MODIFIED GRAVITY THEORIES: ALTERNATIVES TO THE MISSING MASS
AND MISSING ENERGY PROBLEMS

By

MARC EDWARD SOUSSA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2005

To my family

ACKNOWLEDGMENTS

The order in which I shall make my acknowledgements in no way reflects the relative importance each represents. There are those who directly and indirectly impacted my life as a graduate student here, and therefore their influences are unique in almost every respect.

Professionally, I thank my committee members (James Fry, David Groisser, Pierre Ramond, Pierre Sikivie, and Richard Woodard) for agreeing to serve and making the necessary time commitments required. Of those members, I make special acknowledgments to my advisor, Richard P. Woodard. His dedication, excellence, and unwavering commitment to his work and his students serve as an example to which any physicist could aspire. My life and my mind will forever be influenced by his tutelage, and for that I am eternally grateful.

I also make a special acknowledgment to Pierre Ramond for his constant and continuing support of my pursuits, and his committed effort to ensuring that my physics education received a well-rounded balance from many fields. His ability to find a unique perspective in which to view problems (related and especially unrelated to his own field of research) is something I will always admire and work to emulate. I thank James Fry and Pierre Sikivie for drafting letters of recommendation on my behalf and for being available to my questions and requests, whether administrative or physics-related.

Of the remaining faculty, I acknowledge Hai-Ping Cheng, my first advisor, and who encouraged me to pursue my interests and took interest in my progress throughout my graduate school career. I thank Mark Meisel who has always been

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODIFIED GRAVITY THEORIES: ALTERNATIVES TO THE MISSING MASS
AND MISSING ENERGY PROBLEMS

By

Marc Edward Soussa

May 2005

Chair: Richard P. Woodard
Major Department: Physics

Modified theories of gravity are examined and shown to be alternative possibilities to the standard paradigms of dark matter and dark energy in explaining the currently observed cosmological phenomenology. Special consideration is given to the relativistic extension of Modified Newtonian Dynamics (MOND) in supplanting the need for dark matter. A specific modification of the Einstein-Hilbert action (whereby an inverse power of the Ricci scalar is added) is shown to serve as an alternative to dark energy.

CHAPTER 1
INTRODUCTION

The advent of precision astrophysics in the past 20 years has provided cosmologists, particle theorists, and general relativists with a healthy volume of data and measurements with which to work and explain. It has become clear that a large fraction of the universe's energy content is unknown to us. Indeed, we are entering an exciting era of astrophysical investigation whereby experiments (past, present, and future) will guide physicists to understand the fundamental nature of the universe.

Our thesis considers two problems: namely, what I will term, with no attempts at originality, the missing mass and the missing energy problem. Each shall be approached by considering first the predominant or orthodox approach to their explanations. In the case of the missing mass problem this corresponds to the concept of dark matter. Respectively, the missing energy problem introduces a substance dubbed dark energy to describe the late-time acceleration of the universe. Each of these approaches will be shown to possess advantageous and seemingly "natural" features which seem to justify their acceptance as the leading candidates. However, these approaches are far from being satisfactory resolutions of their targeted problems. Inconsistencies and ambiguities remain. While that is so, alternative approaches must be vigorously researched.

One of the most important points to underline in this thesis is the fact that neither dark matter nor dark energy have been detected in a laboratory setting; they have only been observed in a gravitational context. Direct determination of dark matter would certainly quell (if not completely put to rest) the notion that perhaps we do not understand gravity even at the classical level. However, until

strong evidence of either particle dark matter or dark energy is obtained we may admit the possibility that it is new gravitational physics, not missing substances which is responsible for our observed universe.

Einstein's equation possesses two sides: the source side and the gravity side.

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.1}$$

where $T_{\mu\nu}$ is the matter-energy stress tensor and $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is the Einstein tensor. Each is obtained by varying the respective actions,

$$G_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\int d^4x\sqrt{-g}R \quad , \quad T_{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta}{\delta g^{\mu\nu}}\int d^4x\sqrt{-g}\mathcal{L}_M(\phi, A_\alpha, \Psi, \dots), \tag{1.2}$$

where $\mathcal{L}_M$ is a matter Lagrangian density. We will use a time-like signature throughout the thesis. $G_{\mu\nu}$ can be thought of as the source of gravity due to the matter components constituting $T_{\mu\nu}$. Clearly, we change the behavior of gravity by altering the sources present in $T_{\mu\nu}$. Indeed, the "dark horses" of modern physics, dark matter and dark energy, are both appendices to the source side, albeit in peculiar forms. Gravitational observation can only tell us about the gravity side, and thus the ability *always* exists to add terms to the source side to make the gravity side true. Only by solving the equations of motion of a new particle field and confirming the solution experimentally can one conclude real existence.

Consider the more general situation,

$$\mathcal{G}_{\mu\nu} = 8\pi G T_{\mu\nu} . \tag{1.3}$$

Here, $\mathcal{G}_{\mu\nu}$ is a not necessarily the Einstein tensor. We will refer to it as the "gravity tensor". Obviously, one must make restrictions when formulating the gravity tensor. General covariance reconciles the requirements of stability with local Lorentz invariance. Higher spin fields are notorious for possessing negative energy degrees of freedom. However, in a generally covariant theory these degrees of freedom are

either unphysical (e.g., the non-transverse modes of the photon in Lorentz gauge) or are constrained degrees of freedom (e.g., the Newtonian potential of gravity). In addition to conforming with general covariance, phenomenological requirements must be met. The most poignant of these is Newton's law of gravitation. Gravity is well understood at the solar system scale. Any proposed modification to Einstein's equation would have to faithfully reproduce Newton's law in this regime. Further, Einstein's relativity (which has proven to be extremely robust over a broad range of physical scale) must emerge from any candidate theory under the appropriate conditions. Gravitational lensing, the bending of light due to matter sources, will be shown to serve as an important phenomenological constraint. Astrophysical data coming from recent experiments such as the Wilkinson Microwave Anisotropy Probe (WMAP) and Cosmic Microwave Explorer (COBE); and data anticipated in the future from the Supernova Acceleration Probe (SNAP) and the Planck Satellite simultaneously furnish us with bizarre challenges now and place more and more stringent restrictions on models in the future.

The thesis is organized as follows: Chapter 2 discusses the missing mass problem and describes the dark matter approach to its explanation. In Chapter 3, the alternative proposition of MOND is introduced and shown as a viable alternative. First, we discuss the nonrelativistic successes of MOND, followed by a thorough analysis of the relativistic extensions currently under consideration. Specifically, the scalar-vector-tensor theories Bekenstein, Milgrom, and Sanders are surveyed in some detail followed by a complete treatment of the author's contribution to the purely metric approach. Chapter 4 introduces the dark energy problem discussing some of the more common approaches. Chapter 5 introduces a specific alteration to the Einstein-Hilbert action one can make to reproduce the same effect as dark energy. The author's contribution to the computation of the force of gravity due to such an alteration is completely treated. It is shown to place

severe limitations on this wide class of models. Chapter 6 summarizes the results with some remarks concerning the implications for present and future work.

CHAPTER 2
DARK MATTER: THE MISSING MASS

2.1   Introduction

When the velocities of satellites orbiting spiral galaxies are measured, they behave in a quite peculiar fashion. The velocities are of the order 10^2km/s (therefore $v/c \sim 10^{-3}$), and one would naively expect a Newtonian description of their gravitational dynamics. From elementary physics, we ascribe centripetal acceleration to a particle in a circular orbit outside a matter distribution $M(r)$,

$$a = \frac{v^2(r)}{r} = \frac{GM(r)}{r^2} \ ,$$

(2.1)

and thus well outside the mass distribution the asymptotic velocity behaves as,

$$v_\infty^2 = \frac{GM}{r} \ .$$

(2.2)

However, rather than a Keplerian fall off of the asymptotic velocity, satellites are observed to asymptote to a constant velocity far from the central galactic bulge,

$$v_\infty = \text{constant} \ .$$

(2.3)

Such behavior, one can easily imagine, can be described by inserting more matter into the system in the appropriate distribution. Indeed, this phenomenon has served as a major impetus for invoking the existence of an unknown matter component (dark matter) that pervades the galactic systems of our universe, and the universe's entirety itself. As we have not directly witnessed this matter via the electromagnetic, electroweak, or strong nuclear forces, it has acquired the "dark" nomenclature. Its only interaction that we have observed (if it were a true

component of our universe) is its gravitational interaction with luminous and nonluminous matter (and of course the cosmological fluid itself).

The application of this idea faces many obstacles from the onset. Immediately, its origins come into question. That is, one must use cosmological motivations and evidence to account for the existence of dark matter. Galaxy formation becomes a critical issue when discussing dark matter since early fluctuations in the CMB give evidence to the density fluctuations that were present at recombination. These imprints constrain the possibilities of how much matter energy density can be involved in galaxy formation.

