MOND as a regime of quantum gravity

Lee Smolin[*]

Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2J 2Y5, Canada

We propose that there is a regime of quantum gravity phenomena, for the case that the cosmological constant is small and positive, which concerns physics at temperatures below the de Sitter temperature, or length scales larger than the horizon. We observe that the standard form of the equivalence principle does not apply in this regime; we consider instead that a weakened form of the equivalence principle might hold in which the ratio of gravitational to inertial mass is a function of environmental parameters. We consider possible principles to determine that function. These lead to behavior that, in the limit of $\hbar \to 0$ and $c \to \infty$, reproduces the modifications of Newtonian dynamics first proposed by Milgrom. Thus modified newtonian dynamics is elucidated as coding the physics of a novel regime of quantum gravity phenomena. We propose also an effective description of this regime in terms of a bimetric theory, valid in the approximation where the metric is static. This predicts a new effect, which modifies gravity for radial motions.

I. INTRODUCTION

As discovered by Vera Rubin and her colleagues [1], and confirmed by many subsequent observations, the rotational velocities of stars and gas in the outer regions of spiral galaxies depart from the $\frac{1}{\sqrt{r}}$ behavior we would expect given Newtonian gravity and the observed distributions of baryonic matter. Instead, the velocities flatten out to constant values, v, given by a simple function of the total baryonic mass of the galaxy, M_b,

$$v^4 = G a_0 M_b, \tag{1}$$

where a_0 is an acceleration scale, which can be read off the data to be

$$a_0 = 1.2 \times 10^{-8} \text{ cm/s}^2. \tag{2}$$

This empirical relation is called the *baryonic* Tully-Fisher law [2]. There are several remarkable features of this law:
  (i)   There is remarkably little scatter, given that this is a summary of astronomical data, and galaxies are messy objects, with strongly nonlinear dynamics and histories [3]. The relation appears to hold in a diverse range of disk galaxy types.
 (ii)   The value of a_0 appears universal.
(iii) a_0 is close to the acceleration of the Universe, $a_\Lambda = c^2 \sqrt{\Lambda}$,

$$a_0 \approx \frac{a_\Lambda}{8.3}. \tag{3}$$

 (iv)   The relation involves the baryonic mass of a galaxy.

In 1983 Milgrom proposed [4] that the discrepancy of rotational velocities from Newtonian expectations could be expressed by a universal relation between the measured radial acceleration

$$a_{\text{obs}} = \frac{v^2}{r} \tag{4}$$

and the acceleration predicted by Newtonian theory on the basis of the observed baryonic masses,

$$a_N^i = \nabla^i U \tag{5}$$

of the form

$$a_N = a_{\text{obs}} G^{-2} \left(\frac{a_{\text{obs}}}{a_0} \right). \tag{6}$$

We can invert this to find a function $F^2(\frac{a_N}{a_0})$ such that

$$\frac{a_{\text{obs}}}{a_N} = F^2 \left(\frac{a_N}{a_0} \right). \tag{7}$$

For small accelerations, compared to a_0, Milgrom proposed that this be chosen to reproduce the Tully-Fisher relation. This requires that for small $a_N \ll a_0$

$$a_{\text{obs}} = \sqrt{a_N a_0}. \tag{8}$$

On the other hand, for large a, Newtonian gravity should be recovered.

In a recent paper, McGaugh, Lelli, and Schombert (MLS) report [5] strong confirmation of an empirical relation of this form, first proposed by Milgrom [4]. They measure $F(\frac{a_N}{a_0})$ in a survey of rotation curves of 153 galaxies in the SPARC database [6]. They measure

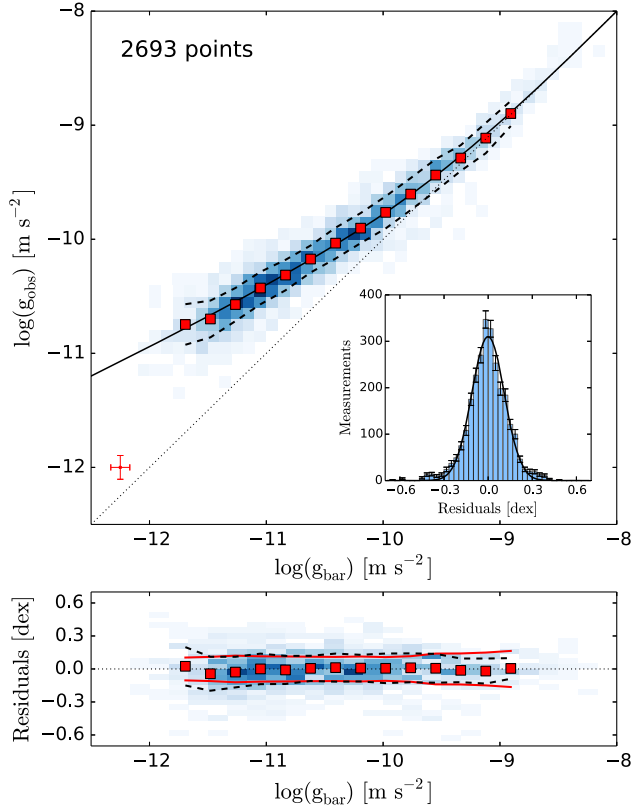[*]lsmolin@perimeterinstitute.ca

FIG. 1. The empirical radial acceleration relation, as shown in Fig. 3 of [5]. Data are taken from the SPARC data base [6].

a_{obs}, the actual radial acceleration by (4) at 2693 radii on these rotation curves. At the same radii they estimate the Newtonian gravitational potential from baryons as observed in stars, gas, and dust, and so determine a_N. They discover that the data are well described by a simple empirical relation of the form of (7), as shown in Fig. 1. As they note, it is amazing that such a relation exists over a wide range of galaxy types, sizes, and morphologies, as this represents the observed accelerations only by a function of the Newtonian accelerations due to baryons.

