THE ASTROPHYSICAL JOURNAL, 306:9–15, 1986 July 1


# CAN THE HIDDEN MASS BE NEGATIVE?

MORDEHAI MILGROM
Department of Physics, Weizmann Institute of Science, Rehovot, Israel
*Received 1985 April 29; acceoted 1985 December 18*

## ABSTRACT

If the mass discrepancy in galactic systems is due to a breakdown of Newtonian dynamics (rather than to hidden mass), it may be possible to find configurations in which the required "hidden mass" is negative. A clear-cut instance of this kind will demonstrate the inadequacy of Newton's laws in the realm of the galaxies and rule out the hidden mass hypothesis.

We show that if the modified nonrelativistic dynamics is basically correct, such configurations do exist and, in fact, are commonplace. It is however very difficult, at present, to detect such a negative "phantom mass" because of the limitations of astronomical measurements.

We also derive the exact expression for the Oort discrepancy in very thin disks.

*Subject headings:* cosmology — gravitation

*In Astronomy we see the bodies whose motion we are studying, and in most cases we grant that they are not subject to the action of other invisible bodies.*

H. Poincaré, in *Science and Hypothesis*

### I. INTRODUCTION

It has been suggested that hidden mass does not exist in appreciable quantities in galaxies and galaxy systems (Milgrom 1983*a*, hereafter Paper I). The large discrepancy between the dynamically determined masses and those which are observed directly is blamed on inadequacy of the Newtonian laws of dynamics. Modified nonrelativistic dynamics (MOND) has been put forth, and its consequences for galaxies and galaxy systems have been discussed in Paper I and in Milgrom (1983*b*, *c*; 1984). A nonrelativistic Lagrangian theory for MOND is described in Bekenstein and Milgrom (1984, hereafter Paper II).

Most of the many predictions and consequences of MOND which were discussed in Milgrom (1983*a*, *b*, *c*, 1984) can also be reproduced in the framework of the hidden mass hypothesis (HMH). This is the case provided one is willing to assume that hidden mass is present in whatever quantities and space distributions are required, ad hoc for each system. An observable contradiction to these predictions will rule out MOND, while HMH can deny its paternity of these offspring. It is important to find consequences of MOND which cannot be mimicked with hidden matter. A verification of such a prediction will contradict the HMH. We discussed briefly two such consequences in Paper I: (*a*) the violation of the strong equivalence principle required by MOND; and (*b*) the possibility that the "mass" of an object, as determined dynamically (assuming conventional dynamics) from its effects on massive particles, will give the wrong amount of light-bending produced by this object (such as is observed, for example, in gravitational lenses). The first effect is expected to occur in full strength in the nonrelativistic regime, and its nature in MOND is rather well understood in light of the Lagrangian theory of Paper II (see discussion in Milgrom 1986). Knowledge of the exact nature and degree of the second effect will have to await the development of a consistent relativistic theory for MOND.

In the present paper we discuss a third consequence of MOND which is incompatible with HMH. If MOND is basically correct, there are mass configurations which, when probed dynamically using Newtonian dynamics will appear to contain negative "hidden mass." In other words, the deduced mass density is such systems is, in some regions of space, smaller than the actual density (which can be observed directly).

The purpose of the present paper is to discuss generalities concerning the possible appearance of negative hidden mass. We do not present detailed predictions for any particular system. There are difficulties in detecting such an effect due to the limitations of astronomical observations. As a result I cannot offer, at the moment, a specific method and system where the effect can be looked for.

Section II describes the general procedure for determining gravitational masses and the conditions for the generation of negative phantom masses. Section III presents actual configurations where negative phantom mass is expected by MOND, and in § IV we discuss prospects for detecting such "negative" densities. Section V summarizes the conclusions.

### II. MASS DETERMINATION AND THE NOTION OF PHANTOM MASS

Various techniques have been employed for measuring masses and mass densities in astronomical systems via their gravitational effects on the trajectories of massive test particles (see, e.g., the review of Faber and Gallagher 1979) (mass determination using light-bending will not be discussed here). Although it is not always obvious, these methods are special cases of the following procedure. Determine the acceleration field $\boldsymbol{g}(\boldsymbol{r})$ for massive test particles. Assume Newtonian dynamics, i.e., the validity of the second law $[\boldsymbol{g}(\boldsymbol{r}) = -\nabla\varphi]$ and the law of gravity (Poisson's equation $\nabla^2\varphi = 4\pi G\rho^*$); then determine the gravitational mass density $\rho^*(\boldsymbol{r})$ from

$$\rho^*(\boldsymbol{r}) = -(4\pi G)^{-1}\nabla \cdot \boldsymbol{g}(\boldsymbol{r}) . \tag{1}$$

If Newtonian dynamics is not applicable, the dynamically determined density $\rho^*(\boldsymbol{r})$ as given by equation (1) is not the true density $\rho(\boldsymbol{r})$ which give rise to the measured $\boldsymbol{g}(\boldsymbol{r})$. We term the excess $\rho_p(\boldsymbol{r}) \equiv \rho^*(\boldsymbol{r}) - \rho(\boldsymbol{r})$, the phantom mass density (PMD). HMH considers $\rho_p$ to be a real but yet unobserved mass density and thus implies $\rho_p(\boldsymbol{r}) \geq 0$ everywhere.

In both the archetype formulation of Paper I and the Lagrangian formulation of Paper II, the true density $\rho(\boldsymbol{r})$ is related to the measured test particle acceleration field $\boldsymbol{g}(\boldsymbol{r})$ by

$$\rho = -(4\pi G)^{-1}\boldsymbol{\nabla} \cdot [\mu(g/a_0)\boldsymbol{g}] , \tag{2}$$

where $\mu(x)$ is described in Papers I and II.