Many candidates have been proposed, of which a few will be discussed later — but we may temporarily spotlight the necessity to quantify dark matter's standing in the particle description of matter. Specifically, does dark matter consist of the usual suspects (i.e. the Standard Model particles)? Or is dark matter the implementation of "new physics" — particles and fields not currently captured in the Standard Model? Its distribution and "symbiotic" relationship with the luminous matter in the galaxy and the universe must be identified. Further, with the recent augmentation of our experimental abilities in measuring galactic observables, dark matter must be embedded into our galactic systems in a self-consistent fashion such that the current observables are not voided by the introduction of a new exotic.

2.2   Dark Matter Taxonomy

We will limit the scope of our discussion of dark matter to issues concerning galactic systems. However, it should be noted that there is an enormous amount of theoretical and experimental work extant dealing with dark matter's possible role in cosmology. Dark matter, if existent, occupies far more of the intergalactic medium than the galactic one. This is evident from recent and quite constraining data from the WMAP probe [1], which reveals the large discrepancy between dark

and luminous matter critical fractions, $\Omega_{CDM} \sim 0.27$ and $\Omega_B \sim 0.03$, respectively (here CDM and B refer to cold dark matter and baryonic matter, respectively, explained below). Any cosmological component represents some fraction of the critical mass density (the density for which the universe expands to a critical radius and freezes) $\rho_c = 3H^2/8\pi G$,

$$\Omega_X \equiv \frac{\rho_X}{\rho_c} \, , \tag{2.4}$$

where H is the Hubble expansion parameter. The WMAP data confirms the hot big bang theory followed by a period of inflation giving rise to a flat universe. In terms of critical fractions, the total fraction is $\Omega_{Total} \sim 1$ with a dark energy component, $\Omega_\Lambda \sim 0.7$, which will be discussed later in greater detail. The question is then whether the 30% energy component (i.e. that which is not coming from dark energy) is a true matter component or whether there is new physics at the level of new particles and fields, or fundamental spacetime principles.

In addition, proof that dark matter really does exist must come from its observed interactions. Many dark matter candidates have emerged through the years. Roughly, one may divide them into two categories: nonluminous baryonic matter such as brown dwarfs, black holes, and large planets (MACHOS — Massively Compact Halo Structures); and weakly interacting particles such as neutrinos, axions, and neutralinos (WIMPS — Weakly Interacting Massive Particles) which pervade large portions of the universe. Of these two candidates, it has been experimentally and phenomenologically determined that if MACHOS do exist, they constitute very little of the total possible dark matter observed [2].

WIMPS can be further categorized by their cosmological history. Some particles formed in the big bang are relativistic for some period of time. Depending on their masses and couplings to other particles, we are able to predict and observe their transition from the relativistic regime to the nonrelativistic one. Those that are relativistic at the onset of galaxy formation are classified as hot

dark matter (HDM), whereas those which have dipped into the nonrelativistic regime are classified as cold dark matter (CDM). A light neutrino ($\lesssim 20\,\text{eV}$) and a heavy neutrino ($\sim 100\,\text{GeV}$) serve as candidates for CDM and HDM, respectively [3]. Finally, a third type of particle dark matter, the axion [4], arises from the Peccei-Quinn mechanism to solve the strong CP problem of QCD [5]. The axion is a particle which exhibits a particular symmetry that ensures (by reaching the minimum of its potential) that CP-symmetry is not violated in any strong nuclear interactions. Depending on the values of the axion mass and couplings, it is possible to account for a large fraction of the dark matter [6].

Further analysis using WMAP data, however, strongly tilts the favor toward a CDM scenario [7]. Because galactic formation depends upon the nature of its dark matter constituency [8], CDM has emerged as it is able (unlike HDM) to provide sufficient clumping on galactic scales we observe today. Therefore, we will survey the conservative approach to galactic dark matter and regard it as CDM.

Galactic dark matter has been most commonly introduced itself in the literature and in scientific investigations (both theoretical and phenomenological) as halos in which one may either view the galaxy embedded in the halo, or the halo embedded in the galaxy. The amount of dark matter projected is often on the order of 90% greater than that of ordinary matter. Since the rotation curves of the inner regions of galaxies are well reproduced by considering only luminous matter, it must be that the majority of the dark matter resides outside the central bulge to ensure the asymptotically constant satellite velocities.

We will further restrict ourselves to rotation curves of spiral galaxies, as they have been by far the most studied. The problem is to find how the dark matter distributes itself in and around the luminous matter of the galactic system. It cannot have a great impact on dynamics in the inner region since luminous matter accounts for the rotation curves there. It also faces the challenge of reproducing

the rotation curves outside the inner region for a wide range of physical scales —
distance, luminous mass, mass-to-light ratios, etc. Thus, dark matter profiles face
the further challenge of being universal (or at least exhibiting universality).

2.3   Dark Matter Distribution

A thorough review of the different halo models is given by [9]. We consider
only two here, which is more than sufficient to capture the most important features
of halos. The simplest CDM distribution which can successfully account for the
rotation curves of spiral galaxies is the isothermal sphere proposed by Gunn and
Gott [10] with mass density,

$$\rho(r) = \frac{\rho_0}{1 + (r/r_c)^2} \ . \tag{2.5}$$

Immediately from Equation 2.5 we see there are two parameters which must be
determined, the central density ρ_0 and the core radius r_c. Clearly, this profile
gives rise to the observed rotation curves. The circular velocity of a particle in the
isothermal distribution is,

$$v^2(r) = \frac{4\pi G}{r} \int_0^r dr' r'^2 \rho(r') \ , \tag{2.6}$$

$$= 4\pi G \rho_0 r_c^2 \left[1 - \frac{r_c}{r} \arctan\left(\frac{r}{r_c}\right) \right] \ . \tag{2.7}$$

In the limit of $r \gg r_c$, Equation 2.7 reduces to,

$$v_\infty \longrightarrow \sqrt{4\pi G \rho_0 r_c^2} \ , \tag{2.8}$$

the asymptotic velocity required to explain the rotation curves. A possibly un-
fortunate feature of Equation 2.8 is that the velocity never ceases to fall off, and
therefore certainly this scenario can only serve as a first approximation.

Another popular class of formulations is the so-called "universal" Navarro, Frenk, and White (NFW) profiles [11, 12],

$$\frac{\rho(r)}{\rho_c} = \frac{\delta_c}{\frac{r}{r_s}(1 + r/r_s)^2} \ , \tag{2.9}$$

where ρ_c is the same as in Equation 2.4. There are again two parameters to determine, as is always the case with a CDM profile: a density parameter (in this case a dimensionless characteristic number δ_c); and a length scale (here represented by the scale radius r_s). These profiles have inherited the classification "universal" for the similarities in the profiles between halos of widely varying mass, which came as a surprise, in light of the power-spectrum data [12].

Identical to the isothermal halo, we calculate the velocity from the NFW profile Equation 2.9,

$$v^2(r) = \frac{4\pi G\delta_c\rho_c r_s^3}{r}\left[\ln(1 + r/r_s) - \frac{1}{1 + r_s/r}\right] \tag{2.10}$$

The square velocity has the limiting behaviors,

$$v^2(r) \longrightarrow \begin{cases} 2\pi G\delta_c\rho_c r_s r & r \ll r_s \ , \\ \frac{4\pi G\delta_c\rho_c r_s^3}{r}\ln(r/r_s) & r \gg r_s \ . \end{cases} \tag{2.11}$$

These limiting behaviors are clearly superior in the phenomenological sense to the isothermal halo on account of the quasi-Keplerian fall off at large enough distances.

The nontrivial solution to the transcendental equation,

$$\bar{x}(1 + 2\bar{x}) - (1 + \bar{x})^2\ln(1 + \bar{x}) = 0 \ , \tag{2.12}$$

where $\bar{x} = \bar{r}/r_s$, gives the ratio of radii necessary to achieve the maximum velocity. This equation can be numerically solved to give $\bar{x} \simeq 2.16$. The velocity approximately drops to $0.82v_{\max}$ and $0.85v_{\max}$ for $x = 0.1\bar{x}$ and $x = 10\bar{x}$, respectively. Therefore, it is quite evident that flat rotation curves can be described with the

NFW profile, with the added feature that at asymptotically large distances, a fall off is exhibited.

2.4   Problems and Drawbacks to the CDM Halo Models

Although the isothermal halo model enjoys success in reproducing many of the observed rotation curves of disk galaxies, the assumptions leading to it exposes its lack of universality. The isothermal model assumes what has been called the "maximum-disk" hypothesis, which assigns to the disk the maximum mass-to-light ratio consistent with velocity measurements in the inner region [11]. The maximum-disk hypothesis effectively separates the disk from the halo. Therefore, satellites in the inner region inherit all but a negligible fraction of their velocities from the disk alone. This has the effect, of course, of limiting the core density of the halo. However, inclusion of data from dwarf galaxies shows this assumption to break down [13]. The rotation curves for these samples were of quite a different shape and could not be explained using the isothermal plus maximum-disk model. Further, velocity-dispersion measurements in the normal direction to the disk showed that the disk only contributed approximately 60% to the inner velocity of satellites [14], thereby nullifying the maximum-disk hypothesis.