Furthermore MLS are able to fit a simple form for $F(a)$ (previously proposed in [7]), to the data which are [5,8]

$$F^2(a_N) = \frac{1}{1 - e^{-\sqrt{\frac{a_N}{a_0}}}}. \quad (9)$$

In addition to these galactic phenomena, there is evidence for departures from Newtonian gravity on still larger scales, in observations of galactic clusters and large scale structure. These diverse phenomena suggest two hypotheses:

(1) *Dark matter:* Much of the matter in the Universe is nonluminous and, likely, nonbaryonic.
(2) *Gravity is modified* in a regime of small acceleration.

Diverse theories have been introduced as elaborations on one or the other hypothesis. Two remarks need to be made.

First, these are not mutually exclusive, as the galactic and extragalactic phenomena may have different explanations. In the early 1900s there were two anomalies observed in the planetary orbits; both Mercury and Neptune were off track. The first of these was explained by a modification of gravity, the latter by the discovery of heretofore dark matter, also known as Pluto [9].

Second, determined efforts have failed to result in a nongravitational detection or production of dark matter.

The hypothesis that the galactic scale phenomena are explained by a modification of Newtonian gravity was first made by Milgrom in 1983 [4], who suggested that the universal relation (7), (8), (9) he proposed was due to the action of a fundamental modification in the laws of gravity.

Many years later, it appears to be fair to say that at the scale of galaxies, *Milgrom's hypothesis has held up well, when compared with data* [10]. Indeed, Milgrom's hypothesis must be credited with a number of predictions which were subsequently confirmed, such as the behavior of low surface brightness and dwarf galaxies.[1] To the extent to which this is true, this represents a challenge to the dark matter hypothesis. While it is certainly conceivable that the success of Milgram's empirical relation (7), (8), (9) might be explained by a dark matter model, that model would have to explain the following:

(1) Why is the scatter in the Tully-Fisher relation smaller when expressed in terms of the baryonic mass than in terms of the presumed total mass [3]?
(2) Why does the acceleration relation hold widely, with small scatter, over many galaxy types, when the baryonic and dark matter distributions are weakly coupled?
(3) Why do specific features of individual galactic rotation curves reflect the observed baryonic matter distributions, if most of the mass is dark matter, only weakly coupled to baryons?
(4) Where does the observed acceleration scale a_0 come from, why is it universal, and why is it close to the cosmological acceleration?

At the same time, as soon as one looks beyond the context of galactic rotation curves, the MOND hypothesis faces severe challenges of its own:

(i) It is so far expressed as a modification of non-relativistic, Newtonian dynamics. While there are attempts to embed MOND in a special or general relativistic framework [17,18], none so far are compelling theoretically. One issue is stability, due to the incorporation of a nongauge invariant vector field. Another issue is nonlocality, which appears necessary to code a potential that falls off slower than inverse distance.

---

[1]For reviews of modified newtonian dynamics (MOND), see [11]; some other relevant papers on MOND are [12–16].

   (ii) These extended versions of MOND do not do well explaining the observed behaviors of galactic clusters, lensing, and large scale structure.

   (iii) Given the acceleration relation, the theoretical proposal is underdetermined. One could regard MOND either as a modification of the gravitational field equations or as a universal modification of the principle of inertia, for small acceleration. The evidence also underdetermines the theory, because we mostly observe the low acceleration regime by steady, circular motion.

It is the aspect of nonlocality, in particular, that suggests that in the, still unlikely, case that MOND is true, it expresses a fundamental departure from known principles. This of course seems unlikely, but we may note several interesting aspects of this problem.

First, we have no other probe of the regime of ultrasmall accelerations, $a < a_\Lambda$, except the motions of stars and gas in galaxies.

Second, we may note that an accelerating observer has, in Minkowski spacetime, an acceleration horizon at a distance, $l_a = \frac{c^2}{a}$. This is, among other things, the peak wavelength of the corresponding Unruh radiation [19]. A weakly, but uniformly, accelerating detector is then, possibly, a very large object, at least from the viewpoint of its effect on the vacuum of the quantum fields. It is then very interesting that the regime $a < a_\Lambda$ is also the regime when the peak Unruh wavelength l_a is of the order of the distance to the cosmological horizon, R, defined by $\frac{\Lambda}{3} = \frac{1}{R^2}$.

Recent research has highlighted the far infrared physics of gauge and gravitational physics [20]. We suggest that MOND may be a surprise hiding in the nonlinear dynamics of these soft modes on the scale of the cosmological horizon.

Third, we expect that fundamental phenomena, related to quantum gravity, are to be expressed in terms of all four of the fundamental constants, $\hbar, G, c$, and Λ. But it is conceivable that for a_Λ fixed there is a regime of phenomena delineated by a_Λ which survives a limit in which $\hbar \to 0$ and $c \to \infty$. (Indeed the peak wavelength of Unruh radiation is a quantum phenomena in which two $\hbar$'s have canceled.) We may call this, very tentatively, the cosmological constant dominated regime of quantum gravity, for it concerns length and acceleration scales comparable to R and a_Λ.

We proceed to investigate what can be said about such a regime of quantum gravity.[2]

---

[2]The idea that MOND is an expression of quantum gravity has been considered earlier from various points of view by Milgrom [21], van Putten [22], Verlinde [23], Hossenfelder [24], Woodard [25], Modesto and Randono [26], Minic *et al.* [27], Hendi and Sheykhi [28], McCulloch [29], Klinkhamer [30], Pikhitsa [31], and others [32], including the author [33,34]. The argument below is, I believe, on the whole, novel, but in places it overlaps some of these discussions.

