior to r, and in fact is given by $\mu(g/a_0)\boldsymbol{g} = -M(r)G\boldsymbol{r}/r^3$.

The field equation (3) is analogous to the equation for the electrostatic potential in a nonlinear isotropic medium in which the dielectric coefficient is a function of the electric field strength.

It may also be useful to note that our field equation is equivalent to the stationary flow equations of an irrotational fluid which has a density $\hat{\rho} = \mu(|\nabla\varphi|/a_0)$, a negative pressure $\hat{P} = -2^{-1}a_0^2\mathscr{F}[(\nabla\varphi)^2/a_0^2]$, flow velocity $\hat{\boldsymbol{v}} = \nabla\varphi$, and a source distribution $\hat{S}(\boldsymbol{r}) = 4\pi G\rho$. The fluid satisfies an equation of state $\hat{P}(\hat{\rho}) = -2^{-1}a_0^2\mathscr{F}\{[\mu^{-1}(\hat{\rho})]^2\} \equiv f(\hat{\rho})$.

An equation of the same form as equation (3) has been studied in a different context to describe classical models of quark confinement using a very different form of the function μ at both large and small values of its argument (see Adler and Piran 1984 and also Lehman and Wu 1983 for a review).

III. THE CONSERVATION LAWS

The conservation laws follow from the symmetry of the Lagrangian under spacetime translations and space rotations. We find it instructive, however, to derive them explicitly from the field equation.

Let $\boldsymbol{g} = -\nabla\varphi$ be the modified acceleration field produced by a bound mass distribution $\rho(\boldsymbol{r})$. We consider the motion of a subsystem made of all the mass within the arbitrary Eulerian volume α (we take α to be such that ρ vanishes on its surface). Each infinitesimal mass element within α moves as a test particle, i.e., with an acceleration $\dot{\boldsymbol{v}} = -\nabla\varphi$ (for any field quantity q, $\dot{q} \equiv \partial_t q + \boldsymbol{v}\cdot\nabla q$.) Note that here we assume a pressureless system. The changes in the argument when pressure is important will be discussed briefly at the end of the section.

Using the continuity equation $\partial_t\rho + \nabla\cdot(\rho\boldsymbol{v}) = 0$ we can write the time derivative of any quantity Q of the form

$$Q = \int_\alpha d^3r q(\boldsymbol{r})\rho(\boldsymbol{r}) , \quad (9)$$

as

$$\dot{Q} = \int_\alpha d^3r \dot{q}(\boldsymbol{r})\rho(\boldsymbol{r}) , \quad (10)$$

provided there is no flux of q across the surface of α (for which it is sufficient to have $\rho q\boldsymbol{v}$ vanish everywhere on this surface).

The total momentum within α is

$$\boldsymbol{P} = \int_\alpha d^3r\boldsymbol{v}(\boldsymbol{r})\rho(\boldsymbol{r}) . \quad (11)$$

The center of mass (c.o.m.) velocity is $\boldsymbol{V} = \boldsymbol{P}/m$, where m is the total mass within α. The c.o.m. acceleration $\boldsymbol{A}$ is determined from equations (10, 11) and Euler's (momentum) equation to be

$$m\boldsymbol{A} = \dot{\boldsymbol{P}} = -\int_\alpha d^3r\rho(\boldsymbol{r})\nabla\varphi . \quad (12)$$

The right-hand side of equation (12) is the gravitational force $\boldsymbol{F}$ acting on α. Substituting from the field equation (3), we get

$$\dot{\boldsymbol{P}} = -(4\pi G)^{-1}\int_\alpha d^3r\nabla\varphi\nabla\cdot[\mu(\nabla\varphi/a_0)\nabla\varphi] . \quad (13)$$

Integrating by parts and then writing volume integrals of a gradient and a divergence as surface integrals, over the surface Σ of α, we find

$$4\pi G\dot{\boldsymbol{P}} = -\int_\Sigma \mu\nabla\varphi(\nabla\varphi\cdot d\boldsymbol{s}) + \frac{a_0^2}{2}\int_\Sigma \mathscr{F}\,d\boldsymbol{s} . \quad (14)$$

Note, in particular, that for an isolated system $\dot{\boldsymbol{P}} = 0$ because in the limit $r \to \infty$ the integrand in each of the terms in equation (14) decreases like r^{-3} (see eq. [7]). Thus the integrals themselves must vanish if the surface Σ surrounds all the mass in the system. In an isolated system made up of separate bodies the (vanishing) net force on the system is the sum of the forces on the parts (from eq. [12]). Hence "action equals reaction" in this theory.