The maximum-disk hypothesis attempts to regard rotation curves as functions of the luminosity alone. This was tendentiously proposed [14] in response to the observation that in low-luminosity galaxies, rotation curves rise slowly and continue rising past the optical radius; whereas in higher luminosity galaxies, the curves rise more sharply to their maximum, leveling off and sometimes even declining past the optical radius. However, it was discovered [14] that within the subset of observed galaxies that exhibit similar luminosities, differently shaped curves were measured (distinguished by the galaxy's surface brightness).

The NFW profiles for galactic halos and X-ray clusters have been studied extensively using N-body/gasdynamical simulations of CDM in a flat, low-density,

cosmological constant-dominated universe [15, 16]. Inserting nonsingular isothermal halos into N-body simulations shows these structures to poorly fit the data [15]. The NFW profiles by far constitute a more sophisticated approach and are able to overcome some of the problems of the isothermal model. That said, the NFW approach currently requires further dynamical input into determining the specific profiles that can account for the broad range of observed galaxy brightnesses and masses observed. For example, low surface-brightness galaxies are not as well fit by NFW profiles [17] where rotation curves rise more sharply in the inner region than the profile predicts.

Lastly, one of the more disturbing features of CDM halos is their large parameter range. Although it is by no means an indicator of futility of this approach, it certainly spurs the advent of more sophisticated models (if not radically new ideas). By trading off disk mass for halo concentration, one is always capable of producing similar velocity curves. In fact, *any* rotation curve can be fit by setting the disk's mass-to-light ratio to zero and tuning the halo parameters. Thus, differentiating among a class of halo models becomes a daunting task that can only be simplified by injecting new data or new fundamental concepts into the profile-building process.

This fine-tuning problem comes to light when considering the Tully-Fisher relation: the observation that a galaxy's luminosity is correlated to its peak rotation velocity via,

$$L \propto v_{\max}^4 \ . \tag{2.13}$$

The rotation velocity is mostly set by the halo, whereas the infrared luminosity comes from the visible matter in the galaxy [18]. The fine-tuning that arises from the symbiotic relationship between the disk and halo must somehow consistently reproduce Equation 2.13. The observational precision of Equation 2.13, however, is not to be easily expected from statistical processes involved in galaxy formation.

Therefore, the Tully-Fisher relation remains an issue for halo profiles that clearly needs to be addressed before any specific profile or halo mechanism is deemed satisfactory.

2.5   Concluding Remarks

Evidence for dark matter, regardless of the galaxy rotation curves, is quite extensive. The Standard Cosmology cannot do without it; at least not without a radical change to fundamental physics. Processes such as Big Bang Nucleosynthesis (BBN) [19, 20], structure formation [21, 22], and the cosmic microwave background (CMB) [23, 24] all indicate the Lambda CDM (cold dark matter with a cosmological constant) approach to be the superior scenario. Successes in these areas cannot be ignored. The conservative approach to favor dark matter as the leading candidate of the "missing mass" phenomenon. It can be said that halo models are just extending past their infancy. Indeed, any gasdynamical process incorporates an enormous amount of complexity. Attempting to find universality among galaxies is a daunting task if dynamical histories have distinct imprints in the rotation curves. Currently, NFW profiles offer the most universal approach in halo models. Despite their inability to accurately fit low-surface-brightness galaxies, their success encourages us to add new galactic dynamics to the model instead of abandoning it altogether. These profiles are inferred using N-body simulations, and therefore questions as to the validity of these simulations certainly enter. For example, one could argue that the limited number of particles employed prevents an accurate reproduction of the dynamics; and the singular nature of NFW profiles is certainly an unattractive feature that does not exist in nature.

The enormous parameter space also leads one to conclude that dark matter's existence cannot be *proved* by galaxy rotation curves alone. Even more narrowly, perhaps neither can any particular halo model (at least not any from the current arsenal). Along with cosmological evidence and particle searches for dark matter

properties, abundances, and composition, galaxy rotation curves will serve as one of the approaches to investigate its possibility. In the event of a "smoking gun" observation of dark matter's existence, rotation curves will play an important role in understanding galaxy formation and dynamical evolution.

CHAPTER 3
MODIFIED NEWTONIAN MECHANICS

MOND was proposed by Milgrom in 1983 [25] as an empirical alternative to dark matter in explaining the rotation curve phenomena. By altering gravity at low acceleration scales, one can reproduce the asymptotically constant velocities of satellites outside the central galactic bulge [26]. This chapter will first introduce the nonrelativistic formulation of Milgrom and Bekenstein. Next, relativistic extensions of this theory will be discussed: first the scalar-tensor varieties whose major proponents have been Bekenstein and Sanders, and secondly the purely metric approach of Soussa and Woodard. We will end with the phenomenological constraints and implications of each of these relativistic approaches.

3.1 Nonrelativistic Formulation

3.1.1 Motivation

One may formulate MOND by altering Newton's second law to be nonlinear in the acceleration $\vec{a}$,

$$\vec{F}_{\text{Newt}} = m\mu\left(\frac{a}{a_0}\right)\vec{a} \qquad \text{where} \qquad \mu(x) = \begin{cases} 1 & \forall\, x \gg 1 \, , \\ x & \forall\, x \ll 1 \, . \end{cases} \tag{3.1}$$

The function $\mu(x)$ is constructed to reproduce Newton's 2nd law, $F = ma$ for accelerations $a \ll a_0$ (corresponding to $\mu \gg 1$); and $F = ma^2/a_0$ for accelerations $a \gg a_0$ (corresponding to $\mu \ll 1$). The numerical value of a_0 has been determined by fitting to the rotation curves of nine well-measured galaxies [27],

$$a_0 = (1.20 \pm 0.27) \times 10^{-10} \text{m s}^{-2} \, . \tag{3.2}$$

However, when one considers the enormous internal accelerations of galactic and stellar constituents relative to the center-of-mass acceleration, it becomes preferable to regard MOND as a modification of the *gravitational* force at low accelerations [1] ,

$$\vec{F}_{\text{MOND}} = f\left(\frac{F_{\text{Newt}}}{ma_0}\right)\vec{F}_{\text{Newt}} \qquad \text{where} \qquad f(x) = \begin{cases} 1 & \forall x \gg 1 \ , \\ x^{-\frac{1}{2}} & \forall x \ll 1 \ . \end{cases} \qquad (3.3)$$

This empirical law is constructed to ensure the asymptotically constant velocities observed in the galactic rotation curves. That is, a particle orbiting a mass M at an acceleration $a \ll \frac{GM}{r^2}$ will follow a trajectory governed by the force $F_{\text{MOND}} = m\sqrt{a_0 GM}/r$. Setting $mv^2/r = F_{\text{MOND}}$ gives,

$$\frac{\sqrt{a_0 GM}}{r} = \frac{v_\infty^2}{r} \qquad \implies \qquad a_0 GM = v_\infty^4 \ . \qquad (3.4)$$

In the absence of dark matter, a galaxy's luminosity L should be a constant times its mass where the constant depends on the type of galaxy. Therefore, MOND is able to automatically reproduce the expectation $L \sim v_\infty^4$, which is the observed Tully-Fisher relation [29].

When the data are analyzed, MOND is shown to be an impressively robust predictive tool. Using only the measured distributions of gas and stars, and the fitted mass-to-luminosity ratios for gas and stars, MOND has accurately matched the data of more than 100 measured galaxies. A review by Sanders and McGaugh [30] summarizes the data and lists the primary sources. Two significant things should be noted: first, MOND agrees in detail, even with low-surface-brightness

---

[1] For example, if we were to consider neutral Hydrogen as a classical planetary system, the proton would be accelerating with a value of about 10^{19}ms^{-2} about the atomic barycentre and therefore well above MOND's characteristic acceleration scale. Milgrom has considered strongly nonlocal nonrelativistic particle actions in which the onset of MOND might be governed by the center-of-mass acceleration [28]

galaxies [17, 31]; second, the fitted mass-to-luminosity ratios are not unreasonable
[32].

On the other hand, when MOND is applied to intergalactic scales, or clusters
of galaxies, it has proven as yet to be less successful. Some dark matter must be
invoked to explain the temperature and density profiles at the cores of large galaxy
clusters [33], and data from the Sloan digital sky survey claims that satellites of
isolated galaxies violate MOND when care is taken to exclude interloper galaxies
(defined as dwarfs with large physical distances from the primary galaxies which
are claimed to make the halo mass profile difficult to measure) [34]. This objection,
however, has serious difficulty in being applied to all the rotation curves. Prada et
al. suggest that this is what leads to the systematic effects that fools the observer
into measuring a constant velocity dispersion in a "few" instances [34]. To regard
100 rotation curves — most of which have not been systematically checked for
the presence of this purported interloper phenomenon — as a few instances is
currently an overestimation of the interloper hypothesis. Recently, a paper by
Peñarrubia and Benson [35] analyzed the effects of dynamical evolution on the
distribution of dark halo substructures using semi-analytic methods (checked
with the latest N-body simulations). Their goal was to disentangle the effects of
processes acting on the substructures. They conclude that orbital properties of
substructure components are determined *a priori* by the intergalactic environment
[35] precluding the interloper hypothesis.