II. THE POSITIVE COSMOLOGICAL CONSTANT DOMINATED REGIME OF QUANTUM GRAVITY

Whatever is the right quantum theory of quantum gravity, it will depend on four dimensional constants, $\hbar, G, c$, and Λ. Different regimes of quantum gravity phenomena can be captured and delineated by studying limits such as $\hbar \to 0$ that must recover general relativity and $G, \Lambda \to 0$ that recovers[3] quantum field theory on flat spacetime. The holographic regime explored by the AdS/CFT correspondence is for negative Λ with $\frac{1}{N} = \hbar G |\Lambda|$ small [35].

In the past few years, new regimes have been explored, such as the relativity locality regime in which G and $\hbar$ are both taken to zero, but holding the Planck energy, given by their ratio, fixed [36,37].

We want to explore the physics of a novel regime of physical phenomena associated with length scales, or wavelengths greater than or comparable to R, and accelerations small compared to a_Λ. This new regime involves novel physical phenomenon which survive the limits $\hbar \to 0$ and $c \to \infty$.

What principles might govern such phenomena?

One that does not directly is the equivalence principle. To see why, let us review it.

As presented by Einstein, the equivalence principle consists of two statements:

EP1: Freely falling observers, whose extent, L, is much smaller than the radius of curvature, $\mathcal{R}$ (and hence R), observe special relativity to hold to zeroth order in $\frac{L}{\mathcal{R}}$. That is, the zeroth order effects of gravity can be eliminated by free fall.

EP2: Uniformly accelerated observers see themselves to be in a uniform gravitational field, to zeroth order in $\frac{L}{\mathcal{R}}$. That is, gravity can be mocked up by uniform acceleration, to zeroth order in $\frac{L}{\mathcal{R}}$.

These are equivalent classically, so long as we respect the restriction to phenomena on scales much smaller than the radius of curvature.

The equivalence principle has something to say about the relationship of the inertial mass, m_I, to the gravitational mass, m_g. The first may be defined by invoking the conservation of momentum in scattering experiments. The inertial mass is then defined (for $v \ll c$) as the ratio of momentum to velocity. The gravitational mass is defined independently, as a measure of the strength by which a body is affected by the gravitational field.

The universality of free fall is normally taken to imply $m_g = m_I$. But consider the possibility that free fall is universal at a given time and place, but that the response of a body to the gravitational field could differ over time and space, as a function of environmental or other parameters. In this case it is natural to extend the principle of equivalence from $m_g = m_I$ to [33]

---

[3]From a perturbative perspective.

$$\frac{m_I}{m_g} = Z(\text{universal function of environment,}$$

$$\text{independent of masses}). \tag{10}$$

This suffices to let us transform gravity away by going into free fall, or mock gravity up by uniformly accelerating. This defines a universal function Z which can depend on global or environmental parameters, but does not depend on the masses themselves.

To deduce more, we have to add that the physics seen by a freely falling observer includes Newtonian gravity in the limit of small velocities. This then implies that

$$m_g = m_I, \quad \text{so } Z = 1. \tag{11}$$

Now notice that while consistent classically, there is a tension between the two parts of the equivalence principle. This has to do with the different ways in which one can try to extend them from a limiting case for which $\frac{L}{\mathcal{R}} \to 0$.

The tension is due to the fact that EP1 becomes exact in the limit $L \ll \mathcal{R}$, whereas EP2 becomes exact in the limit of perfectly uniform acceleration; i.e. it requires not a small observer, but a static spacetime. But uniform acceleration implies a length $l_a = \frac{c^2}{a}$. This is the distance from the detector to the horizon created by the detector's uniform acceleration.

This tension becomes a conflict quantum mechanically, because that length is the peak wavelength of the Unruh radiation, which is then the scale over which the accelerating detector disturbs the vacuum.

We already know that the equivalence principle does not easily coexist with quantum field theory because the vacuum state of a quantum field is not a local object. One cannot localize a quanta of a massive field to smaller than its Compton wavelength, $\lambda_C = \frac{\hbar}{mc}$. But the Unruh effect makes this conflict more fraught.

For a massless field, the Unruh effect involves for small acceleration an arbitrarily low temperature,

$$T_U = \frac{\hbar a}{2\pi c}, \tag{12}$$

whose corresponding thermal distribution has a peak wavelength,

$$\lambda_U = \frac{\hbar c}{T_U} = l_a = \frac{c^2}{a}. \tag{13}$$

Thus, the limit of small acceleration or low temperature involves arbitrarily long length scales, which brings us into conflict with the condition required by the equivalence principle that the phenomena associated with the detector have to be smaller than the radius of curvature.

EP1 requires that all lengths are much smaller than $\mathcal{R}$ and hence are smaller than R. We are consistent with this only if the acceleration is large enough that the peak wave length of the Unruh radiation is much smaller than the radius of curvature, and hence R. This defines a phase of quantum gravity, dominated by EP1.

But there could be another phase of quantum gravity not dominated by EP1 for small acceleration, $l_a > R$ or $a < a_\Lambda$. Here, a restricted form of EP2 can hold, but not EP1.

Thus, when there is a positive cosmological constant, the resulting length scale, R, and acceleration scale, a_Λ, divide physical phenomena into two regimes.

(i) *The equivalence principle dominated regime.* Those phenomena, all of whose lengths are less than R, and all of whose accelerations are greater than a_Λ, fall into the normal regime dominated by the equivalence principle. The classical description of this physics is general relativity.

(ii) *The cosmological constant dominated regime* includes phenomena whose length scales are greater than R and/or whose acceleration scales are less than a_Λ.

One way to delineate the two regimes is to express them in terms of the relationship between the inertial and gravitational masses,

$$\frac{m_I}{m_g} = Z. \tag{14}$$

In the equivalence principle dominated regime, $Z = 1$. In the Λ dominated regime this can depend on other, global or environmental, quantities (but not the masses themselves).

We need a principle to determine how Z depends on various quantities. The idea is that all the nonlocal and far infrared physics connected with the cosmological horizon scale are reflected in the dependence of Z on global or environmental parameters. Thus, so far as local physics is concerned, we may apply classical reasoning to the physics in the cosmological constant dominated regime, with the exception of the renormalization of the ratio between the gravitational and inertial masses of a body.