Equation (2) does not fix $\boldsymbol{g}(\boldsymbol{r})$ uniquely for a given $\rho(\boldsymbol{r})$. If in addition to equation (2) we required $\boldsymbol{g} = -\boldsymbol{\nabla}\varphi$, $\boldsymbol{g}$ is determined uniquely (Milgrom 1986), and equation (2) is then the field equation for the Lagrangian theory (Paper II). If, on the other hand, we take $\mu(g/a_0)\boldsymbol{g} = \boldsymbol{g}_{\rm N}$, we get the MOND formulation of Paper I, where $\boldsymbol{g}_{\rm N}$ is the Newtonian acceleration produced by $\rho$.

Now write equation (2) as

$$\begin{aligned}\rho &= -(4\pi G)^{-1}\{\boldsymbol{\nabla} \cdot \boldsymbol{g}\mu(g/a_0) + \boldsymbol{g} \cdot \boldsymbol{\nabla}[\mu(g/a_0)]\} \\ &= \rho^*\mu(g/a_0) - (4\pi G)^{-1}\mu'(g/a_0)a_0^{-1}\boldsymbol{g} \cdot \boldsymbol{\nabla} g .\end{aligned} \tag{3}$$

The phantom mass density can then be written as

$$\rho_p(\boldsymbol{r}) = \rho(\boldsymbol{r})(1/\mu - 1) + (4\pi G)^{-1}L\boldsymbol{e}_g \cdot \boldsymbol{\nabla} g . \tag{4}$$

Here $L = d \ln [\mu(x)]/d \ln (x)$ at $x = g/a_0$, and $\boldsymbol{e}_g$ is a unit vector in the direction of $\boldsymbol{g}$. The quantity $L$ varies between 0 (in the limit of very high accelerations) and 1 (in the opposite limit), as we assume that $\mu$ is increasing and convex. Of the two terms in equation (4), the first is never negative ($\mu \leq 1$). The second is negative if $|g|$ decreases in the direction of $\boldsymbol{g}$.

The total phantom mass of a finite system as measured in a large sphere surrounding the system is of course positive and in fact diverges with increasing radius of the sphere. For such a system, at large distances, $g \propto r^{-1}$ (and $\rho = 0$ at large $r$); thus from equation (4), $\rho_p \propto r^{-2}$ as expected (since $\rho_p$ is to produce a flat asymptotic rotation curve).

Negative phantom mass is perhaps best looked for in regions of space where the actual density is negligible. Setting $\rho = 0$ in equation (4) yields a criterion by which $\rho_p < 0$ is equivalent to $\boldsymbol{e}_g \cdot \boldsymbol{\nabla} g < 0$. To find the regions of $\rho_p < 0$, we have to solve the field equation of Paper II (by numerical means in practically all cases) and substitute the resulting $\boldsymbol{g}(\boldsymbol{r})$ in equation (4). I found it very useful, in looking for negative PMD configurations to employ the phenomenological relation used in Paper I. Although this is not the theory which we now use for MOND, it is straightforward to solve, and it gives the correct qualitative results in all the problems I have checked. (In many cases the results are very good quantitatively.) I use this formulation only as a useful and quick guide. Equation (4) is still valid in this formulation, but $\boldsymbol{g}$ is calculated differently. In this formulation $\boldsymbol{g}$ is parallel to the Newtonian acceleration $\boldsymbol{g}_{\rm N}$, and $\boldsymbol{\nabla} g_{\rm N} = \mu(1 + L)\boldsymbol{\nabla} g$, so we get from equation (4)

$$\hat{\rho}_p(\boldsymbol{r}) = \rho(\boldsymbol{r})(1/\mu - 1) + [L/4\pi G\mu(1 + L)]\boldsymbol{e}_{\rm N} \cdot \boldsymbol{\nabla} g_{\rm N} , \tag{5}$$

where the caret on $\hat{\rho}_p$ marks the fact that it is gotten from the approximate theory and $\boldsymbol{e}_{\rm N}$ is a unit vector in the direction of $\boldsymbol{g}_{\rm N}$.

The field $\boldsymbol{g}_{\rm N}$ is of course easily calculable for any mass configuration, and the regions with $\rho_p < 0$ can be mapped immediately. In general, these turn out to have the same structure as the $\rho_p < 0$ regions for the exact theory (as we have found numerically; Milgrom 1986).

The field equation can be solved in closed form and an analytic expression for $\rho_p$ obtained for a class of important configurations, viz., mass distributions in a constant external field $\boldsymbol{g}_0$ such that the total acceleration deviates only a little from $\boldsymbol{g}_0$ in most of space. Examples of astronomical systems which fall under this description are open clusters and very wide binary stars in the field of a galaxy and a satellite dwarf elliptical in a field of a parent galaxy. Let $\boldsymbol{g} = \boldsymbol{g}_0 + \boldsymbol{\gamma} = \boldsymbol{g}_0 - \boldsymbol{\nabla}\eta$ be the solution of the field equation subject to the boundary condition $\boldsymbol{g} \to \boldsymbol{g}_0 (r \to \infty)$ and assume $\gamma \ll g_0$ everywhere. We can then expand the field equation in $\boldsymbol{\nabla}\eta$, retaining only the lowest contributing terms, to obtain

$$\nabla^2\eta + L_0\,\partial^2\eta/\partial z^2 \approx (4\pi G/\mu_0)\rho . \tag{6}$$