3.1.2 Action Principle

Milgrom and Bekenstein were able to obtain the MOND force law from a
nonrelativistic Lagrangian that respected all the symmetry (and hence conserva-
tion) principles we demand of nonrelativistic theories [36]. Considering a general
gravitational potential ϕ sourced by a mass density ρ, they proposed the following

Lagrangian,

$$L = -\int d^3r \left\{ \rho(\vec{r})\phi(\vec{r}) + (8\pi G)^{-1}a_0^2 F\left[\frac{(\vec{\nabla}\phi(\vec{r}))^2}{a_0^2}\right] \right\} , \qquad (3.5)$$

where the interpolating function F is related to the function μ of Equation 3.1 via,

$$\mu(x) = F'(x^2) . \qquad (3.6)$$

Assuming that the potential vanishes on the boundary, varying L with respect to ϕ gives the equation of motion,

$$\vec{\nabla} \cdot [f(\|\vec{\nabla}\phi\|/a_0)\vec{\nabla}\phi] = 4\pi G\rho, \qquad \text{where} \qquad f(x) = F'(x^2) . \qquad (3.7)$$

Consider an isolated mass M. Using Gauss's law and spherical symmetry we trivially have,

$$f(\|\vec{\nabla}\phi\|/a_0)\vec{\nabla}\phi = \frac{GM}{r^3}\vec{r} . \qquad (3.8)$$

in view of our empirical requirement set forth in Equation 3.3, we see that the asymptotic behavior of the acceleration of an object due to an object of mass M must behave as,

$$\vec{a} \longrightarrow -\frac{\sqrt{a_0 GM}}{r^2}\vec{r} + \mathcal{O}(r^{-2}) . \qquad (3.9)$$

Note the Equation 3.9 applies in all regimes: MOND and Newtonian. Requiring that,

$$\vec{\nabla}\phi = \frac{\sqrt{a_0 GM}}{r^2}\vec{r} , \qquad (3.10)$$

leads to the trivial solution,

$$\phi \longrightarrow \sqrt{a_0 GM} \ln(r/r_0) + \mathcal{O}(r^{-1}) , \qquad (3.11)$$

where r_0 is an arbitrary radius. Solving the field Equation 3.7 can be done by getting the first integral leaving one with an algebraic problem. From the rotational, space, and time translational symmetry of Equation 3.5 we immediately obtain

the conserved quantities of linear momentum, angular momentum, and energy, respectively.

The form of the interpolating function is obviously not unique and many different forms may be taken as long as the limiting behavior matches the requirement Equation 3.3. A typical example is,

$$f(x) = \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4}{x^2}}} \,, \tag{3.12}$$

and thus,

$$F(x) = \sqrt{x + x^2} - \ln(\sqrt{x} + \sqrt{1 + x}) \,. \tag{3.13}$$

3.2 Relativistic Formulation

3.2.1 Motivation

The purpose of the previous section, and that of the original authors, was to demonstrate: first, the MOND force law is derivable from an action principle; and second (and a direct consequence of the first), the action principle would possess the spacetime symmetries associated with conservation of energy and momentum.

The need to extend MOND into the relativistic domain is easily seen. Esthetically, the nonrelativistic formulation of MOND leaves a theorist almost aching to extend it to a fully relativistic description. However, and perhaps even fortunately, it is phenomenology which serves as the greatest impetus. As is, MOND proclaims itself as an *alternative* to dark matter, and therefore any phenomena to which the presence of dark matter has succeeded in explaining must now be equally or better described by MOND.

If one is interested in cosmology and gravitational lensing, there is no escape from the requirement of a generally covariant formulation of MOND that includes at least the usual metric. If MOND is indeed a viable alternative approach, it must

account for the deficiency observed in the general relativity with no dark matter prediction [37]. Milgrom, Bekenstein, and Sanders have a scalar-tensor approach to this end [36]

3.2.2 Scalar-tensor Approach

We will not in anyway here attempt to thoroughly consider scalar-tensor theories. Rather, we will take a more taxonomic approach, and in the process list their respective strengths and weaknesses from the theoretical perspective. Bekenstein [18] gives a more exhaustive review of these theories. The main phenomenological issue is whether the metric encodes the MOND force law or whether it is coming from a scalar field. All the scalar-tensor approaches possess the latter feature, and therefore one may immediately see that a new kind of dark matter emerges. Namely, if these added fields are real then we are again faced with the challenge of detecting them as any other dark matter candidate.

Aquadratic Lagrangian: AQUAL

Bekenstein and Milgrom first proposed a relativistic formulation of MOND as an appendix to their principal theme of devising a nonrelativistic potential theory in [26]. Their original approach was to introduce a dynamical degree of freedom in the form of a scalar field ψ in the spirit of scalar-tensor theories. Particles no longer follow geodesics of the Einstein metric $g_{\mu\nu}$, but rather that of a conformally related "physical" metric $\tilde{g}_{\mu\nu} = e^{2\psi}g_{\mu\nu}$. MOND physics comes from the scalar field Lagrangian density,

$$\mathcal{L}_\psi = -\frac{1}{8\pi GL^2}\tilde{f}(L^2 g^{\mu\nu}\psi_{,\mu}\psi_{,\nu}) \,, \tag{3.14}$$

where $\tilde{f}$ is an *a priori* known function which is constructed so as to reproduce MOND in the appropriate regimes and L is a constant length. In this theory, particles follow geodesics of $\tilde{g}_{\mu\nu}$. That is, if we parameterize a particle's worldline

to be $\chi^\mu(\tau)$, it has the following action,

$$S_m = -m \int d\tau e^\psi \sqrt{g_{\mu\nu}(\chi(\tau))\dot{\chi}^\mu(\tau)\dot{\chi}^\nu(\tau)} \,, \tag{3.15}$$

where a dot indicates differentiation with respect to τ. To make contact with the nonrelativistic theory, expand the particle action,

$$S_m = -m \int d\tau \left(1 + \Phi_N + \psi + \frac{\vec{v}^2}{2} + \dots \right). \tag{3.16}$$

identifying $\Phi_N = -(1 + g_{tt})/2$ as the Newtonian potential, determined by mass density ρ via the linearized Einstein equations. If we further restrict ourselves to the quasi-static case, it is straightforward to recover the MOND equation of motion Equation 3.7 in the weak field limit, with the acceleration of a particle governed by,

$$\vec{a} = -\vec{\nabla}(\Phi_N + \psi) \,. \tag{3.17}$$

Phase coupling gravitation: PCG

AQUAL was discovered to possess the debilitating feature that ψ could propagate superluminally [26]. To see this, consider the wave equation for free propagation of ψ that follows from Equation 3.14,

$$[\tilde{f}'(L^2 g^{\mu\nu}\psi_{,\mu}\psi_{,\nu})g^{\alpha\beta}\psi_{,\beta}]_{;\alpha} = 0 \,. \tag{3.18}$$

Here a semicolon represents covariant differentiation with respect to $g_{\mu\nu}$. Now linearize Equation 3.18 for small perturbations in ψ and consider the highest derivative terms. Following Bekenstein's coordinate prescription [26], it is possible to find a local Lorentz frame to point in the x-direction, allowing one to expand Equation 3.18 as,

$$0 = -\delta\psi_{,tt} + (1 + 2\xi)\delta\psi_{,xx} + \delta\psi_{,yy} + \delta\psi_{,zz} + \dots \,, \tag{3.19}$$

where $\xi \equiv d \ln \tilde{f}'(y)/d \ln y$ and the dots indicate terms with only one derivative. To determine whether ψ can propagate acausally, we need only consider the highest derivative term and their respective coefficients. Since $\xi \geq 0$, the coefficients in Equation 3.19 clearly display $\delta\psi$'s ability to violate causality.

Another downfall of AQUAL comes from the conformal nature in which the field couples to the Einstein and the physical metrics. It cannot influence gravitational lensing. We simply state this fact here: any conformal transformation of the form $g_{\mu\nu} \longrightarrow \Omega^{-2}g_{\mu\nu}$ has no impact on the bending of light. We will revisit this statement and discuss it at greater length when considering a purely metric theory. This means galaxies induce gravitational lensing only to the extent predicted by general relativity without dark matter. This is far too small [38].

In order to prevent the superluminal propagation of the scalar field inherent to the relativistic AQUAL theory, one can add a second scalar field which couples to ψ to ensure causality. This incarnation of MOND, PCG (Phase Coupled Gravity) [39], has now for a scalar Lagrangian density,

$$\mathcal{L}[\psi, A] = -\frac{1}{2}\left[g^{\mu\nu}(A_{,\mu}A_{,\nu} + \eta^{-2}A^2\psi_{,\mu}\psi_{,\nu}) + \mathcal{V}(A^2)\right] , \qquad (3.20)$$

and equation of motion for A,

$$A^{;\alpha}_{;\alpha} - \eta^{-2}A\psi_{,\alpha}\psi^{,\alpha} - A\mathcal{V}'(A^2) = 0 . \qquad (3.21)$$

Including a point mass M, the equation of motion for ψ follows from Equation 3.20,

$$(A^2g^{\alpha\beta}\psi_{,\beta})_{;\alpha} = \eta^2 e^{\psi}M\delta^3(\vec{r}) . \qquad (3.22)$$

Small values of $|\eta|$ justifies dropping the first term in Equation 3.20 and allows us to solve for A in terms of ψ. Inserting this into Equation 3.22 reproduces the same type of equation exhibited by AQUAL for ψ.