The rate of change of the angular momentum,

$$\boldsymbol{J} = \int_\alpha d^3r\rho(\boldsymbol{r})\boldsymbol{r} \times \boldsymbol{v} , \quad (15)$$

is similarly given by

$$4\pi G\dot{\boldsymbol{J}} = -\int_\alpha \nabla\cdot(\mu\nabla\varphi)\boldsymbol{r} \times \nabla\varphi d^3r . \quad (16)$$

Again integrating by parts, we write $\dot{\boldsymbol{J}}$ as a surface integral:

$$4\pi G\dot{\boldsymbol{J}} = -\int_\Sigma \boldsymbol{r} \times \nabla\varphi\mu\nabla\varphi\cdot d\boldsymbol{s} + \frac{a_0^2}{2}\int_\Sigma \mathscr{F}\boldsymbol{r} \times d\boldsymbol{s} . \quad (17)$$

In obtaining equation (17) we made use of the identity

$$2(\nabla\varphi\cdot\nabla)(\boldsymbol{r} \times \nabla\varphi) = (\boldsymbol{r} \times \nabla)(\nabla\varphi)^2 , \quad (18)$$

which is straightforward to prove.

To show that $\dot{\boldsymbol{J}} = 0$ for an isolated system, we take a spherical surface centered at the origin and of radius $r \to \infty$. The second term in equation (17) vanishes since $d\boldsymbol{s}$ is parallel to $\boldsymbol{r}$. The integrand in the first term decreases at least as fast as r^{-3}

by equation (7). Thus, the integral must vanish for a surface surrounding all the mass in the system (and sum of torques on separate bodies within an isolated system thus vanish).

Consider now the energy of an isolated self-gravitating system. It would be natural to define the total energy of the system as $E = -L + E_K$, where the kinetic energy is

$$E_K = \tfrac{1}{2} \int d^3 r v^2 \ . \tag{19}$$

Note, however, that here the Lagrangian, formally defined in equation (2b), diverges logarithmically because $\mathscr{F}[(\nabla\varphi)^2/a_0^2] \rightarrow Ar^{-3}$ as $r \rightarrow \infty$. This divergence reflects the fact that the potential is confining. The binding energy of any system of finite mass is thus logarithmically divergent (compared with the energy of the system with the different elements infinitely apart), and so also is the difference in total energy between two systems with different total masses. We can thus speak only of the energy difference of two systems *with the same total mass* (see also the discussion of the energy of isothermal spheres in Paper IV). To make L finite it suffices to subtract from the integrand in equation (2b) its expression for an arbitrary fixed system (such as that of an homogeneous sphere of finite radius) and a given mass M. The Lagrangian then describes only systems with a fixed total mass M.

Consider an infinitesimal change $\delta\rho$ in the density of a system, such that $\delta M = \int_\alpha \delta\rho d^3 r = 0$. This change induces a change $\delta\varphi$ in φ, through the field equation. We have for the change in L: $\delta L = \delta L_\varphi + \delta L_\rho$, where the first term results from changing only φ and the second from changing only ρ in the expression for L. Because $\delta M = 0$, equation (8) implies that $\delta\varphi$ vanishes at infinity. For such a change $\delta\varphi$ about a φ which is a solution of the field equation, L is stationary. Thus $\delta L_\varphi = 0$, and we have

$$\delta L = -\int \varphi \delta\rho d^3 r \ . \tag{20}$$

We now show that the total energy E, as defined above, is conserved, for an isolated self-gravitating system. We describe the system, as we did earlier, as a continuous fluid of density $\rho(\boldsymbol{r}, t)$ and velocity $\boldsymbol{v}(\boldsymbol{r}, t)$. The result of the previous paragraph tells us (see eq. [20]) that $\dot{L} = -\int_\alpha \varphi \partial\rho/\partial t d^3 r$. On the other hand (using eq. [10]),

$$\begin{aligned} \dot{E}_K &= \tfrac{1}{2} \int_\alpha \rho (\dot{v^2}) d^3 r = -\int_\alpha d^3 r \rho \boldsymbol{v} \cdot \nabla\varphi \\ &= -\int_\alpha d^3 r \{\nabla \cdot (\varphi\rho\boldsymbol{v}) - \varphi \nabla \cdot (\rho\boldsymbol{v})\} \ . \end{aligned} \tag{21}$$

The first term can be written as a surface integral and vanishes, and the second can be written via the continuity equation as $-\int_\alpha \varphi \partial\rho/\partial t d^3 r$. We thus have $\dot{L} = \dot{E}_K$ or $\dot{E} = 0$.