Here $\mu_0$ and $L_0$ are the values of $\mu$ and $L$ at $g_0$, and the $z$-axis is taken in the direction of $\boldsymbol{g}_0$. The PMD is thus given by

$$\rho_p = (4\pi G)^{-1}\nabla^2\eta - \rho = (1/\mu_0 - 1)\rho - (4\pi G)^{-1}L_0\,\partial^2\eta/\partial z^2 . \tag{7}$$

Because equation (6) is linear, the PMD is the sum of the PMDs produced by the individual elements of $\rho(\boldsymbol{r})$. In the coordinates $\hat{x} \equiv x$, $\hat{y} \equiv y$, $\hat{z} \equiv z(1 + L_0)^{-1/2}$, equation (6) is identical to the Poisson equation with $\rho$ replaced by $\rho/\mu_0$, and $\eta$ can thus be written in closed form.

$$\begin{aligned}\eta(\boldsymbol{r}) &= -\left(\frac{G}{\mu_0}\right)\int \frac{\rho(\hat{\boldsymbol{r}}')d^3\hat{r}'}{|\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}'|} \\ &= -\left[\frac{G}{\mu_0(1 + L_0)^{1/2}}\right]\int \frac{d^3r'\rho(\boldsymbol{r}')}{R(|\boldsymbol{r} - \boldsymbol{r}'|)} ,\end{aligned} \tag{8}$$

where $R(|\boldsymbol{r} - \boldsymbol{r}'|) = [(x - x')^2 + (y - y')^2 + (z - z')^2/(1 + L_0)]^{1/2}$. Taking the second $z$-derivative of $\eta$ and substituting in equation (7), we get the phantom density

$$\rho_p(\boldsymbol{r}) = (1/\mu_0 - 1)\rho(\boldsymbol{r}) - [L_0/4\pi\mu_0(1 + L_0)^{5/2}]\int d^3r'\rho(\boldsymbol{r}')R^{-5}(|\boldsymbol{r} - \boldsymbol{r}'|)\{(1 + L_0)[(x - x')^2 + (y - y')^2] - 2(z - z')^2\} . \tag{9}$$

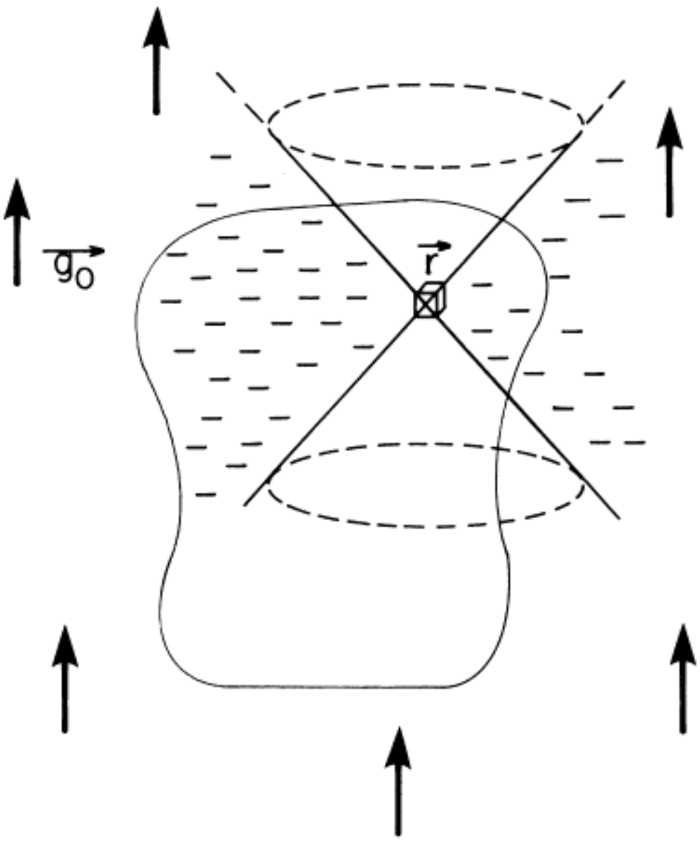


FIG. 1.—The region where a mass element at $\boldsymbol{r}$ contributes negatively to the phantom mass density in a low surface density configuration in an external field (region marked with minus signs).

Thus each element of mass $dM(\boldsymbol{r}') = \rho(\boldsymbol{r}')d^3r'$ gives rise to a negative contribution to $\rho_p$ everywhere outside a cone with its vertex at $\boldsymbol{r}'$, axis along $z$, and half-opening angle $\theta$ with $tg\theta = [2/(1 + L_0)]^{1/2}$ (where the integrand in eq. [9] is positive, see Fig. 1).

Coming back to the general case, the following theorem helps us find configurations with negative PMD. A system which contains an isolated point $O$ of zero acceleration in vacuum (such as the midpoint between two equal masses or the center of a ring) must contain a region of negative PMD, the boundary of which pases thorugh the point $O$.

To prove this statement, choose a small, empty, closed, simply connected surface $\Sigma$ of constant $g \equiv |\boldsymbol{g}| \neq 0$ around $O$ (that this is possible follows from our assumptions on $O$). We now show that $\boldsymbol{g} \cdot \nabla g$ must be negative in some regions of $\Sigma$ and positive in others. Integrating equation (2) (with $\rho = 0$) in the volume contained in $\Sigma$, we have $\int_\Sigma \mu(g/a_0)\boldsymbol{g} \cdot d\boldsymbol{\sigma} = 0$ or, since $g$ is constant on the surface, $\int_\Sigma \boldsymbol{g} \cdot d\boldsymbol{\sigma} = 0$. Again, because $g$ is constant on $\Sigma$, $d\boldsymbol{\sigma}$ is parallel to $\nabla g$ (we take $\Sigma$ small enough so that $g$ which vanishes at $O$ increases outward). Thus $\boldsymbol{g} \cdot \nabla g$ cannot have the same sign everywhere on $\Sigma$. The quantity $\boldsymbol{g} \cdot d\boldsymbol{\sigma}$ cannot vanish everywhere on $\Sigma$, because this would mean that $\boldsymbol{g}$ is everywhere parallel to $\Sigma$, leading to field lines which do not start and end on masses. This is forbidden in both formulations of MOND.