For the choice $\mathcal{V}(A^2) = -\frac{1}{3}\epsilon^{-2}A^6$ with ϵ a constant one can show that a particle acceleration for spherically symmetric solutions behaves as,

$$a_r = -\frac{GM}{r^2} - \frac{\eta^2 M}{4\pi\epsilon\kappa r} \ , \tag{3.23}$$

where $\kappa \equiv 2^{-3/2}\left(1 + \sqrt{1 + 4\left(\frac{\eta M}{\pi\epsilon}\right)^2}\right)$. Making an appropriate choice of ϵ and thus κ and identifying the critical acceleration scale in terms of our PCG parameters,

$$a_0 = \frac{\eta^3}{4\pi G\epsilon} \ , \tag{3.24}$$

will satisfy the rotation curve requirement. Thus PCG is capable of reducing to the AQUAL behavior and thus the nonrelativistic regime which is responsible for rotation curves. It also removes the acausal propagation of the scalar field. Naively, one would assume that since first derivatives of ψ in Equation 3.21 enter quadratically that causality ensues. A more thorough analysis by Bekenstein [40] shows that although this is not sufficient, considering only stable backgrounds enforces the desired property of causal propagation.

The parameters η and ϵ are stringently constrained by solar system tests. The accuracy to which we know the perihelion precession of Mercury proves enough to marginally rule out PCG (See [18] for a detailed discussion).

Finally, PCG suffers the same problem as AQUAL: the conformal coupling of the metrics leads to no enhancing of gravitational lensing in the general relativity with no dark matter hypothesis.

Disformally transformed metrics: Stratified gravitation

The failure of both the AQUAL and PCG theories stem from the conformal relation between the Einstein and physical metrics. Consider, rather a "disformal" relation [41],

$$\tilde{g}_{\mu\nu} = e^{-2\psi}(\mathcal{A}g_{\mu\nu} + \mathcal{B}L^2\psi_{,\mu}\psi_{,\nu}) \ , \tag{3.25}$$

where $\mathcal{A}$ and $\mathcal{B}$ are functions of the invariant $g^{\mu\nu}\psi_{,\mu}\psi_{,\nu}$ and L is a constant length. The second term in Equation 3.25 is responsible for the additional light deflection needed to explain observed galaxy lensing without dark matter. However, if one demands causality then it was found [42] that the sign of $\mathcal{B}$ would have to be such that the effect of the disformal transformation would be to decrease the amount of gravitational lensing. One way to overcome this shortcoming is to replace the second term in Equation 3.25 by a non-dynamical vector field which is purely time-like [43]. This stratified framework, which chooses

$$\tilde{g}_{\mu\nu} = e^{-2\psi} g_{\mu\nu} - 2U_\mu U_\nu \sinh(2\psi) \, , \tag{3.26}$$

$$g^{\mu\nu} U_\mu U_\nu = -1 \, , \tag{3.27}$$

is able to successfully describe observed gravitational lensing phenomena, and satisfies local solar system tests of gravity. However, it is clearly a preferred-frame theory, and has no *a priori* principle in which to select the preferred direction. Further, U_α must be of a bizarre form to satisfy Equation 3.27 at all spacetime points — undoubtedly an unnatural feature.

Tensor-Vector-Scalar Theory: TeVeS

Recently, Bekenstein has built on the successes of Sanders' stratified theory [18], by formulating a theory incorporating a vector, a scalar, and a tensor field which work cooperatively to exhibit the desired phenomenology — gravitational lensing beyond general relativity alone and causal propagation of scalar modes. In addition, this tensor, vector, scalar (TeVeS) theory is no longer a preferred-frame theory like the stratified framework theory of Sanders by virtue of the vector field's dynamics.

The TeVeS theory proposes the following transformation between the physical and Einstein metric,

$$\tilde{g}_{\mu\nu} = e^{-2\phi}g_{\mu\nu} - 2U_\mu U_\nu \sinh(2\phi) \ , \tag{3.28}$$

$$g^{\mu\nu}U_\mu U_\nu = -1 \ . \tag{3.29}$$

Now, however, the vector follows from the action [18],

$$S_V = -\frac{K}{32\pi G} \int d^4x \sqrt{-g} \left[g^{\alpha\beta}g^{\mu\nu}U_{[\alpha,\mu]}U_{[\beta,\nu]} - 4(\lambda(x)/K)(g^{\mu\nu}U_\mu U_\nu + 1) \right] , \tag{3.30}$$

where $\lambda(x)$ is a Lagrange multiplier enforcing Equation 3.29, and K is a constant of dimension zero. Further, we employ the notation $V_{[\alpha,\beta]} = V_{\alpha,\beta} - V_{\beta,\alpha}$. [2]

The scalar action is,

$$S_\sigma = -\frac{1}{2} \int d^4x \sqrt{-g} \left[\sigma^2(g^{\mu\nu} - U^\mu U^\nu)\phi_{,\mu}\phi_{,\nu} + \frac{G}{2\ell^2}\sigma^4 F(kG\sigma^2) \right] , \tag{3.31}$$

where F is a function again constructed to reproduce MOND behavior and k and ℓ are constants of length dimension zero and one, respectively. The gravity action is the usual Einstein-Hilbert action with metric $g_{\mu\nu}$, but the matter action is constructed coupling to $\tilde{g}_{\mu\nu}$ instead of $g_{\mu\nu}$.

TeVeS's scalar equation of motion is,

$$\left[\mu(k\ell^2 g^{\mu\nu}\phi_{,\mu}\phi_{,\nu}g^{\alpha\beta}\phi_{,\alpha} \right]_{;\beta} = kG_N(\tilde{\rho} + 3\tilde{p})e^{-2\phi} \ , \tag{3.32}$$

where $\mu(y)$ is defined by the equation,

$$-\mu F(\mu) - \frac{1}{2}\mu^2 F'(\mu) = y \ . \tag{3.33}$$

---

[2] It should be noted that Jacobson et al. [44] have considered models such as Equation 3.30 (Einstein-Aether Theories).

Quantities with tildes are constructed using the physical metric. Equation 3.32 is exact. To exhibit AQUAL behavior, we take $g^{\alpha\beta} \to \eta^{\alpha\beta}$ and $e^{-2\phi} \to 1$. Further, we neglect $\tilde{p}$ relative to $\tilde{\rho}$. Equation 3.32 can then be approximated,

$$\vec{\nabla} \cdot \left[\mu(k\ell^2(\vec{\nabla}\phi)^2\vec{\nabla})\phi \right] = kG\tilde{\rho} \ . \tag{3.34}$$

By properly constructing μ, Equation 3.34 reproduces the nonrelativistic scheme of AQUAL. Similarly, one can work out from Equation 3.32 the MOND limit and the GR limit — in fact, Equation 3.32 is the starting point for most analyses.

Parameterizing the metric as,

$$g_{\alpha\beta}dx^\alpha dx^\beta = -e^{\nu(\varrho)}dt^2 + e^{\zeta(\varrho)}\left[d\varrho^2 + \varrho^2(d\theta^2 + \sin^2\theta d\varphi^2) \right] \ . \tag{3.35}$$

One can show [18] that gravitational lensing in TeVeS is achieved by,

$$\Delta\varphi = \frac{b}{2} \int_{-\infty}^{\infty} \frac{\nu' - \zeta' + 4\phi'}{\varrho}dx \ , \tag{3.36}$$

where $x = \sqrt{\varrho^2 - b^2}$ is the Cartesian coordinate along the light-ray characterized by distance ρ and impact parameter b from the source. Further, Equation 3.36 can be approximated to leading order via the relation [18],

$$\Delta\varphi = 2b \int_{-\infty}^{\infty} \frac{\Phi'}{\varrho}dx \ , \tag{3.37}$$

where $\Phi = \phi + \Phi_N$. This result is consistent with the GR plus dark matter prediction.

TeVeS's major setback is that (like dark matter) it introduces new parameters to which one must then examine their origins. Taking appropriate limits of TeVeS's three parameters k, ℓ, K allows one to properly go from general relativity to MOND, etc. At least two (we may assume that at least one of these parameters are related explicitly to a_0 which is determined from rotation curve data) added experiments have to be performed to determine these parameters. We can imagine

that perhaps solar system tests provide one, lensing another, and rotation curves a third. Thus, like many current problems in current physics, we may recast an old problem in terms of new unknown parameters.

Another drawback is that if the scalar and vector fields are real entities, they are in essence a new form of dark matter which need explanation. One can thus argue that this is simply dark matter in a peculiar guise. Undoubtedly, however, the TeVeS is an improvement over the previous scalar-tensor approaches, and its phenomenological success alone makes it a viable and serious candidate explanation for the current astrophysical data.