If we wish to account for some of the kinetic energy by internal energy and include the effects of pressure, we must use for the momentum equation of motion $\dot{\boldsymbol{v}} = -\nabla\varphi - \rho^{-1}\nabla p$. In the expression for $\dot{\boldsymbol{P}}$ (eq. [12]) the pressure term becomes a surface integral $\int_\Sigma p\boldsymbol{ds}$, which is just the external pressure force on the system (which vanishes for an isolated system). Similarly, in the equation for $\dot{\boldsymbol{J}}$ we get an additional term which amounts to the external moment $\int_\Sigma p\boldsymbol{ds} \times \boldsymbol{r}$. In equation (21) for $\dot{E}_K$ we get two additional terms: a surface work term $\int_\Sigma p\boldsymbol{v} \cdot \boldsymbol{ds}$ and a term $\int_\alpha p\rho(\dot{\rho^{-1}}) d^3 r$ which is just $-\dot{E}_{\text{in}}$, where E_{in} is the internal energy (we assume no heat transfer). Hence the total energy $E_K + E_{\text{in}} - L$ is conserved if there is no surface work.

Thus we find that in the proposed non-Newtonian gravitation theory all the usual conservation laws hold.

IV. THE CENTER-OF-MASS MOTION OF BODIES IN AN EXTERNAL FIELD

The modified dynamics was used in Papers II–IV to calculate rotation curves of galaxies, to derive masses of galaxies and galaxy systems, to calculate the structure of isothermal spheres, etc. In doing so it was assumed that the various bodies the motion of which was considered, such as stars in the field of a galaxy or galaxies accelerated in the field of a cluster, have the same acceleration as does a test particle. Observational evidences supporting the validity of this assumption, to a good an approximation, were discussed in Paper I.

Here we show that in the framework of the present theory it is justified to use the test-particle equation of motion to determine the acceleration of the objects mentioned above to high accuracy. We consider the center-of-mass acceleration of a body of total mass m (which for convenience we shall refer to as the "star") in the presence of some large mass M (the "galaxy"). We seek to show that the acceleration of the "star" is independent of its internal structure to so good an approximation as to justify the procedures used in Papers II–IV. We shall assume in what follows that the acceleration field of the "galaxy" alone can be taken as constant across the "star" (no tidal effect).

We have shown above that an isolated object does not accelerate itself [i.e., it has A(c.o.m) = 0]. In a linear theory, such as Newtonian dynamics, this fact is sufficient to make the acceleration of any object in the field of any mass independent of the object's structure. Our proposed theory is nonlinear; in fact, in the general case the acceleration of an object in a given field does depend on its mass and structure, as the following example will show. Let M and m be two point masses ($m \ll M$) isolated from other masses. We show below that m moves (with acceleration a_m) in the field of M very nearly like a test particle, i.e., $\mu(a_m/a_0)\boldsymbol{a}_m = -MG\boldsymbol{R}/R^3$, where $\boldsymbol{R}$ is the radius vector from M to m. From momentum conservation, the acceleration $\boldsymbol{a}_M$ of M is given by $\boldsymbol{a}_M = -\boldsymbol{a}_m(m/M)$. This acceleration is, however, different from the acceleration, $\boldsymbol{a}_t$, of a test particle produced by m alone, at the position of M. The last acceleration is given by $\mu(a_t/a_0)\boldsymbol{a}_t = +mG\boldsymbol{R}/R^3$.

We now show, however, that as long as the mass of the "star" is much smaller than that of the "galaxy," the acceleration of the "star" does not depend on whether the intrinsic accelerations within the "star" are large or small (compared with a_0) or, for that matter, on any property of the "star."

Consider first the case where the "star" can be thought of as being in a *constant* acceleration field $-\nabla\varphi_g$ of the "galaxy." It is thus assumed that in the absence of the "star" $\nabla\varphi$ is exactly constant everywhere in space (not just across the "star" as we assume anyway) and equals $\nabla\varphi_g$.

We start with the expression for the c.o.m. acceleration as a surface integral, equation (14). Denote the first force term in equation (14) by $\boldsymbol{F}_1$ and the second by $\boldsymbol{F}_2$. We can write $\boldsymbol{F}_1$ as $\boldsymbol{F}_1 = \boldsymbol{F}_{11} + \boldsymbol{F}_{12}$ with

$$\begin{aligned} 4\pi G\boldsymbol{F}_{11} &= -\nabla\varphi_g \int_\Sigma \mu\nabla\varphi \cdot \boldsymbol{ds} \\ &= -\nabla\varphi_g \int_\alpha d^3 r \nabla \cdot (\mu\nabla\varphi) = -4\pi G m \nabla\varphi_g \ , \end{aligned} \tag{22}$$

and

$$4\pi G\boldsymbol{F}_{12} = -\int_{\Sigma} \nabla\delta\mu\nabla\varphi \cdot d\boldsymbol{s} . \tag{23}$$

Here $\nabla\varphi$ is the acceleration field produced by the "galaxy" plus "star," and $\nabla\delta = \nabla\varphi - \nabla\varphi_g$. The integration is over a volume α which contains all the mass in the star and no other mass.