We thus show that on every small equi-$g$ surface around $O$ there are regions of positive and negative PMD. Thus $O$ must be on the boundary between positive and negative PMD regions.

Arguing by continuity, if we now add some small density at $O$ we are still left with a region of negative PMD, although, in general, it will now be somewhat different with its boundary detached from $O$.

For a spherically symmetric system we can get $\rho_p(\boldsymbol{r})$ in closed form without using the above formulae. In this case, the formulations of Papers I and II give the same field $\boldsymbol{g}(\boldsymbol{r})$, which is related to the Newtonian acceleration by $\mu(g/a_0)\boldsymbol{g} = \boldsymbol{g}_N$, with $\boldsymbol{g}_N = -M(r)Gr^{-2}\boldsymbol{e}_r$, and by definition $g = M^*(r)Gr^{-2}$. Substituting in the above relation between $g$ and $g_N$ and taking the derivative with respect to $r$, we get, after some algebra,

$$\rho_p(r) = \rho(r)(1/I' - 1) + L(1 + L)^{-1}M(r)(4\pi r^3\mu)^{-1} . \tag{10}$$

Here $I(x) = x\mu(x)$, and the argument of $I'$ is $g/a_0$. We thus see that if $\mu$ satisfies $I'(x) \leq 1$ [such as for $\mu(x) = x/(1 + x)$], $\rho_p$ can never be negative in a spherical system. If, on the other hand, $I'(x) > 1$ in some range of $x$ [such as for $\mu(x) = 1 - e^{-x}$], there can be found spherical distributions $\rho(r)$ for which $\rho_p < 0$ in some region.

What is the sign of the phantom density if one assumes a modification of the distance dependence of gravity (as in Finzi 1963, Tohline 1983, or Sanders 1984)? Such a modification I believe, cannot account for the "hidden mass" as observed in galaxies and galaxy systems but is of some interest in the present context. In such schemes, which are linear, the field produced by many masses is the sum of the fields produced by each mass alone and thus, from equation (1), so is the case for the phantom mass. Thus negative phantom densities can be produced in such formulations if, and only if, they are produced by a point mass. Write generally the modified acceleration of a point mass $M$ as

$$g(r) = -MGr^{-2}e_r[1 + q(r)] , \tag{11}$$

with $q(0) = 0$ [a nonzero $q(0)$ can be absorbed in $G$]. From equation (1) we thus have

$$\rho^*(r) = M\delta^3(r) + (4\pi r^2)^{-1}Mq'(r) , \tag{12}$$

or

$$\rho_p(r) = (4\pi r^2)^{-1}Mq'(r) . \tag{13}$$

Since one usually chooses $q(r)$ to be increasing with $r$ (to make the modification more important on galactic scales than in the laboratory), as is the case in all the references mentioned above, we get $\rho_p \geq 0$ for an arbitrary system.

## III. CONFIGURATIONS WITH NEGATIVE PHANTOM DENSITIES

Of the many astronomical systems which have regions with negative PMD, only a few may eventually lend themselves to its measurement. Thus the best strategy, at the moment, is to go through a large number of cases with negative PDM in the hope of finding one which can also be tested.

The purpose of the present section is to provide an initial library of prototype isolated configurations with negative PMD. A list of a few such schematic configurations follows with a short explanation. Although, in principle, we may find $\rho_p < 0$ in systems with arbitrarily large accelerations, $|\rho_p|$ is significant only in systems with small accelerations ($MG/r^2 \ll a_0$), as is evident, for example, from equation (4) ($L$ becomes small for large accelerations).

### *a) Two Point Masses $m_2 \geq m_1$ (or masses of a size small enough compared with their separation)*

The point $O$ on the connecting line between the masses at which $\nabla\varphi = 0$ is on the boundary of a region with $\rho_p < 0$ (from the general theorem of § II). For $m_1 = m_2$, this region has a figure-**8** shape rotated about the connecting line ($z$-axis) which is perpendicular to the major axis of the **8** (see Fig. 2*a*) (for $m_1 \neq m_2$, the shape is a distorted **8**). Moving from $O$ radially in the $x$–$y$ plane $|\nabla\varphi|$ increases, reaches a maximum at $r_m$, and then decreases (like $r^{-1}$ at large $r$). According to the above criteria, $\rho_p < 0$ between $O$ and $r_m$, and $\rho_p > 0$ beyond $r_m$. Along the $z$-axis, $\nabla\varphi$ and $\nabla|\nabla\varphi|$ are in opposite directions, and hence $\rho_p > 0$. Hence the figure-**8** shape of the $\rho_p < 0$ region. This description has been verified by a numerical solution of the problem.