III. THE THERMAL EQUIVALENCE PRINCIPLE

What should an observer in the cosmological constant dominated regime observe? Guidance is provided by the crucial result of Narnhofer *et al.* [38] and Deser and Levin [39], who find that a uniformly accelerating detector in de Sitter spacetime observes a thermal spectrum with a temperature,

$$T_{\text{DL}} = \sqrt{T_{dS}^2 + \left(\frac{\hbar a}{2\pi c}\right)^2}, \tag{15}$$

where the de Sitter temperature is

$$T_{\text{dS}} = \frac{\hbar c^2}{2\pi}\sqrt{\frac{\Lambda}{3}}. \tag{16}$$

Note that $T_{\mathrm{DL}} \geq T_{\mathrm{dS}}$, so the latter is a minimum temperature for equilibrium.

The peak wavelength is then

$$\lambda_{\mathrm{DL}} = \frac{1}{\sqrt{\frac{1}{l_a^2} + \frac{1}{R^2}}}. \qquad (17)$$

The peak wavelength is a pure quantum phenomenon, but so is the Unruh temperature, which means that both are proportional to $\hbar$. So, when we compute the peak wavelength, the $\hbar$'s cancel and we are left with a seemingly classical criterion marking a boundary of a phase of a quantum phenomenon,

$$a < a_\Lambda. \qquad (18)$$

For $a \gg a_\Lambda$, $\lambda_U \approx l_{\mathrm{acc}} \ll R$ and we are in the equivalence principle dominated phase. But for $a < a_\Lambda$, $\lambda_U \approx R$ and we are in the new, cosmological constant dominated phase. Indeed not only do the $\hbar$'s cancel, but, when expressed in terms of accelerations, the c's also cancel, so we have the possibility of a quantum gravity effect modifying Newtonian dynamics.

Now we do not live in an exact de Sitter spacetime, so how is the Deser-Levin temperature relevant? We propose that it applies to an observer in the cosmological constant dominated regime, for $\Lambda > 0$.

To be precise, let us restrict ourselves to static observers, whose worldlines are generated by timelike killing fields of static spacetimes. Such an observer observes an acceleration, a^a, and a temperature, T. The T is a minimal temperature they can detect, when all sources of thermal radiation, including black hole horizons, but not counting the de Sitter temperature, are turned off or shielded. a^a is the acceleration they observe between their trajectory and the trajectory of a freely falling body at their location.

The *thermal equivalence principle* (TEP) asserts that *the temperature, T, experienced by the static detector is related to the magnitude of the acceleration, a, by the Deser-Levin formula* (15).

This incorporates the idea of the universality of free fall in a gravitational field. a^a is the acceleration the observer has to exert to stay static by following a timeline killing field.

Note that while this TEP extends EP2, there is no extension of EP1 to this regime. This is because the phenomena described by the TEP are on scales larger than R and are thus not compatible with the restriction required by EP1 that the relevant phenomena be small compared to R. Indeed, it appears there is no simple way to characterize the spectrum that freely falling or orbital detectors observe in asymptotically de Sitter spacetimes.[4]

A. Phase transition versus phase boundary

The de Sitter temperature is the minimum temperature that can be measured in equilibrium in a static spacetime with positive cosmological constant. Similarly, the peak wavelength is bounded above by the cosmological horizon,

$$\lambda_{\mathrm{DL}} \leq R. \qquad (19)$$

While we see many stars with $a < a_\Lambda$, in terms of temperature and peak wavelength it appears that the cosmological constant dominated regime lives in the neighborhood of a phase boundary characterized by approaching the minimal temperature or the maximal peak wavelength.

This may then be a kind of quantum critical phenomena, at a minimal temperature, which is as close as an observer in de Sitter spacetime can get to $T = 0$.

If so, we may expect phenomena in this region to scale. Indeed, Milgrom proposes that MOND behavior is characterized by invariance under scaling [40]. He proposes that in the regime of small acceleration, equations of motion are to scale uniformly under

$$x^i \to \lambda x^i, \qquad t \to \lambda t \qquad (20)$$

with masses held fixed, as are fixed constants including Λ, G, and a_0. (c is automatically invariant, while $\hbar$ requires a separate discussion, which we postpone as the factors of $\hbar$ cancel in the limit we are discussing here.)

We note that Newton's law,

$$a_N^{\hat{r}} = -\frac{m_g}{m_I} \frac{GM}{r^2}, \qquad (21)$$

is not homogeneous under scale transformations, so long as $m_g = m_I$.

We can use the possibility that $m_g = Z^{-1} m_I$ to modify the acceleration law to make it scale invariant in the Λ dominated regime.

Since Z characterizes the Λ dominated phase, we expect it may depend on Λ, but as it is dimensionless, it will have to depend on the ratio of R to other lengths. A length of interest is l_a; hence we expect[5]

$$Z\left(\frac{R}{l_{\mathrm{acc}}}\right) = Z\left(\frac{T_U}{T_{\mathrm{dS}}}\right) = Z\left(\frac{a}{a_\Lambda}\right). \qquad (22)$$

We note that when we express Z in terms of accelerations, the c's cancel. If we impose scale invariance, then Z must scale like λ^{-1}. Hence scale invariance requires

$$Z\left(\frac{a}{a_\Lambda}\right) = \frac{a}{a_\Lambda}. \qquad (23)$$

[5]Note here that a is the magnitude of the acceleration of the static observer.

This gives us the MOND acceleration relation,

$$a^i \frac{a}{a_\Lambda} = \nabla^i \phi_N = -\hat{r}^i \frac{GM}{r^2}. \tag{24}$$

Note that the original observation which motivated this whole story, of flat rotation curves, is a verification of scaling, as the velocity, which is scale invariant, becomes independent of radius.