We now make the radius of the integration surface go to infinity. In this limit, $|\nabla\delta| \rightarrow 0$ on the surface, and we retain only the term of lowest order in $\nabla\delta$ in the expression for $\boldsymbol{F}_{12}$ and $\boldsymbol{F}_2$. In fact, we show in § V that $r^2|\nabla\delta| \rightarrow$ const. as $|r| \rightarrow \infty$, so only the terms of first order in $\nabla\delta$ contribute (there is no zeroth-order contribution to F_2). We thus have

$$4\pi G\boldsymbol{F}_{12} = -\mu_g \times \lim_{r\rightarrow\infty} \int_{\Sigma} \nabla\delta(\nabla\varphi_g \cdot d\boldsymbol{s}) ,$$

and (24)

$$4\pi G\boldsymbol{F}_2 = \mu_g \times \lim_{r\rightarrow\infty} \int_{\Sigma} (\nabla\delta \cdot \nabla\varphi_g) d\boldsymbol{s} ,$$

where $\mu_g = \mu(\nabla\varphi_g/a_0)$. The two integrals in equation (24) are equal (as both can be easily shown to be equal to $\int_\alpha(\nabla\varphi_g \cdot \nabla)\nabla\delta d^3r$). Hence $\boldsymbol{F}_{12} + \boldsymbol{F}_2 = 0$, and we obtain the desired *exact* result

$$\boldsymbol{F} = m\boldsymbol{A} = -m\nabla\varphi_g . \tag{25}$$

Consider now the more realistic case of a system of mass m and characteristic size r at a distance R from a mass M. Assume that we can take a surface Σ around m of characteristic radius l such that (1) the radius l is small enough compared with R that the field of M alone ($\nabla\varphi_g$) is approximately constant within Σ, and (2) l is large enough that the acceleration $\nabla\delta = \nabla\varphi - \nabla\varphi_g$ on Σ is much smaller, in absolute value, than $|\nabla\varphi_g|$. If m is small enough compared with M, it is possible to find such a surface. Then equation (25) holds approximately. In the limit $m/M \rightarrow 0$, $M/R^2 =$ const., the acceleration of m becomes totally independent of its structure (when tidal effects can be neglected). For a finite but small ratio m/M the relative correction to equation (25) is bound by a number of order $(m/M)^{1/2}$ when $|\nabla\varphi_g| \lesssim a_0$. When $|\nabla\varphi_g| \gg a_0$, the relative correction to equation (25) is bound by a number of order $\mu'(|\nabla\varphi_g|/a_0)(m/M)^{1/2}$ [remember that $\mu'(x) \rightarrow 0$ for $x \rightarrow \infty$].

V. THE EFFECTS OF AN EXTERNAL FIELD ON INTERNAL DYNAMICS

In this section we consider the internal dynamics of a gravitating system, s, in the presence of an external (modified) gravitational field. Observational effects which are relevant to this question have been discussed in Papers I–III. In particular, it was shown, on phenomenological grounds, that the sought for theory for the MOND must violate the strong equivalence principle in that the internal dynamics of a system embedded in an external field is affected by this field. For example, if the external acceleration is large compared with a_0, the internal dynamics of the systems will be approximately Newtonian even if the relative internal accelerations are much smaller than a_0. We will show that the present theory satisfies this requirement.

The external field manifests itself through the boundary condition on the accelerating field at infinity. Solutions are sought for the field equation (3) with $\nabla\varphi \rightarrow \nabla\varphi_g = -\boldsymbol{g}_g$ as $r \rightarrow \infty$ (we use the same notation as in § IV). The external acceleration $\boldsymbol{g}_g$ is to be thought of as the (modified) field of an enveloping system, S, in the absence of s. For example, s can be a binary or a cluster of stars in the field of a galaxy (S) or a galaxy (s) in the field of a cluster of galaxies (S). In general, $\boldsymbol{g}_g$ may vary across s, thus influencing the internal dynamics of s via tidal effects. Since we want to concentrate here on the nontidal effects, we take $\boldsymbol{g}_g$ to be constant.