Consider a two point mass system with $q = m_1/m_2$ and a large separation $l$ ($MG/l^2 \ll a_0$, $M \equiv m_1 + m_2$). From the scaling laws discussed in Milgrom (1986), we have in this limit (outside the masses) $\rho_p(r) = h(q, r/l)\rho_s$, where $h(q, \lambda)$ is a dimensionless function independent of $M$ or the separation, and $\rho_s = \rho_0(r_t/l)^2$, where $\rho_0 = M/r_t^3$ and $r_t = (MG/a_0)^{1/2}$. Thus $\rho_s = M^{1/2}a_0^{1/2}G^{-1/2}l^{-2}$. Using, for example, equation (5), based on the formulation of Paper I we get a closed expression for $\hat{\rho}_p$ in the symmetry plane of the equal mass case

$$\hat{h}(1, \lambda) = -(2^{3/2}\pi)^{-1}\lambda^{-1/2}(1 - 8\lambda^2)(1 + 4\lambda^2)^{-7/2} , \tag{14}$$

where $\lambda = r/l$. We note that $\hat{\rho}_p$ diverges like $r^{-1/2}$ at $r = 0$ and remains negative up to $r = l/2^{3/2}$. Note that $r_t$ is the transition radius from the Newtonian to the modified regime, and we assumed $l \gg r_t$.

Any sufficiently isolated pair of galaxies, such as the Milky Way and Andromeda, is an instance of such a system. In fact, any small group of galaxies contains a point (usually outside the galaxies themselves) where $\nabla\phi = 0$ and thus must also have a $\rho_p < 0$ region by the general theorem.

### *b) A Disk with a Surface Density Depression around Its Center*

Consider first a ring or an arbitrary thin planar axisymmetric disk with a hole around its center; the center then is an isolated point where $g = 0$ and $\rho = 0$ and must thus be on a boundary of a $\rho_p < 0$ region by the theorem proved in § II. Arguments similar to those given for two point masses show that this region has an **8** shape rotated about its major axis, which is perpendicular to the disk (see Fig. 2*b*).

More realistically, take any axisymmetric planar disk with a central depression in the surface density. Let the disk's typical size (say the half-mass radius) be $r_s$, its average surface density $\sigma_{av}$, and its central surface density $\sigma_c$. Along the $z$ symmetry axis the maximum value of $u = \mu(g/a_0)g$ is of order $\pi\sigma_{av}G$ and is reached at a distance of order $r_s$. The exact value of $u$ at the surface of the disk at its center is $2\pi\sigma_c G$ (from eq. [2] the change in the $u_n$, the perpendicular component of $u = \mu g$, across the surface is $4\pi\sigma_c G$, and the above result is then obtained from the symmetry of the problem).Thus if $\sigma_c$ is small enough compared with $\sigma_{av}$, there must be a segment along the $z$-axis where $g$ increases (against the direction of $g$) as we go away from the disk. Hence on this segment (and from continuity in some region containing it), $\rho_p < 0$.

Photometry of disk galaxies has shown some with such a central depression in disk surface density (Freeman 1970; Kormendy 1977). The bulge, which they also have, may sometimes (but not always) cause the $\rho_p < 0$ region to disappear.

### *c) The Phantom Density inside and Just outside a Thin Galactic Disk*

Studies of the dynamical density distribution with height above the Galactic plane in the solar neighborhood which go sufficiently high above the plane have indicated negative densities above a few hundred parsecs (Hill 1960; King 1977) [some analyses impose the constraint $\rho > \rho$(observed), and they, of course, do not encounter $\rho < 0$]. A result like this, being nonsensical in Newtonian mechanics, has been dismissed as an artifact of incomplete or inaccurate data. A question is raised, in the context of the present discussion, as to what MOND predicts for this case. In general, $\rho_p$ at any point in the field of a galaxy depends on the overall mass distribution in the galaxy. As it turns out, however, we can express $\rho_p$ at a point in or just outside a very thin galactic disk in terms of quantities which are directly measurable *near that point* (rotation velocity and surface density).