3.2.3 Purely Metric Approach

Motivation

One interested in simplicity, viz. by pure degrees of freedom would certainly want to consider a purely metric extension of MOND, if for no other reason than theoretical completeness. Further, it is certainly arguable that a purely metric approach is closer to the "spirit" of general relativity. However, as is usually the case in physics, one gains simplicity in one facet of a theory only to lose it in another. We discovered that no local theory can reproduce MOND behavior. Consider the weak-field expansion of general relativity about a Minkowski background,

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} R \longrightarrow \frac{1}{16\pi G} \int d^4x \left\{ h_{\mu\nu}{}^{,\mu\nu} - h_{,\mu}{}^{\mu} + \mathcal{O}(h^2) \right\} , \qquad (3.38)$$

We want MOND corrections to "turn on" at a characteristic gravitational acceleration. The Ricci scalar, however, vanishes for general relativity outside a source. So there is no way a putative MOND correction term based upon R can "know" to turn on. You can get terms which don't vanish by using the Kretchman invariant, $R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$; these can indeed tell when to turn on but functions of this invariant automatically inherit the Ostrogradskian instability [45].

The only local actions which can avoid the higher derivative problem involve functions of the Ricci scalar alone. The Gauss-Bonnet invariant avoids this instability but is purely topological. Since the nonrelativistic MOND force law involves $\|\vec{\nabla}\phi\|^2$, the weak-field expansion must start at cubic order in the action.

Abandoning locality is certainly not an uncommon occurrence in today's theoretical physics. Effective theories have become more and more commonplace in regimes where ignorance of fundamental principles dominates. Quantum gravity's effective action, of course, falls in this category; and although we are unable to even say what the full effective action is, nothing prevents us from guessing its form. We chose the simplest class of guesses which would be capable of satisfying our nonrelativistic constraint Equation 3.7, acting with the inverse covariant d'Alembertian on the Ricci scalar[3] . We will refer to this as the *small potential*,

$$\varphi[g] = \frac{1}{\Box}R \qquad \text{where} \qquad \Box = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu) \,. \qquad (3.39)$$

(We use the convention $R_{\mu\nu} \equiv \Gamma^\rho_{\nu\mu,\rho} - \Gamma^\rho_{\rho\mu,\nu} + \Gamma^\rho_{\rho\sigma}\Gamma^\sigma_{\nu\mu} - \Gamma^\rho_{\nu\sigma}\Gamma^\sigma_{\rho\mu}.$) Embedding MOND in a nonlocal Lagrangian has the form,

$$\mathcal{L} = \frac{c^4}{16\pi G}\Big[R + c^{-4}a_0^2\mathcal{F}\Big(c^4a_0^{-2}g^{\mu\nu}\varphi_{,\mu}\varphi_{,\nu}\Big)\Big]\sqrt{-g} \,, \qquad (3.40)$$

where $\mathcal{F}(x)$ is an interpolating function whose form for small x controls the onset of MOND behavior as in the nonrelativistic case considered previously.

Although the field equations are nonlocal, they do not possess additional graviton solutions in weak-field perturbation theory. To see this, expand the metric

---

[3] This is not an unprecedented approach, but has been utilized in examining the physics of the post-inflationary universe [46].

about a Minkowski background,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} . \tag{3.41}$$

The Ricci scalar follows easily,

$$R = h^{\mu\nu}{}_{,\mu\nu} - \frac{1}{2}h_{,\nu} + \mathcal{O}(h^2) , \tag{3.42}$$

where graviton indices are raised and lowered using the Lorentz metric. Using our gauge freedom we choose de Donder gauge,

$$h^{\mu}{}_{\nu,\mu} - \frac{1}{2}h_{,\nu} = 0 , \tag{3.43}$$

to show that the small potential is local in the weak-field limit,

$$\varphi[\eta + h] = -\frac{1}{2}h + \mathcal{O}(h^2) . \tag{3.44}$$

Since the Lagrangian depends upon the first derivative of φ, this theory contains no higher derivative solutions in weak-field perturbation theory. Further, all solutions to the source-free Einstein equations are solutions to this theory since $R = 0$ throughout spacetime, which implies $\varphi = 0$ as well. Therefore, our modification to the Einstein-Hilbert action in Equation 3.40 is the change in response to sources, without adding new weak-field dynamical degrees of freedom — a clear distinction from the scalar-tensor theories so far discussed.

The class of nonlinear gravity theories we are considering are known to have a connection to scalar-tensor theories through a conformal rescaling of the metric, thereby introducing a minimally coupled, massive scalar field [47]. Regardless, the number of dynamical degrees of freedom remain equivalent. The scalar-tensor theories of Milgrom, Bekenstein, and Sanders all introduce *new* degrees of freedom, and therefore these approaches are truly distinct. Were the small potential $\varphi[g]$ an independent dynamical variable, rather than a functional of the metric, the purely

gravitational sector of these models would be identical, and thus the matter sector would serve to distinguish them. Matter couples in the usual way to $g_{\mu\nu}$ in our class of models whereas it couples to $\varphi^2 g_{\mu\nu}$ in the Bekenstein-Milgrom formalism. Note in particular that the field equation associated with a truly dynamical scalar,

$$(\varphi_{,\mu}\mathcal{F}')^{;\mu} = 0 \ , \tag{3.45}$$

is not generally solved by $\varphi = \Box^{-1}R$.

Phenomenological constraints

The two physical processes we will use to guide us in formulating our relativistic extension of MOND are the rotation curves and gravitational lensing. We will consider circular orbits as an approximation to the typical orbit a satellite of a spiral galaxy follows. The invariant length element in a static, spherically symmetric geometry can be expressed as,

$$ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu = -B(r)dt^2 + A(r)dr^2 + r^2 d\Omega^2 \ . \tag{3.46}$$

The worldline of a test particle moving in this geometry may be parameterized by $\chi^\mu(t)$ and obeys the geodesic equation,

$$\ddot{\chi}^\mu(t) + \Gamma^\mu{}_{\rho\sigma}(\chi(t))\dot{\chi}^\rho(t)\dot{\chi}^\sigma(t) = 0 \ . \tag{3.47}$$

It is a straightforward exercise to obtain the nonzero connection coefficients from Equation 3.46,

$$
\begin{aligned}
&\Gamma^t{}_{tr} = \frac{B'}{2B'}, &\qquad &\Gamma^r{}_{tt} = \frac{B'}{2A}, &\qquad &\Gamma^r{}_{rr} = \frac{A'}{2A}, \\
&\Gamma^r{}_{\theta\theta} = -\frac{r}{A}, &\qquad &\Gamma^r{}_{\phi\phi} = -\frac{r}{A}\sin^2\theta, &\qquad &\Gamma^\theta{}_{\theta r} = \frac{1}{r}, \\
&\Gamma^\phi{}_{\phi r} = \frac{1}{r}, &\qquad &\Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta, &\qquad &\Gamma^\phi{}_{\phi\theta} = \cot\theta \ .
\end{aligned}
\tag{3.48}
$$

Now parameterize the worldline for the case of circular motion,

$$(\chi^t, \chi^r, \chi^\theta, \chi^\phi) = \left(ct, r, \frac{\pi}{2}, \phi(t)\right) . \tag{3.49}$$

The only nontrivial geodesic equation in Equation 3.47 is,

$$\frac{B'}{2A} - \frac{r}{A}\frac{\dot{\phi}^2}{c^2} = 0 . \tag{3.50}$$

The $A(r)$ is thus irrelevant. In circular orbits, the velocity has the relation to the angular velocity $v = r\dot{\phi}$. In the MOND limit v^2 approaches the constant $v_\infty^2 = \sqrt{a_0 GM}$, and thus the MOND limiting form of $B(r)$ must obey,

$$B'(r) \longrightarrow \frac{2}{r}\sqrt{\frac{a_0 GM}{c^4}} . \tag{3.51}$$

For weak-fields we can write,

$$A(r) = 1 + a(r), \qquad B(r) = 1 + b(r), \tag{3.52}$$

where $|a(r)| \ll 1$ and $|b(r)| \ll 1$. A large spiral galaxy has on the order of 10^{11} solar masses in dust and gas, or $M \sim 10^{41}$kg. Therefore, such a galaxy would enter the MOND limit at a radius,

$$R_{\text{gal}} \sim \sqrt{\frac{GM}{a_0}} \sim 10^{20}\text{m} . \tag{3.53}$$

It is significant that the natural length scale corresponding to the MOND acceleration,

$$R_{\text{MOND}} \sim \frac{c^2}{a_0} \sim 10^{27}\text{m}, \tag{3.54}$$

is greater than the Hubble radius. This has the important consequence that,

$$\frac{r}{R_{\text{MOND}}} \ll 1, \tag{3.55}$$

on galaxy and galactic cluster scales, and therefore powers of r do not necessarily distinguish 'large' and 'small' terms.

We will now propose a phenomenological Ansatz for the asymptotic behavior of the weak-fields in terms of four order one parameters [4] ,

$$a(r) \longrightarrow \delta_1 \frac{GM}{c^2 r} + \epsilon_1 \sqrt{\frac{a_0 GM}{c^4}}, \qquad b(r) \longrightarrow \delta_2 \frac{GM}{c^2 r} + \epsilon_2 \sqrt{\frac{a_0 GM}{c^4}} \ln\left(\frac{r}{R_{\text{gal}}}\right) . \quad (3.57)$$

From Equation 3.51 we see that MOND predicts $\epsilon_2 = 2$. General relativity without dark matter predicts $\epsilon_1 = \epsilon_2 = 0$ and $\delta_1 = -\delta_2 = 2$. The insertion of an isothermal dark halo whose density is chosen to reproduce $v_\infty^2 = \sqrt{a_0 GM}$ would lead to the general relativity prediction of $\epsilon_1 = \epsilon_2 = \delta_1 = -\delta_2 = 2$.