IV. ENTROPIC DEFINITION OF INERTIAL MASS

How could there arise, from fundamental theory, a scaling of the ratio of inertial to gravitational mass? To explore this, we bring in the entropic gravity hypothesis of Verlinde [41].[6] However, we interpret Verlinde's idea as an entropic elucidation of the concept of—and principle of—inertia. The idea is that the inertial mass arises from consideration of how well a particle can be localized.

Following [22,41], we assign an entropy to how localizable a particle is by a static detector confined within an horizon. The minimal localization a static, accelerating detector can make of a particle is that it is within some distance scale, L, defined by a detector. The detector should be smaller than her horizon, so $L \leq R$. The best localization of the particle she can make is that the particle is within a Compton wavelength, λ_C. We can then define an entropy of localization which is the negative of the information gained by such a localization, which is proportional to the ratio of the best and worst possible localization,

$$S_{\text{loc}} = 2\pi \frac{L}{\lambda_C}. \tag{25}$$

This entropy counts the information potentially available by localizing the particle.

First we check that this gives Newtonian dynamics when the cosmological constant is turned off, which means we take $R \to \infty$ and $T_{\text{DL}} = T_U$. The corresponding free energy is (for vanishing cosmological constant)

$$W = TS = \frac{\hbar a}{2\pi c} S_{\text{loc}} = \frac{L}{\lambda_C} \frac{\hbar a}{c}. \tag{26}$$

Now we follow Verlinde in defining an entropic force as T times the differential of S gotten by moving the particle a distance ΔL within by the detector,

$$F = T \frac{\Delta S}{\Delta L} = \frac{\hbar}{\lambda_C c} a. \tag{27}$$

We can now *define* the inertial mass in terms of the Compton wavelength, by

$$m_I \equiv \frac{\hbar}{\lambda_C c} \tag{28}$$

to find Newton's second law,

$$F = m_I a. \tag{29}$$

Hence, we have derived the principle of inertia, and inertial mass, from entropic considerations.

But this argument assumed the cosmological constant vanishes. What happens when we turn on a small Λ? We consider the case of positive Λ, in which case there is a cosmological horizon, at a distance R defined by $\frac{\Lambda}{3} = \frac{1}{R^2}$.

In the presence of Λ there is an irreducible de Sitter temperature T_{dS}. Because the temperature in equilibrium cannot be reduced to below, T_{dS}, hence we can posit that the change in free energy is

$$\Delta W = (T_{\text{DL}} - T_{\text{dS}}) \frac{\Delta S}{\Delta L}. \tag{30}$$

It is necessary to subtract off T_{dS} in the entropic derivation of the force law; otherwise a freely falling detector with $a = 0$ would experience a force.

To get the force we take again the derivative with respect to ΔL to find

$$F = (T_{\text{DL}} - T_{\text{dS}}) \frac{\Delta S}{\Delta L} = \frac{a^2}{2a_\Lambda} \frac{\hbar}{\lambda_C c}. \tag{31}$$

This should be set equal to the Newtonian force law, $F = m_g a_N$.

Assuming that $m_g = m_I$, this leads to the MOND dynamics

$$a^2 = 2a_N a_\Lambda. \tag{32}$$

We can explain this by noting that when there is no cosmological horizon the entropic force is proportional to $T_U \propto a$, whereas in the presence of a positive cosmological constant we have, for small a,

$$F \propto (T_{\text{DL}} - T_{\text{dS}}) \propto \frac{a^2}{2a_\Lambda}. \tag{33}$$

Alternatively, the deviation from Newtonian physics seen in MOND can be expressed by a renormalization of the relation between inertial and gravitational masses, inserted into the standard Newtonian laws of gravity and inertia,

---

[6]The general relativistic version of this idea is described in [42–44]. A version of Verlinde's argument valid in loop quantum gravity is in [45]. Other proposals to derive MOND from entropic gravity are described in [22,23,26].

$$m_I = m_g Z\left[\frac{a}{a_\Lambda}\right], \qquad (34)$$

with $Z[\frac{a}{a_\Lambda}] = \frac{a}{a_\Lambda}$ for $a < a_\Lambda$.

Notice several things: The temperature is always that seen by a static observer. This is an accelerated observer, locally. It is not on a geodesic, so the acceleration is nonzero for Newton, special, and general relativity. The static observer may observe particles in free fall and in orbit. We have no reason to worry about what temperatures those particles would see were they equipped with detectors. This explains how the results can be valid for circular motion as seen in spiral galaxies.

Second, the thermodynamics we employ is equilibrium thermodynamics. There must be periods of nonequilibrium behavior, while a galaxy is forming. Our simple model so far has nothing to say about such behavior; it assumes the galaxy has been static for a long time.

Third, from this argument we recover the scaling (20) in the cosmological constant dominated regime. This supports the picture that MOND is a kind of critical phenomena related to being near the boundary of minimal temperature at $T = T_{\mathrm{dS}}$.

Finally, this entropic origin of inertia is reminiscent of the idea that inertial motions are those that see the least thermal fluctuations [46,47].

V. EFFECTIVE DESCRIPTION

The conclusion we come to, from the foregoing, is that MOND can be described as a modification of the principle of inertia, arising from the insertion, into the ratio of gravitational to inertial mass, of a renormalization factor, Z, which is a function of global or environmental variables. So the equations of motion for a star in a galaxy, with trajectory $x^i(t)$, become modified to

$$\ddot{x}^i = Z^{-1} g^{ij} \nabla_j \phi, \qquad (35)$$

where g^{ij} is the inverse of the spatial metric. Z can be read as

$$Z = \frac{m_I}{m_g} = Z\left(\frac{T_U}{T_{dS}}\right) = Z\left(\frac{a}{a_\Lambda}\right), \qquad (36)$$

where T_U is the Unruh temperature observed by a static detector held at that point, and a is the magnitude of the acceleration necessary to hold that detector in place.