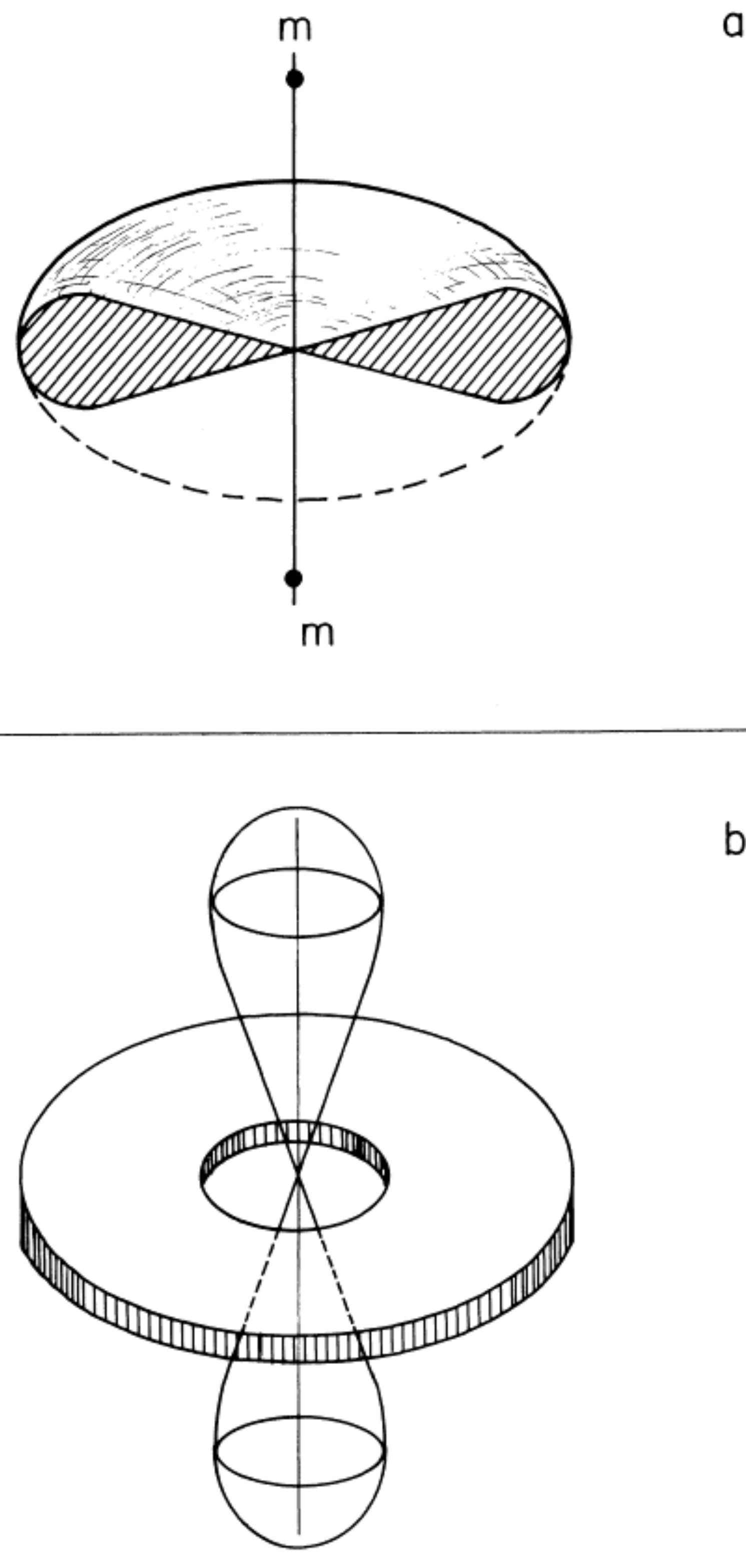


FIG. 2.—Schematic depiction of the negative phantom density regions in (*a*) a two equal mass configuration and (*b*) a flat axisymmetric disk with a central hole

The galaxy model which we use consists of a very thin disk and a bulge. We assume that the galaxy is axisymmetric and mirror symmetric about the $z = 0$ galactic plane. Consider a region in the disk around galactocentric distance $r$ where the density of the bulge can be neglected (more precisely, where the density in the bulge is much smaller than the average density of the disk within a sphere of radius $r$). Let $\Sigma(r, z)$ be the disk surface density between $-z$ and $+z$, $V(r)$ the rotation velocity in the galactic plane, and $g_r$, $g_z$ respectively the $r$ and $z$ components of the modified acceleration field. Applying the Gauss theorem to the field equation in a cylindrical volume with its bases rings of radius $r$ and width $dr$ at $\pm z$ such that $r \gg dr$, we obtain, in the $z > 0$ side of the disk (using the reflection symmetry),

$$\mu(g/a_0)g_z = -2\pi G\Sigma(r, z) . \tag{15}$$

*This equation is a good approximation as long as* $z \ll rg_z[d(rg_r)/dr]^{-1}$ (as we neglected contributions to the surface integral from the sides of the cylinder). Taking the derivative of equation (15) with respect to $r$ and defining $\chi = L(g_z/g)^2$, we find

$$g_{z,r} = -(1 + \chi)^{-1}[2\pi G\mu^{-1}\partial\Sigma/\partial r + L(g_r g_z/g^2)g_{r,r}] . \tag{16}$$

From the field equation we have

$$4\pi G\rho = -\nabla \cdot [\mu(g/a_0)\mathbf{g}] = -\mu[g_{z,z} + r^{-1}(rg_r)_{,r}] - a_0^{-1}\mu'\mathbf{g} \cdot \nabla|g| , \tag{17}$$

from which we isolate $g_{z,z}$, by using $g_{r,z} = g_{z,r}$ ($\mathbf{g}$ is a gradient field), and substituting $g_{z,r}$ from equation (16), to get

$$g_{z,z} = (1 + \chi)^{-1}\{-4\pi G\mu^{-1}\rho(r, z) - [1 + L(g_r/g)^2(1 - \chi)/(1 + \chi)]g_{r,r} - r^{-1}g_r + 2L(g_r g_z/g^2)2\pi G[\mu(1 + \chi)]^{-1}\partial\Sigma/\partial r\} . \tag{18}$$

Finally, the Newtonian density $\rho^* = -(4\pi G)^{-1}[g_{z,z} + r^{-1}(rg_r)_{,r}]$ is given by

$$\rho^* = (1 + \chi)^{-1}\mu^{-1}\rho + (4\pi G)^{-1}(1 + \chi)^{-1}\{[-\chi + L(g_r/g)^2(1 - \chi)/(1 + \chi)]g_{r,r} - \chi r^{-1}g_r + 2\chi g_r r^{-1}(1 + \chi)^{-1}\partial \ln \Sigma/\partial \ln r\} . \tag{19}$$

This is the general relation between the Laplacian of the potential field ($\rho^*$) and measurable parameters $\Sigma$ and $g_r$ [which we get from $V(r)$] of the disk in the vicinity of $r$. In the Newtonian case, $L = \chi = 0$, $\mu = 1$, and we get the Poisson equation $\rho^* = \rho$. Note that in general $0 \leq \chi \leq 1$.

In the limit considered in Milogrom (1983*b*) for the Oort problem in the solar neighborhood, we took $(g_z/g)^2 \ll 1$; hence $\chi \approx 0$. We also assumed $G^{-1}g_{r,r} \ll \rho/\mu$ (which amounts to assuming that the local disk density is much larger than the density of the galaxy averaged within radius $r$). The first term on the right-hand side of equation (19) thus dominates the others. This is a good approximation within a height $z$ of order of the scale height $z_0$ of the disk near the Sun. Thus, for small enough $z$ we obtain near the Sun $\rho^* = \mu^{-1}\rho$, as used in Milgrom (1983*b*), and the phantom density (Oort discrepancy) $(\mu^{-1} - 1)\rho$ is positive and proportional to $\rho$ itself.