As we have shown in § IV, the acceleration A of the c.o.m. of s is $\boldsymbol{A} = \boldsymbol{g}_g$. The acceleration of a test particle within s, with respect to the c.o.m. of s, is thus given by

$$\boldsymbol{a} = -\nabla\varphi - (-\nabla\varphi_g) = -\nabla\delta . \tag{26}$$

It is thus the field $\nabla\delta$ (hereafter, the internal acceleration field) which determines the internal dynamics and which concerns us in this section. In Newtonian dynamics $\nabla\delta$ is, of course, identical to the solution of the field equation with the boundary conditions of vanishing gradient at infinity. The internal dynamics is totally unaffected by a constant external acceleration field in this case. This result is valid for any theory which satisfies the strong equivalence principle. This, however, is not the case in the present theory.

We first consider the asymptotic behavior of the internal field of an arbitrary finite mass distribution (of total mass M) in the presence of an external field. Applying Gauss's theorem to the field equation (3) for a surface Σ which surrounds all the mass, we get

$$4\pi GM = \int_{\Sigma} \mu(\nabla\varphi/a_0)\nabla\varphi \cdot d\boldsymbol{s} . \tag{27}$$

We take Σ to be a large enough radius that everywhere on it $|\nabla\delta| \ll |\nabla\varphi_g|$. Expanding in $\nabla\delta$, and taking the lowest order which contributes to the right-hand side of equation (27), we get (the zeroth order vanishes)

$$4\pi GM\mu_g^{-1} = \int_{\Sigma} \nabla\delta \cdot d\boldsymbol{s} + L_g \int_{\Sigma} (\boldsymbol{e}_z \cdot \nabla\delta)(\boldsymbol{e}_z \cdot d\boldsymbol{s}) . \tag{28}$$

Here the constants μ_g and L_g are given by

$$\mu_g = \mu(\nabla\varphi_g/a_0); \quad L_g = d\ln(\mu)/d\ln(x)]_{x=|\nabla\varphi_g|/a_0} , \tag{29}$$

and $\boldsymbol{e}_z$ is a unit vector in the direction of $\boldsymbol{g}_g$ which we chose as the positive Z axis (μ_g is between 0 and 1, and the same is true for L_g if we assume that μ is a monotonic and convex function of its argument [see Paper II]). We use a spherical coordinate system with the origin somewhere within the mass distribution.

Consider the asymptotic behavior of $\nabla\delta$ for $r \rightarrow \infty$. Since the area of Σ increases like r^2 as the integration surface becomes large, we must have $|\nabla\delta| \sim r^{-2}$ for the right-hand side of equation (28) to remain a finite constant. We can thus write:

$$\nabla\delta \xrightarrow{r\rightarrow\infty} r^{-2}\boldsymbol{k}(\theta, \psi) + \boldsymbol{O}(r^{-3}) . \tag{30}$$

We write $\boldsymbol{k}$ in a general form:

$$\boldsymbol{k}(\theta, \psi) = q(\theta, \psi)[\boldsymbol{e}_r + f(\theta, \psi)\boldsymbol{e}_z + p(\theta, \psi)\boldsymbol{e}_\theta] . \tag{31}$$

Substituting equations (30) and (31) in the right-hand side of equation (28), we find that, in order for this to be independent of the surface chosen, $p(\theta, \psi) = 0$ and $f(\theta, \psi) = -L_g \cos\theta/(1 + L_g)$. The requirement that $\nabla\delta$ have a vanishing curl determines q up to a normalization constant which, in turn, is fixed by equation (28). The final result is

$$\nabla\delta \xrightarrow{r\to\infty} MGr^{-2} \times \mu_g^{-1}(1+L_g)^{-1/2}(\boldsymbol{e}_r - \alpha_g \cos\theta \boldsymbol{e}_z) \times (1-\alpha_g \cos^2\theta)^{-3/2} + \boldsymbol{O}(r^{-3}) . \quad (32)$$

Here $\alpha_g = L_g/(1+L_g)$. This behavior is to be contrasted with the asymptotic behavior of the field of an isolated mass: $\nabla\varphi \to (MGa_0)^{1/2} r^{-1}\boldsymbol{e}_r$ as $r\to\infty$. Unlike the last which is a confining field, the first is quasi-Newtonian. Its deviations from a Newtonian behavior are: (*a*) There is an effective increase in G (or a decrease in the ratio of the *inertial* mass to the gravitational mass) by a factor μ_g^{-1} which can be very large if $|g_g| \ll a_0$. (*b*) The asymptotic field is not radial, and its radial part is not spherically symmetric. Both its direction and its magnitude carry information on the direction of the external field. Note in particular that if the external field is Newtonian (i.e., $|\nabla\varphi_g| \gg a_0$), $\mu_g \approx 1$, $L_g \ll 1$ and the asymptotic internal field of an arbitrary mass in this field is very nearly Newtonian even though it is (asymptotically) much smaller than a_0.