We can express this as a bimetric theory, in which the metric that governs the motion of particles is not the same as the metric that solves the Einstein equations.

We note that for a particle in free fall or orbital motion in the static gravitational field, $\ddot{x}^i = -a^i$, by the equivalence principle (where a^i is evaluated at their location). But these are conceptually distinct and generally not equal. a^i is the acceleration of a static detector held at a fixed position in the gravitational field, while $\ddot{x}^i$ is the acceleration of a particle in free fall. Moreover, the dependence on a^i is actually a proxy for a dependence of the ratio of gravitational to inertial mass on the local temperature. Thus, we do not have an issue of equations of motion of higher order in time derivatives. We do have a question as to how the dependence of T_U evolves in time in nonstatic configurations or nonequilibrium states.[7]

In earlier sections, I argued that both scale invariance and a thermodynamic or entropic origin of the first law suggests

$$Z\left(\frac{a}{a_\Lambda}\right) = \frac{a}{a_\Lambda} \quad \text{for } a < a_\Lambda \qquad (37)$$

and $Z = 1$ otherwise.

Let us next see if we can describe this phenomena in terms of an effective picture in the language of general relativity. We can, in a limited sense that applies only to static spacetimes, as follows.

We start with (35), which tells us that we can describe MOND as a replacement in the spherically symmetric case

$$g^{rr}\partial_r \phi \to Z^{-1} g^{rr} \partial_r \phi. \qquad (38)$$

Recall that in the Newtonian limit of general relativity

$$g_{00} = 1 - \frac{2\phi}{c^2} = f. \qquad (39)$$

This suggests that the metric that governs the motion of particles in a static gravitational field is modified by

$$g_{rr} \to \tilde{g}_{rr} = Z g_{rr}, \qquad (40)$$

with the rest unmodified.

We may call this the *thermal metric*.[8] We recall that the leading term in the geodesic equation, in the Newtonian limit of the spherically symmetric case, is

$$\ddot{x}^r = \Gamma^r_{00}\dot{x}^0 \dot{x}^0 \approx \Gamma^r_{00} c^2, \qquad (41)$$

where

$$\Gamma^r_{00} = -\frac{1}{2} g^{rr} \partial_r g_{00}. \qquad (42)$$

This gives to leading order

$$\Gamma^r_{00} \to \tilde{\Gamma}^r_{00} = Z^{-1} \Gamma^r_{00}. \qquad (43)$$

---

[7]As a result of the limitation to motion in static spacetimes, we are unable to address the issue of instabilities.

[8]This name was suggested by Matteo Smerlak.

Then we have a prescription for encoding MOND as an effective description within general relativity, in the case that the metric is static:

(1) Leave the field equations in the original metric, g_{ab}, unmodified.

(2) The motion of a particle in a static gravitational field is given by the geodesics of the modified, thermal metric, i.e. from the variational principle

$$S = m_I \int ds \sqrt{\tilde{g}_{ab} \dot{x}^a \dot{x}^b}, \qquad (44)$$

where the metric $\tilde{g}_{ab}$ is defined by the radial component being rescaled by $g_{rr} \to \tilde{g}_{rr} = Z[\mathcal{F}] g_{rr}$.

(3) We then study the variational principle,

$$\delta_{\mathcal{F}} S = 0. \qquad (45)$$

The meaning of $\delta_{\mathcal{F}}$ is that when varying the action, the path is to be varied with $\mathcal{F}$ held as a fixed function. This gives us $x^a(s)_{\mathcal{F}}$, where there is a solution for every fixed function, $\mathcal{F}$. We are then interested in those solutions that satisfy[9] (45) together with

$$\mathcal{F} = \frac{a}{a_\Lambda}. \qquad (46)$$

This may be thought of as a condition of equilibrium. The idea is that Z is defined as a function of ratios of temperatures, as a result of processes of equilibration to the Deser-Levin temperature. In the MOND, or cosmological constant dominated regime, this is close to the de Sitter temperature, so equilibration takes place over cosmological time scales. The particle motion takes place over much shorter time scales and is determined by the variational principle with respect to which the thermal background, and hence Z, can be considered fixed. Equation (46) is a consistency condition that expresses the fact that the dynamical and thermal equilibrium of the galaxy is made up of vast numbers of stars.

This reproduces to zeroth order in $\frac{v^2}{c^2}$ the picture we arrived at above. There is, however, an additional term in the modified geodesic equation which survives the non-relativistic limit $c \to \infty$. It comes from the term in the geodesic equation $\tilde{\Gamma}^r_{rr} \dot{r}^2$. Including it, the full radial part of the geodesic equation reads

$$\ddot{x}^r = -\frac{1}{Z} g^{rr} \partial_r \phi - \frac{\dot{r}^2}{c^2} \partial_r \phi - \frac{\dot{r}^2}{\partial_r \phi} 2\pi G \rho \Theta[Z-1]. \quad (47)$$

_____
[9] In the MOND limit.

We also use the fact that

$$Z = \frac{|\partial_i \phi|}{a_0}, \qquad (48)$$

where $\nabla^2 \phi = 4\pi G \rho$. The second term is standard from the Schwarzschild solution. The third term is novel. It is small for most galaxies because it vanishes for circular motion. In addition, by the time the theta function turns on, indicating we are in the MOND regime, the baryon density will be falling off exponentially. It will make a contribution to the precision of the perihelion for the orbits of stars in disk galaxies, but these do not seem to be easy to measure. But it still may be the source of new effects that might be observable. More worryingly, it may have a nontrivial effect on the radial motions of stars in elliptical galaxies.