To see what the remaining parameters in our Ansatz must be to be consistent with phenomenology, we consider the angular deflection of light from a mass M in terms of the turning point R_0,

$$\Delta\phi = 2 \int_{R_0}^\infty \frac{dr}{r} \sqrt{\frac{A(r)}{\left(\frac{r}{R_0}\right)^2 \frac{B(R_0)}{B(r)} - 1}} - \pi . \quad (3.58)$$

Expanding this expression in powers of the weak-fields and changing variables to $r = R_0 \csc\theta$ yields,

$$\Delta\phi = 2 \int_{R_0}^\infty \frac{dr}{r} \frac{1}{\sqrt{\left(\frac{r}{R_0}\right)^2 - 1}} \left\{ 1 + \frac{a(r)}{2} - \frac{1}{2} \frac{b(R_0) - b(r)}{1 - \left(\frac{R_0}{r}\right)^2} + \cdots \right\} - \pi , \quad (3.59)$$

$$= \int_0^{\frac{\pi}{2}} d\theta \left\{ a(R_0 \sec\theta) - \csc^2\theta [b(R_0) - b(R_0 \sec\theta)] + \cdots \right\} . \quad (3.60)$$

---

[4] One may be worried about the logarithmic growth in $b(r)$, but it is of no practical concern. The change in $b(r)$ from the onset of MOND all the way to the current horizon ($R_{\text{hor}} \sim 10^{26}$m) is

$$b(R_{\text{hor}}) - b(R_{\text{gal}}) \sim -\delta_2 \times 10^{-6} + \epsilon_2 \times 10^{-5} . \quad (3.56)$$

The weak-field regime is therefore applicable throughout the Hubble volume.

Substituting in the Ansatz Equation 3.57 gives,

$$\Delta\phi = (\delta_1 - \delta_2)\frac{GM}{c^2 R_0} + (\epsilon_1 + \epsilon_2)\frac{\pi}{2}\sqrt{\frac{a_0 GM}{c^4}} + \cdots . \qquad (3.61)$$

Without dark matter, general relativity gives too little deflection at large R_0 to be consistent with the frequency of lensing by galaxies. General relativity with an isothermal dark halo is consistent with the existing data [37, 38]. Therefore, for MOND to faithfully adhere to the current observations, it is required to have the sum $\epsilon_1 + \epsilon_2$ to be positive and of order one.

The field equations

In this section the field equations that would be derived from Equation 3.40 are presented. Here we give a more heuristic approach to this end. To avoid digression, we simply state here that the one does not get causal field equations by varying a temporally nonlocal action. Further, if one considers this class of models in the context of quantum field theory, then issues of nonreal operator eigenvalues quickly present themselves. We will here derive the field equations without light of the above concerns. Instead we will use a trick so to speak to obtain causal and conserved field equations from Equation 3.40. Therefore, one may as well regard the resulting equations of motion, rather than Equation 3.40 as defining the model.

The method we present reconciles the requirements of causality and conservation. Using retarded boundary conditions, one may easily add corrections to the field equations to enforce causality. However, it is immensely more difficult to guess symmetric tensors of any complexity that will combine to have a vanishing covariant divergence. Of course, field equations derived from any coordinate invariant action will automatically be conserved; however, varying actions which involve nonlocal operators result in equations at x^μ which depend upon fields in the future as well as in the past of x^μ.

The method is simplest to describe by comparing with the correct variation.
Consider an arbitrary functional of the metric, $f[g](y)$. We can write,

$$f[g](y)\frac{\delta\varphi[g](y)}{\delta g^{\mu\nu}(x)} = f[g](y)\left\{-\frac{1}{\Box_y}\frac{\delta\Box_y}{\delta g^{\mu\nu}(x)}\frac{1}{\Box_y}R(y) + \frac{1}{\Box_y}\frac{\delta R(y)}{\delta g^{\mu\nu}(x)}\right\}, \qquad (3.62)$$

where we used the fact that $\delta\Box^{-1} = -\Box^{-1}(\delta\Box)\Box^{-1}$.

Varying the covariant d'Alembertian and the Ricci scalar gives,

$$\frac{\delta\sqrt{-g(y)}\Box(y)}{\delta g^{\mu\nu}(x)} = \frac{\partial}{\partial y^\rho}\left[\sqrt{-g(y)}\left(\delta^\rho{}_{(\mu}\delta^\sigma{}_{\nu)} - \frac{1}{2}g^{\rho\sigma}g_{\mu\nu}\right)\delta^4(x-y)\right]\frac{\partial}{\partial y^\sigma}, \qquad (3.63)$$

$$\frac{\delta R(y)}{\delta g^{\mu\nu}(x)} = [R_{\mu\nu}(y) + D_\mu D_\nu - g_{\mu\nu}(y)\Box]\delta^4(x-y). \qquad (3.64)$$

We now define the small potential using the retarded Green's function,

$$\varphi[g](y) \equiv \int d^4z\, G_{\text{ret}}(y;z)R(z). \qquad (3.65)$$

Using the well known property $G_{\text{ret}}(x;y) = G_{\text{adv}}(y;x)$, we have,

$$\int d^4y\, f[g](y)\frac{\delta\varphi[g]}{\delta g^{\mu\nu}(x)} = -\frac{1}{2}g_{\mu\nu}R\frac{1}{\Box_{\text{adv}}}f + \left[\delta^\rho{}_{(\mu}\delta^\sigma{}_{\nu)} - \frac{1}{2}g^{\rho\sigma}\right]\varphi_{,\rho}\partial_\sigma\frac{1}{\Box_{\text{adv}}}f$$
$$+ [R_{\mu\nu} + D_\mu D_\nu - g_{\mu\nu}\Box]\frac{1}{\Box_{\text{adv}}}f. \quad (3.66)$$

The trick is to simply replace the advanced Green's function in Equation 3.66 with
the retarded ones,

$$\int d^4y\, f[g](y)\frac{\delta\varphi[g]}{\delta g^{\mu\nu}(x)} \longrightarrow -\frac{1}{2}g_{\mu\nu}R\frac{1}{\Box}f + \left[\delta^\rho{}_{(\mu}\delta^\sigma{}_{\nu)} - \frac{1}{2}g^{\rho\sigma}\right]\varphi_{,\rho}\partial_\sigma\frac{1}{\Box}f$$
$$+ [R_{\mu\nu} + D_\mu D_\nu - g_{\mu\nu}\Box]\frac{1}{\Box}f. \quad (3.67)$$

Since conservation depends only upon the differential equations obeyed by the
Green function, it is not affected by this replacement. The source for our gravi-
tational equations of motion is the stress-energy tensor from the variation of the
matter action S_m,

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}. \qquad (3.68)$$

Finally, taking $16\pi Gc^{-4}/\sqrt{-g}$ times the variation of our nonlocal action Equation 3.40 — in the sense of our trick Equation 3.67 — and defining the *large potential*[5] ,

$$\Phi[g] \equiv \frac{1}{\Box}(\varphi^{,\rho}\mathcal{F}')_{;\rho}, \tag{3.69}$$

gives the following field equations,

$$8\pi Gc^{-4}T_{\mu\nu} = 2[\Phi_{;\mu\nu} - g_{\mu\nu}\Box\Phi] + [1 - 2\Phi]G_{\mu\nu}$$
$$+ [g_{\mu\nu}\varphi^{,\rho}\Phi_{,\rho} - \varphi_{,\mu}\Phi_{,\nu} - \varphi_{,\nu}\Phi_{,\mu}] + \varphi_{,\mu}\varphi_{,\nu}\mathcal{F}' - \frac{a_0^2}{2c^4}g_{\mu\nu}\mathcal{F} . \tag{3.70}$$

It is worthwhile to explicitly demonstrate conservation as we have not rigorously derived Equation 3.70. Taking the covariant divergence D^ν of Equation 3.70 gives,

$$2[\Phi_{;\mu\nu} - g_{\mu\nu}\Box\Phi]^{;\nu} = 2R_\mu{}^\nu\Phi_\nu , \tag{3.71}$$

$$([1 - 2\Phi]G_{\mu\nu})^{;\nu} = -2R_\mu{}^\nu\Phi_\nu + R\Phi_\mu , \tag{3.72}$$

$$[g_{\mu\nu}\varphi^{,\rho}\Phi_{,\rho} - \varphi_{,\mu}\Phi_{,\nu} - \varphi_{,\nu}\Phi_{,\mu}]^{;\nu} = -\varphi_{,\mu}\Box\Phi - R\Phi_{,\mu} , \tag{3.73}$$

$$(\varphi_{,\mu}\varphi_{,\nu}\mathcal{F}')^{;\nu} = \varphi_{;\mu\nu}\varphi^{,\nu}\mathcal{F}' + \varphi_{,\mu}\Box\Phi , \tag{3.74}$$

$$\left(-\frac{a_0^2}{2c^4}g_{\mu\nu}\mathcal{F}\right)^{;\nu} = -\varphi_{;\mu\nu}\varphi^{,\nu}\mathcal{F}', \tag{3.75}$$

which obviously sum to zero. Because $1/\Box$ denotes the retarded Green's function, these equations are causal in the sense that the equations at x^μ depend only upon points within the x^μ's past light-cone[6] .