In the present paper we are interested in what happens at a few scale heights away from the midplane, where the disk density $\rho$ drops below the average density of the galaxy averaged in a sphere of radius $r$. Above this height, the first term becomes negligible, $\Sigma(r, z)$ is saturated, and we can set $\Sigma(r, z) = \Sigma(r) \equiv \Sigma(r, \infty)$, the total disk surface density.

Thus if both $\Sigma(r)$ and the rotation curve $V(r)$ (from which we can determine $g_r$ and $g_{r,r}$, assumed to be independent of $z$ at small $z$) are known in some segment of $r$, we can predict the phantom density along this segment just outside the disk. For a very thin disk equation (19) gives

$$\rho_p(r, z = 0^+) = (4\pi G)^{-1}(1 + \chi)^{-1}(V^2/r^2)\{\chi + (1 - \hat{V}^2)(1 + \chi)^{-1}[L(1 - \chi) - 2\chi] - 2\chi(1 + \chi)^{-1}\hat{\Sigma}\} , \tag{20}$$

where $\hat{V}^2 = d \ln (V^2)/d \ln r$ and $\hat{\Sigma} = d \ln \Sigma d \ln r$. Thus the sign of $\rho_p(0^+)$ depends rather critically on local variations in $V^2$ and $\Sigma$. Places where $\Sigma$ increases abruptly with $r$ (even by a small amount), such as the inner edge of a spiral arm, are potential spots of negative $\rho_p$. When $(g_z/g)^2 \ll 1$ ($\chi \ll 1$, $L$), such as near the Sun, we can write

$$\rho_p(z = 0^+) \approx (4\pi G)^{-1}(V^2/r^2)[\chi(1 - 2\hat{\Sigma}) + L(1 - \hat{V}^2)] . \tag{21}$$

It is important to note that $\rho_p$, as determined by the Oort analysis, is a local quantity, so the relevant values of $\hat{\Sigma}$ and $\hat{V}^2$ should be deduced from variation of $\Sigma$ and $V^2$ over an $r$-range of a few hundred parsecs near the solar position. Large variations over small regions may produce small pockets of $\rho_p < 0$. There may of course be galaxies where equation (20) predicts $\rho_p(0^+) < 0$ over large extents of $r$.

## IV. THE PROSPECTS OF MEASURING NEGATIVE PHANTOM MASS

Nonspherically symmetric isolated configurations exhibit a large variety of negative PMD regions. However, their presence is very difficult to verify, for the following reasons.

1. Global phantom masses are always positive, and $\rho_p < 0$ can be detected only via a local determination of $\nabla \cdot \mathbf{g}$.
2. In the configurations of interest, it is difficult to find test bodies for which the components of the acceleration can be measured.
3. It is not enough to measure one component of the acceleration (such as we do when measuring, for example, the rotation curve of a disk galaxy). For example, suppose we had particles in circular orbits in the symmetry plane of the two equal point masses system. We could then measure the radial component of the acceleration $g_r(r)$ in the symmetry plane. It is, however, easy to see that the contribution of this radial term to $\rho_p$ is positive, as both $g_r$ and $dg_r/dr$ [and hence also $d(rg_r)/dr$] are negative in the $\rho_p < 0$ region in the plane. It is the $dg_z/dz$ contribution which makes $\rho_p$ negative, and so we will have to measure both $g_r(r)$ and $g_z(z)$ (or at least $dg_z/dz$ at $z = 0$). Similarly, in a disk with a depressed central surface density, it is not enough to measure $dg_z/dz$ along the $z$-symmetry axis, as this quantity is negative exactly where $\rho_p < 0$ and thus contributes positively to $\rho_p$.
4. Some of the interesting relevant configurations are not stationary (e.g., that of two point masses). Test particles in the system change their positions on time scales similar to those for global configuration changes. This fact would make it difficult to measure a snapshot acceleration field for the system.

Most of these difficulties can be overcome if a system can be found with X-radiating gas in hydrostatic equilibrium in the region of interest. Forman, Jones, and Tucker (1985) have recently demonstrated the usefulness of such gas as a probe of the mass distribution in systems (early-type galaxies in their case) which are otherwise not accessible to such measurements. Earlier similar studies of the mass distribution around M87 with further references are given in Fabricant, Rybicki, and Gorenstein (1984). If the density and temperature distributions $\rho_g(\mathbf{r})$ and $T_g(\mathbf{r})$, respectively, of such a gas are known and the gas can be assumed to be in hydrostatic equilibrium, we can write for the Newtonian density

$$\rho^* = -(4\pi G)^{-1}(k/m)[T_g\nabla^2 q + \nabla^2 T_g + \nabla T_g \cdot \nabla q] \,, \tag{22}$$

where $m$ is the mean molecular weight and $q = \ln(\rho_g)$. In hydrostatic equilibrium, $T_g$ is a function of $\rho_g$, so we have

$$\rho^* = -(4\pi G)^{-1}\left(\frac{k}{m}\right)\left[\left(T_g + \frac{dT_g}{dq}\right)\nabla^2 q + \left(\frac{dT_g}{dq} + \frac{d^2 T_g}{dq^2}\right)(\nabla q)^2\right]. \tag{23}$$