A. Hamiltonian analysis

To understand better the modification in dynamics implied by (47) we turn to a Hamiltonian analysis of particle motion in the thermal geometry. We proceed as usual [48] and identify two conserved quantities for a particle with trajectory $x^a(s)$ and four velocity $u^a = \frac{dx^a}{ds}$,

$$E = \tilde{g}_{ab} u^a k^b = f\dot{t}, \qquad L = \tilde{g}_{ab} u^a r^b = r^2 \dot{\phi}, \qquad (49)$$

where k^a and r^a are the time and angular Killing vector fields. Setting $-1 = \tilde{g}_{ab} u^a u^b$ yields the conservation law,

$$E^2 = \frac{Z}{2} \dot{r}^2 + V, \qquad (50)$$

where V is the standard Schwarzschild potential,

$$V = \frac{1}{2} \left[1 - \frac{2GM}{r} + \frac{L^2}{r^2} - \frac{2GML^2}{r^3} \right]. \qquad (51)$$

The variational principle (44) is then equivalent to a Lagrangian,

$$\mathcal{L} = \frac{Z}{2} \dot{r}^2 - V. \qquad (52)$$

We turn this into a Hamiltonian in the usual way,

$$\mathcal{H} = \frac{p^2}{2Z} + V, \qquad (53)$$

where the radial momentum is defined by

$$p = \frac{\delta \mathcal{L}}{\delta \dot{r}} = Z\dot{r}. \qquad (54)$$

We also have

$$\dot{p} = -\frac{\delta \mathcal{H}}{\delta r} = -\frac{\partial V}{\partial r} + \frac{p^2}{Zr} \theta[1-Z] \qquad (55)$$

from which we deduce the radial acceleration relation,

$$\ddot{r} = -\frac{1}{Z}\frac{\partial V}{\partial r} - \dot{r}^2 \frac{2\pi G\rho}{\partial_r \phi}\theta[1-Z], \qquad (56)$$

using $Z = \partial_r \phi$, this reproduces (47).

Outside the matter distribution, where $\phi = \frac{GM}{r}$, this can also be expressed as

$$\ddot{r} = -\frac{1}{Z}\frac{\partial V}{\partial r} - \frac{\dot{r}^2}{r}\theta[1-Z]. \qquad (57)$$

VI. DISCUSSION

We have arrived, very tentatively, at a kind of effective, bimetric description, so far valid only for static spacetimes. The standard metric, g_{ab}, is a solution to Einstein's equations, whose source is baryonic matter. However, the metric that governs the motion of particles is a different metric, $\tilde{g}_{ab}$, which differs from the standard metric only by a scaling of the radial component by a function of the temperature seen by a static observer.

This is justified by a rough argument based on the entropic gravity hypothesis of Verlinde and others, modified by the presence of a positive cosmological constant. It can also be justified by a scaling argument based on the scaling hypothesis of Milgrom. That hypothesis can in turn perhaps be justified by seeing the MOND regime as critical phenomena due to the proximity of a phase boundary.

It is painfully clear that the ideas discussed here are tentative and unlikely. Surely dark matter is a far simpler and less challenging hypothesis, and we will all breathe a sigh of relief if and when it is detected. Meanwhile, there are several challenges and opportunities which are raised by the present proposal:

  (i) The far field is a big issue. MOND cannot go on forever. For one thing we need to make contact with the large scale successes of general relativity. Z needs to level off at a constant at some distance, hopefully with a scaled mass. Until this is addressed the theory has nothing to say about the relevance of MOND scaling at extragalactic scales relevant for clusters and large scale structure, where the dark matter hypothesis appears to work well.

 (ii) Quantum gravity must ultimately explain the effective bimetric structure. One possibility is to employ disordered locality [49,50] as studied in [34].

(iii) In the above analysis we assumed disk galaxies are spherical, which they are not.

(iv) We assumed that there is an immediate transition between $Z(\frac{a}{a_0})$ between $z = 1$ for $a > a_0$ and $Z = \frac{a}{a_0}$ when $a < a_0$. In most of the MOND literature this transition is softened by an interpolating function as in (9).

 (v) Most interestingly, we have found a new prediction for dynamics in the MOND regime, given by the third term in (47). This does not affect circular motion, but it does introduce a damping of radial motion which may play a role in the dynamics of elliptical galaxies or in the formation of disk galaxies. We note that it is present only in the MOND regime, but it does not depend on the value of a_0.

(vi) What we have studied here is an effective theory, limited to test matter propagating on static spacetimes. One issue to be clarified is whether the biometric theory described here has instabilities when the restriction to static spacetimes is lifted. After that is clarified, the next step is to stationary spacetimes and should be straightforward. Beyond that we need to consider how to consistently couple matter to the gravitational field in the MOND regime. Equivalently, we want to consistently couple the two metrics to each other and to matter. If the picture presented here is in the right direction, that dynamics will be nonlocal and out of equilibrium. It is unlikely this can be expressed in a local or field theoretic form. Indeed, it is possible that the processes of equilibration on a cosmological scale reveal irreversible aspects of quantum gravity, as discussed in [51–58].

[1] V. C. Rubin, N. Thonnard, and W. K. Ford, Astrophys. J. **225,** L107 (1978); V. C. Rubin, W. K. Ford, and N. Thonnard, Astrophys. J. **238,** 471 (1980).

[2] R. B. Tully and J. R. Fisher, Astron. Astrophys. **54,** 661 (1977).

[3] S. McGaugh, F. Lelli, and J. Schombert, Astrophys. J. Lett. **816,** L14 (2016).

[4] M. Milgrom, Astrophys. J. **270,** 365 (1983).

[5] S. McGaugh, F. Lelli, and J. Schombert, Phys. Rev. Lett. **117,** 201101 (2016).

[6] F. Lelli, S. S. McGaugh, and J. M. Schombert, Astron. J. **152,** 157 (2016).

[7] M. Milgrom and R. H. Sanders, Astrophys. J. **678,** 131 (2008).

[8] S. S. McGaugh, Astrophys. J. **683,** 137 (2008).

[9] J. Magueijo (personal communication).

[10] X. Li, L. Tang, and H.-N. Lin, Chin. Phys. C **41,** 055101 (2017).