In general, it is difficult to describe in simple terms the differences between the internal field of a given system with and without an external field. There are, however, some aspects of this question about which we can make clear-cut statements (such as the asymptotic behavior we have just discussed).

VI. DISCUSSION

We have presented a nonrelativistic potential theory for gravity which deviates from Newtonian theory in the limit of small accelerations. The theory was built on the basic rules of MOND, and is in some sense the simplest such theory of gravity which is derivable from a Lagrangian.

It is worth noting that if we relax the requirement of a single potential in the theory, a more general family of such theories can be constructed. For example, we could write a theory which involves two potentials φ_1 and φ_2 determined by the field equations

$$\nabla^2\varphi_1 = 4\pi G(1-\lambda)\rho(\boldsymbol{r}) ,$$
$$\nabla\cdot[\bar{\mu}(\nabla\varphi_2 \times \lambda/a_0)\nabla\varphi_2] = 4\pi G\lambda\rho(\boldsymbol{r}) , \quad (33)$$

where $0 \le \lambda \le 1$ is an arbitrary parameter and

$$\bar{\mu}(x) \overset{x\gg 1}{\approx} 1 , \quad \bar{\mu}(x) \overset{x\ll 1}{\approx} x .$$

Again we supplement the field equations with the boundary condition $|\nabla\varphi_i| \to 0$ as $r\to\infty$. The way λ appears in the field equations is chosen in accordance with assumption (*c*) of MOND (§ I). The acceleration $\boldsymbol{a}$ of a test particle is here regarded as given by $\boldsymbol{a} = -\nabla(\varphi_1+\varphi_2)$. It is easy to show that the conservation laws hold in this theory, and it is also straightforward to generalize the results of the previous sections for this theory and show, in particular, that the c.o.m. acceleration of an arbitrary system in a constant external field is independent of internal properties of the system.

The two-potentials theory is derivable from the Lagrangian density

$$\mathscr{L} = -\rho(\varphi_1+\varphi_2) - (8\pi G)^{-1}\{(1-\lambda)^{-1}(\nabla\varphi_1)^2 + \lambda^{-3}a_0^2\overline{\mathscr{F}}[(\nabla\varphi_2)^2\lambda^2/a_0^2]\} . \quad (34)$$

The relativistically invariant theory which we sketch briefly in Appendix B reduces to a theory of this type in the nonrelativistic limit in a static universe.

Using the present (single potential) theory instead of the semiempirical approach of Papers I–IV makes practically no difference for the applications described in these papers. The latter are based essentially only on the basic requirements of MOND which are satisfied by the present theory. Some of the main results of this theory which cannot be obtained from the formalism used in Papers I–IV are: (*a*) the demonstration of the consistency of the usual conservation laws with the assumptions of MOND; (*b*) the justification for the use of MOND to describe the c.o.m. motion of composite objects in an external field; and (*c*) its consistency with the conclusion (based on the study of open clusters) that a full theory of MOND must violate the strong equivalence principle in that the internal dynamics of a system is affected by a constant external acceleration field.

APPENDIX A

PROOF THAT $\boldsymbol{u} \equiv \nabla\varphi_N - \mu(|\nabla\varphi|/a_0)\nabla\varphi$ VANISHES AT LARGE DISTANCES FROM A MASS AT LEAST AS $\boldsymbol{O}(r^{-3})$

Consider a bound density distribution of total mass M with the origin at the center of mass. The vector field $\boldsymbol{u}$ defined above satisfies $\nabla\cdot\boldsymbol{u} = 0$ and vanishes at infinity. We can thus write $\boldsymbol{u}$ in terms of the vector potential $\boldsymbol{A}$:

$$\boldsymbol{u} = \nabla\times\boldsymbol{A} ; \qquad \boldsymbol{A}(\boldsymbol{r}) = (4\pi)^{-1}\int \frac{\nabla'\times\boldsymbol{u}(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}d^3r' . \quad (\mathrm{A1})$$

Thus the only term with an r^{-2} behavior at infinity which $\boldsymbol{u}$ can have is $\boldsymbol{u}^{(2)} = \nabla\times(r^{-1}\boldsymbol{B}) = r^{-3}\boldsymbol{r}\times\boldsymbol{B}$, where $\boldsymbol{B} = (4\pi)^{-1}\int\nabla'\times\boldsymbol{u}'d^3r'$, which is the lowest order term in the multipole expansion of equation (A1). If we then show that $\boldsymbol{B} = 0$, we get that the lowest contributing multiple term to $\boldsymbol{u}$ vanishes at least as fast as r^{-3}.