Now let us turn to spherical systems where a negative PMD is manifested by a faster-than-Keplerian fall-off in the rotation curve [i.e., a range where $V^2(r)r$ decreases with $r$]. From equation (10) and the paragraph below it, we see that it is never possible to find $\rho_p < 0$ in spherical systems if $I'(x) = \mu(x) + x\mu'(x) \leq 1$ for all $x$. Suppose, however, that the inverse inequality holds in some range of $x$. This can only happen near $x = 1$, since for $x \ll 1$, $I'(x) \approx 2x \ll 1$, and for $x \gg 1$ we have $I' \approx 1$. If $\mu$ is concave as we always assume (i.e., $L \leq 1$), the maximum value which $I'$ can have is 2, but this occurs only for an extreme choice of $\mu$ ($\mu = x$ for $x \leq 1$ and $\mu = 1$ for $x > 1$). In general, $I'_{\max} - 1 \ll 1$ (for example, for $\mu = 1 - e^{-x}$, $I'_{\max} = 1 + e^{-2}$). As a result, the first term in equation (10) which may make $\rho_p < 0$ and which contains a factor $(I')^{-1} - 1$ is smaller in absolute value than the second positive term (always larger than a third of the average density within $r$). For example, if the density $\rho(r)$ decreases with $r$ so that $\rho(r) \leq 3M/4\pi r^3$, $\rho_p$ is always positive if $I'_{\max} \leq 3/2$. Thus, even in the most favorable cases the effect is small, and the maximum value of $|\rho_p|$ we can get for $\rho_p < 0$ is $\rho/6$.

## V. CONCLUSION

Every consequence of MOND which we have tested so far can be reproduced in the framework of the hidden mass hypothesis by assuming ad hoc the existence of the necessary distribution of hidden matter. It is thus particularly important to identify consequences of MOND which are in clear conflict with HMH.

In the present paper we show that not all the gravitational fields produced by realistic mass configurations, according to MOND, can be reproduced in Newtonian dynamics. For some fields, insistence on Newtonian dynamics will require negative masses as sources.

We have shown that the regions of such negative " phantom masses " are predicted by MOND to appear, among other systems, in low-acceleration systems of the following types:

a) In the symmetry plane of any two equal mass system or in general for any two point mass systems such as a binary galaxy.

b) Along the azimuthal axis of a disk with a central depression in surface density.

c) Just outside a thin galactic disk in regions where the surface density increases sharply, such as at the inner edge of a spiral arm.

d) Near any gravitating system which lie in an (approximately) constant external low-acceleration field. For example, the details of the stellar dynamics in open clusters, dwarf spheroidal galaxies, and the outer (low surface density) regions of globular clusters are expected to be strongly affected by the field of the parent galaxy and in particular involve regions of negative phantom densities. The same holds for the Oort cloud of long-period comets in the combined Sun-Galaxy field.

We are not able, at present, to suggest a concrete feasible measurement that will test the existence of negative phantom mass. The main purposes of this paper are to point out the effect itself and provide a basis for further searches, in the hope that a practical test will eventually be found. Some hope is offered by the use of hot X-radiating gas as a collection of probes for measuring the acceleration fields of galaxy systems, as described recently by Forman, Jones and Tucker (1985).

*Note added in manuscript.*—All the results of this paper can be carried over to the case where MOND is formulated in terms of more than one potential. Say the MOND acceleration is given by $\mathbf{g} = \Sigma_i \mathbf{g}_i$, with each of the $\mathbf{g}_i$ satisfying an equation like (2) with corresponding $\mu_i$ and $G_i$. We have by definition $G = \Sigma G_i$. The Newtonian density is given by $\rho^* = -(4\pi G)^{-1}\Sigma_i \nabla \cdot \mathbf{g}_i$. Defining $\rho_i^* = -(4\pi G)^{-1}\nabla \cdot \mathbf{g}_i$ and $\rho_i = (G_i/G)\rho$, we apply the preceding discussion to $\rho_i^*$ and $\rho_i$ and use $\rho = \Sigma_i \rho_i$, $\rho^* = \Sigma_i \rho_i^*$.

This work has branched out of my work with Jacob Bekenstein on the Lagrangian theory for MOND. I thank Jacob for many discussions on the topics treated in this paper. The work was supported by grants from the Israeli Academy of Science and a grant from Minerva.

## REFERENCES

Bekenstein, J., and Milgrom, M. 1984, *Ap. J.*, **286**, 7 (Paper II).
Faber, S. M., and Gallagher, J. S. 1979, *Ann. Rev. Astr. Ap.*, **17**, 135.
Fabricant, D., Rybicki, G., and Gorenstein, P. 1984, *Ap. J.*, **86**, 186.
Finzi, A. 1963, *M.N.R.A.S.*, **27**, 21.
Forman, W., Jones, C., and Tucker, W. 1985, *Ap. J.*, **293**, 102.
Freeman K. C. 1970, *Ap. J.*, **139**, 284.
Hill, E. R. 1960, *Bull. Astr. Inst. Netherlands*, **15**, 1.
King, I. R., 1977, in *Highlights Astr.*, **4**, 41.
Kormendy, J. 1977, *Ap. J.*, **217**, 406.
Milgrom, M. 1983*a*, *Ap. J.*, **270**, 365 (Paper I).
———. 1983*b*, *Ap. J.*, **270**, 371.
———. 1983*c*, *Ap. J.*, **270**, 384.
———. 1984, *Ap. J.*, **287**, 571.
———. 1986, *Ap. J.*, **302**, 617.
Sanders, R. H. 1984, *Astr. Ap.*, **136**, L21.
Tohline, J. E. 1983, in *IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Dordrecht: Reidel), p. 205.

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