[11] For a review, see B. Famaey, and S. McGaugh, Living Rev. Relativity **15,** 10 (2012).

[12] S. S. McGaugh, Can. J. Phys. **93,** 250 (2015).

[13] R. H. Sanders, Astrophys. J. **473,** 117 (1996).

[14] E. M. Di Teodoro, F. Fraternali, and S. H. Miller, Astron. Astrophys. **594,** A77 (2016).

[15] S. H. Miller, R. S. Ellis, M. Sullivan, K. Bundy, A. B. Newman, and T. Treu, Astrophys. J. **753,** 74 (2012).

[16] J. Bekenstein and J. Magueijo, Phys. Rev. D **73,** 103513 (2006).

[17] J. D. Bekenstein, Phys. Rev. D **70,** 083509 (2004); **71,** 069901(E) (2005).

[18] J. W. Moffat, J. Cosmol. Astropart. Phys. 05 (2005) 003; 03 (2006) 004; M. Carmeli, Int. J. Theor. Phys. **37,** 2621 (1998).

[19] W. G. Unruh, Phys. Rev. D **14,** 870 (1976).

[20] For a review see, A. Strominger, arXiv:1703.05448.

[21] M. Milgrom, Phys. Lett. A **253,** 273 (1999); Ann. Phys. (N.Y.) **229,** 384 (1994).

[22] M. H. P. M. van Putten, arXiv:1411.2665.

[23] E. P. Verlinde, arXiv:1611.02269.

[24] S. Hossenfelder, arXiv:1703.01415.

[25] R. P. Woodard, Can. J. Phys. **93,** 242 (2015).

[26] L. Modesto and A. Randono, arXiv:1003.1998v1.

[27] C. M. Ho, D. Minic, and Y. J. Ng, Gen. Relativ. Gravit. **43,** 2567 (2011).

[28] S. H. Hendi and A. Sheykhi, Phys. Rev. D **83,** 084012 (2011).

[29] M. E. McCulloch, Europhys. Lett. **115,** 69001 (2016); Galaxies **2,** 81 (2014).

[30] F. R. Klinkhamer, Mod. Phys. Lett. A **26,** 2783 (2011).

[31] P. V. Pikhitsa, arXiv:1010.0318.

[32] C. M. Ho, Phys. Lett. B **693,** 567 (2010); D. Edmonds, Astrophys. J. **793,** 41 (2014); Int. J. Mod. Phys. A **32,** 1750108 (2017); J. Ng, arXiv:1602.00055.

[33] L. Smolin, Fundam. Theor. Phys. **187,** 427 (2017); contribution to Paddy@60, a book in honor of Thanu Padmanabhan.

[34] F. Markopoulou, I. Premont-Schwarz, and L. Smolin (unpublished).

[35] J. M. Maldacena, arXiv:hep-th/9711200; Phys. Rev. Lett. **80,** 4859 (1998); E. Witten, Adv. Theor. Math. Phys. **2,** 253 (1998); **2,** 505 (1998); S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Phys. Lett. B **428,** 105 (1998).For a review, see O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz, Phys. Rep. **323,** 183 (2000).

[36] G. Amelino-Camelia, L. Freidel, J. Kowalski-Glikman, and L. Smolin, Phys. Rev. D **84,** 084010 (2011).

[37] L. Freidel and L. Smolin, arXiv:1103.5626.

[38] H. Narnhofer, I. Peter, and W. Thirring, Int. J. Mod. Phys. B **10,** 1507 (1996).

[39] S. Deser and O. Levin, Classical Quantum Gravity **14,** L163 (1997).

[40] M. Milgrom, Fortschr. Phys. **65,** 1600046 (2017); Scholarpedia **9,** 31410 (2014).

[41] E. P. Verlinde, J. High Energy Phys. 04 (2011) 029.

[42] T. Jacobson, Phys. Rev. Lett. **75,** 1260 (1995).

[43] T. Jacobson, Phys. Rev. Lett. **116,** 201101 (2016).

[44] T. Padmanabhan, Rep. Prog. Phys. **73,** 046901 (2010); Classical Quantum Gravity **21,** 4485 (2004); Mod. Phys. Lett. A **25,** 1129 (2010).

[45] L. Smolin, arXiv:1001.3668.

[46] L. Smolin, Classical Quantum Gravity **3,** 347 (1986); in *Quantum Concepts in Space and Time*, edited by C. J. Isham and R. Penrose (Oxford University Press, Oxford, UK, 1985).

[47] S. Kolekar and T. Padmanabhan, Classical Quantum Gravity **32,** 202001 (2015).

[48] R. Wald, *General Relativity* (University of Chicago Press, Chicago, IL, 1984), pp. 139–142.

[49] F. Markopoulou and L. Smolin, Classical Quantum Gravity **24,** 3813 (2007).

[50] C. Prescod-Weinstein and L. Smolin, Phys. Rev. D **80,** 063505 (2009).

[51] R. Mangabeira Unger and L. Smolin, *The Singular Universe and the Reality of Time* (Cambridge University Press, Cambridge, UK, 2015).

[52] L. Smolin, *Time Reborn* (Houghton Mifflin Harcourt, New York, 2013).

[53] M. Cortês and L. Smolin, Phys. Rev. D **90,** 084007 (2014).

[54] M. Cortês and L. Smolin, Phys. Rev. D **90,** 044035 (2014).

[55] M. Cortês and L. Smolin, Phys. Rev. D **93,** 084039 (2016).

[56] M. Cortês and L. Smolin, arXiv:1703.09696.

[57] M. Cortes, H. Gomes, and L. Smolin, Phys. Rev. D **92,** 043502 (2015); M. Corts, A. R. Liddle, and L. Smolin, Phys. Rev. D **94,** 123514 (2016).

[58] L. Smolin, Classical Quantum Gravity **33,** 025011 (2016).