To this end we make use of the fact that $\boldsymbol{u}$ is not an arbitrary divergenceless field but has the special form given above. In the limit of large r,

$$\mu(|\nabla\varphi|/a_0)\nabla\varphi = \nabla\varphi_N - \boldsymbol{u} = r^{-3}(MG\boldsymbol{r} - \boldsymbol{r}\times\boldsymbol{B}) + \boldsymbol{O}(r^{-3}) \quad (\mathrm{A2})$$

Taking the absolute value of this relation, we can express $|\nabla\varphi|$ as a function of the absolute value of the right-hand side. Then dividing by $\mu(|\nabla\varphi|/a_0)$ (which for $r \to \infty$ can be replaced by $|\nabla\varphi|/a_0$), we get

$$\nabla\varphi = a_0^{1/2} r^{-2}(MG\boldsymbol{r} - \boldsymbol{r} \times \boldsymbol{B})/(M^2G^2 + B^2 \sin^2 \theta)^{1/4} + O(r^{-2}) . \tag{A3}$$

Here θ is the angle between $\boldsymbol{r}$ and $\boldsymbol{B}$ which we can take along the z-axis. Now, requiring that the azimuthal component of $\nabla \times (\nabla\varphi)$ vanishes gives $\boldsymbol{B} = 0$.

APPENDIX B

A RELATIVISTIC THEORY WHICH SATISFIES THE REQUIREMENTS OF MOND

The main ingredients which must go into a relativistic theory for the MOND are (*a*) recovery of the non-Newtonian behavior in the nonrelativistic limit and (*b*) a violation of the strong equivalence principle as deduced from the phenomenological study of open clusters in the solar neighborhood (Paper I). To introduce the second ingredient into the theory, we include a scalar field ψ as a dynamical degree of freedom in addition to the usual metric tensor $g_{\mu\nu}$, in the spirit of scalar-tensor theories. The theory we sketch here (we intend to publish a more detailed account separately) can be given two equivalent formulations involving a metric tensor and one scalar field. In one, $g_{\mu\nu}$ satisfies the usual Einstein equation but test particles do not follow geodesics of $g_{\mu\nu}$. In the other formulation, particles do follow geodesics of a metric tensor $\tilde{g}_{\mu\nu}$ which, however, does not satisfy the usual Einstein equation. Instead, the field equation for $\tilde{g}_{\mu\nu}$ contains the scalar field explicitly. This dual description corresponds to the use of two different sets of units (gravitational units in the first description and atomic units in the second) and has been discussed extensively in the literature in connection with theories with variable rest masses (e.g., Dicke 1962; Hoyle and Nalikar 1974; Bekenstein and Meisels 1980). The field equations in the atomic units for example, are derivable from the action

$$S = S_g + S_\psi + S_m , \tag{B1}$$

where the gravitational action is

$$S_g = c^4(16\pi G_0)^{-1} \int e^{-2\psi/c^2}[R(\tilde{g}_{\mu\nu}) + 6c^{-4}\psi_{,\alpha}\psi_{,}{}^{\alpha}](-\tilde{g})^{1/2} d^4x , \tag{B2}$$

the " scalar field " action is

$$S_\psi = -a_0{}^2\beta(1 + \beta)^2(8\pi G_0)^{-1} \int e^{-4\psi/c^2} F\left[\frac{e^{2\psi/c^2}\psi_{,\mu}\psi_{,}{}^{\mu}}{a_0{}^2(1 + \beta)^2}\right](-\tilde{g})^{1/2} d^4x , \tag{B3}$$

with $F(\chi) \approx \chi$ for $\chi \gg 1$ and $F(\chi) \approx \frac{2}{3}\chi^{3/2}$ for $\chi \ll 1$, and the particle action is $S_m = -mc^2 \int d\tau$ with $d\tau$ the element of proper time. The form of S_m implies that particles move on geodesics of $\tilde{g}_{\mu\nu}$.

In the strong field limit ($\chi \gg 1$) this theory goes exactly to the Brans-Dicke scalar-tensor theory with $\beta = 2\omega + 3$, where ω is the parameter of the Brans-Dicke (BD) theory (Brans and Dicke 1961; Dicke 1962; Misner, Thorner and Wheeler 1973). All observational constraints on the BD theory to date (e.g., Will 1979, 1982) are in the form of lower limits on ω. Such constraints were obtained in field regimes corresponding to $\chi \gg 1$. As in such cases our theory gives the same predictions as the BD theory, the lower limits on ω give lower limits on our parameter β.

For small velocities and weak potentials (not to be confused with the limit of small accelerations) and assuming a static universe (so that the cosmological contribution to $\partial\psi/\partial t$ can be assumed to vanish) the theory involves two potentials and is of the type described in § VI. The nonrelativistic limit of the theory thus satisfies the requirements of MOND.

We now consider gravitational waves (GW). All systems which are thought to emit gravitational waves, in quantities which may be detectable, involve very large accelerations and so the emission for such systems can be calculated to high accuracy using the BD theory. We have not attempted the calculation of GW emission by low acceleration systems where our theory may differ from that of BD.

There are some aspects of GW *propagation* distinctive to our theory which are of great importance in principle. In particular, under certain circumstances, we find *acausal* propagation of a certain component of the wave. Consider the propagation of weak GW in the background of a weak static acceleration field (e.g., in intergalactic space). The wave is assumed to be a small disturbance on the background, and we linearize the field equations in the wave amplitude. We also assume that the wave length is short compared with the scale over which the background field varies.

The measurable metric (atomic units) can in this approximation be written as a sum of a conformally flat part and a transverse part $h_{\mu\nu}$

$$\tilde{g}_{\mu\nu} \approx (1 - 2\psi/c^2)\eta_{\mu\nu} + h_{\mu\nu} , \tag{B4}$$

where $\eta_{\mu\nu}$ is the Minkowski metric. The transverse contribution is found to propagate with the speed of light. The propagation of ψ is anisotropic. If θ is the angle between the wave vector and the direction of the background field, the ψ wave propagates with velocity

$$v(\theta) = c(1 + 2v_b \cos^2 \theta)^{1/2} . \tag{B5}$$

Here $v_b = d \ln F'(\chi)/d \ln (\chi)$ at the value of χ corresponding to the background field.

It is not clear whether the acausal propagation rules out the relativistic theory. First, we have not yet thoroughly investigated whether the acausal component can be emitted. Second, it appears that the acausal waves cannot induce acausal effects in the behavior of particles or electromagnetic fields. Of the two types of waves, only one propagates acausally. This component is, however, only a "conformal factor" multiplying a causally propagating metric. The fact that the light cone is the same for conformally related metrics means that the acausal propagation of the conformal factor in $\tilde{g}_{\mu\nu}$ does not affect the light cone structure of the measurable metric. Thus, particles will not start crossing the light cone due to a metric distortion propagating faster than c and shoving the light cone past them. The waves cannot accelerate particles to transluminal velocities. For a similar reason a transluminal pulse of ψ passing by does not affect electromagnetic phenomena.

It is possible that the present theory may be modified so as to avoid the acausal behavior while retaining the desired nonrelativistic behavior and relativistic invariance. This is a subject for the future.

We have not yet studied cosmology in the present theory. The study of cosmology has a twofold importance. First, it may provide us with additional constraints on the theory, and expose it to further tests. Second, it may provide the key to an improved theory in which the function F and the constants a_0 and β will be determined from first principles.

REFERENCES

Alder, S. L., and Piran, T. 1984, *Rev. Mod. Phys.*, **56**, 1.
Bekenstein, J., and Meisels, A. 1980, *Phys. Rev.*, **D22**, 1313.
Brans, C., and Dicke, R. 1961, *Phys. Rev.*, **124**, 925.
Dicke, R. 1962, *Phys. Rev.*, **125**, 2163.
Felten, J. E. 1984, *Ap. J.*, **286**, 3.
Finzi, A. 1963, *M.N.R.A.S.*, **127**, 21.
Hoyle, F., and Narlikar, J. 1974, *Action at a Distance in Physics and Cosmology* (San Francisco: Freeman).
Lehman, H., and Wu, T. T. 1983, preprint.
Milgrom, M. 1983*a*, *Ap. J.*, **270**, 365 (Paper I).
———. 1983*b*, *Ap. J.*, **270**, 371 (Paper II).
Milgrom, M. 1983*c*, *Ap. J.*, **270**, 384 (Paper III).
———. 1984, *Ap. J.*, in press (Paper IV).
Misner, C., Thorne, K., and Wheeler, J. 1973, *Gravitation* (San Francisco: Freeman).
Tohline, J. E. 1983, in *IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Dordrecht: Reidel).
Weinberg, S. 1972, *Gravitation and Cosmology* (New York: Wiley).
Will, C. M. 1979, in *General Relativity—An Einstein Centenary Survey*, ed. S. W. Hawking and W. Israel (Cambridge: Cambridge University Press).
———. 1982, *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press).

JACOB BEKENSTEIN: Department of Physics, Ben-Gurion University, Beer-Sheva 84105, Israel

MORDEHAI MILGROM: Department of Physics, Weizmann Institute, Rehovot 76100, Israel