---

[5] Note that the large potential would vanish identically had φ been a fundamental scalar as in the scalar-tensor models of Bekenstein and Milgrom [36].

[6] This issue of causality should be distinguished from causal propagation. Field equations for which the highest derivative term is nonlinear can admit superluminal propagation as in the relativistic model of Bekenstein and Milgrom [36].

The Schwarzchild-MOND solution

Here we work out the small and potentials: Equation 3.39 and Equation 3.69, respectively, for a spherically symmetric and static metric Equation 3.46. They give rise to two independent equations of motion deriving from Equation 3.70 for this geometry.

The generally spherically symmetric and static geometry Equation 3.46, Equation 3.48 gives rise to the following Ricci scalar,

$$R = -\frac{B''}{AB} + \frac{B'}{2AB}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{2}{rA}\left(\frac{A'}{A} - \frac{B'}{B}\right) + \frac{2}{r^2}\left(1 - \frac{1}{A}\right) . \quad (3.76)$$

For this case, and acting upon a function only or r, the covariant d'Alembertian reduces to,

$$\Box = \frac{1}{r^2\sqrt{AB}}\frac{d}{dr}\left(r^2\sqrt{\frac{B}{A}}\frac{d}{dr}\right) . \quad (3.77)$$

The differential equation which defines the small potential therefore takes the form,

$$\left(r^2\sqrt{\frac{B}{A}}\varphi'\right)' = \left(r^2\sqrt{\frac{B}{A}}\left[-\frac{B'}{B} + \frac{2}{r}(A-1)\right]\right)' - r\sqrt{AB}\left(1 - \frac{1}{A}\right)\left(\frac{A'}{A} + \frac{B'}{B}\right) . \quad (3.78)$$

Assuming the parenthesized terms above vanish at $r = 0$, we can write,

$$\varphi'(r) = -\frac{B'}{B} + \frac{2}{r}(A-1) - \frac{1}{r^2}\sqrt{\frac{A}{B}}\int_0^r dr' r'\sqrt{\frac{B}{A}}(A-1)\left(\frac{A'}{A} + \frac{B'}{B}\right) . \quad (3.79)$$

The differential equation that defines the large potential is,

$$\partial_\mu\left(\sqrt{-g}g^{\mu\nu}\Phi_{,\nu}\right) = \partial_\mu\left(\sqrt{-g}g^{\mu\nu}\varphi_{,\nu}\mathcal{F}'\right) \implies \left(r^2\sqrt{\frac{B}{A}}\Phi'\right)' = \left(r^2\sqrt{\frac{B}{A}}\varphi'\mathcal{F}'\right)' . \quad (3.80)$$

Assuming again that the parenthesized terms vanish at $r = 0$ we can write,

$$\Phi'(r) = \varphi'(r)\mathcal{F}'\left(\frac{c^4\varphi'^2(r)}{a_0^2 A(r)}\right) . \quad (3.81)$$

The second covariant derivatives we shall need are,

$$\Phi_{;tt} = -\frac{B'}{2A}\Phi' , \; \Phi_{;rr} = \Phi'' - \frac{A'}{2A}\Phi' , \; \Box\Phi = \frac{1}{A}\left[\Phi'' + \frac{2}{r}\Phi' + \frac{1}{2}\left(\frac{B'}{B} - \frac{A'}{A}\right)\right] . \quad (3.82)$$

Only the diagonal components of the field equations Equation 3.70 are nontrivial in this geometry. The $\theta\theta$ and $\phi\phi$ components are proportional to one another, and are identically obtained from the rr and tt equations from conservation. We have therefore the two independent equations of motion,

$$\frac{8\pi GA}{c^4 B}T_{tt} = 2\Phi'' + \frac{4}{r}\Phi' + \frac{A}{B}G_{tt}(1 - 2\Phi) + \frac{a_0^2}{2c^4}A\mathcal{F} - \frac{A'}{A}\Phi' - \varphi'\Phi' \, , \qquad (3.83)$$

$$\frac{8\pi G}{c^4}T_{rr} = -\frac{4}{r}\Phi' + G_{rr}(1 - 2\Phi) - \frac{a_0^2}{2c^4}A\mathcal{F} - \frac{B'}{B}\Phi', \qquad (3.84)$$

and the two from conservation,

$$\frac{T_{\phi\phi}}{\sin^2(\theta)} = T_{\theta\theta} = \frac{r^3}{2A}\left\{\frac{B'}{2B}\frac{A}{B}T_{tt} + \left[\frac{d}{dr} + \frac{2}{r} - \frac{A'}{A} + \frac{B'}{2B}\right]T_{rr}\right\} . \qquad (3.85)$$

The tt and rr components of the Einstein tensor are,

$$\frac{A}{B}G_{tt} = \frac{A'}{rA} + \left(\frac{A-1}{r^2}\right) \qquad , \qquad G_{rr} = \frac{B'}{rB} - \left(\frac{A-1}{r^2}\right) . \qquad (3.86)$$

At this point we begin our perturbative analysis in this geometry, and recall that we may express $A(r)$ and $B(r)$ in terms of the weak-fields $a(r)$ and $b(r)$,

$$A(r) = 1 + a(r) \, , \qquad (3.87)$$

$$B(r) = 1 + b(r) \, . \qquad (3.88)$$

In the MOND regime, we have $|a(r)| \ll 1$ and $|b(r)| \ll 1$. Therefore, to leading order in the weak-fields Equation 3.79 becomes,

$$\varphi' \longrightarrow \frac{2a}{r} - b' + \cdots . \qquad (3.89)$$

Notice that the integrand in Equation 3.79 vanishes exactly in the general relativity regime when $A = B^{-1}$. In the MOND regime, the integrand is no longer zero, but it is second order in the weak fields and we are therefore justified in ignoring this term altogether for our present analysis.

In the asymptotic regime we can assume that each derivative adds a factor of $1/r$. Hence $\varphi'(r)$ goes like $1/r$ times the small numbers $a(r)$ or $b(r)$. It follows that Φ'/r is much larger in magnitude than $\varphi'\Phi'$. By similar reasoning we recognize that Φ'/r and Φ'' dominate the other MOND corrections,

$$\left|\frac{1}{r}\Phi'\right| \sim \left|\Phi''\right| \gg \left|\varphi'\Phi'\right| \, , \, \left|\frac{A'}{A}\Phi'\right| \, , \, \left|\frac{B'}{B}\Phi'\right| \, , \, \left|\frac{a_0^2}{c^4}\mathcal{F}\right| \, . \tag{3.90}$$

We will assume $T_{rr} = 0$ in the weak-field limit and allow for a nonzero $A/B T_{tt} = \rho$. Including the first two terms in Equation 3.90 with the general relativity terms allows us to express Equation 3.83 and Equation 3.84 to leading order in the weak-fields,

$$2\Phi'' + \frac{4}{r}\Phi' + \frac{a'}{r} + \frac{a}{r^2} + \ldots = \frac{8\pi G}{c^4}\rho(r) \, , \tag{3.91}$$

$$-\frac{4}{r}\Phi' + \frac{b'}{r} - \frac{a}{r^2} + \ldots = 0 \, . \tag{3.92}$$

The first of these equations Equation 3.91 can be integrated to give,

$$\frac{4}{r}\Phi' + \frac{2a}{r^2} + \cdots = \frac{K}{r^3} + \frac{16\pi G}{c^4 r^3} \int_{R_{\text{gal}}}^r dr' r'^2 \rho(r') \, , \tag{3.93}$$

where K is the constant of integration. Adding Equation 3.92 and Equation 3.93 cancels the leading MOND corrections,

$$\frac{b'}{r} + \frac{a}{r^2} + \cdots = \frac{K}{r^3} + \frac{16\pi G}{c^4 r^3} \int_{R_{\text{gal}}}^r dr' r'^2 \rho(r') \, . \tag{3.94}$$

Notice that Equation 3.94 is independent of the still unknown interpolating function $\mathcal{F}$. We can therefore make general statements about all models of the type Equation 3.40. In the absence of dark matter, the mass integral must eventually stop growing, for which case the left hand side of Equation 3.93 must fall as $1/r^3$. To satisfy this situation, $b'(r)$ must go as a constant times $1/r$ and $a(r)$ must go like minus this same constant. In terms of our Ansatz of the previous section

Equation 3.57, we have just demonstrated that,

$$\epsilon_1 + \epsilon_2 = 0 \ . \tag{3.95}$$

We acquire *no* lensing at leading order — a phenomenologically unacceptable result.

It is still worthwhile to see if we can find an interpolating function $\mathcal{F}(x)$ to reproduce MOND rotation curves. We will consider a sphere of mass M and radius R with very low, constant density,

$$\rho(r) = \frac{3Mc^2}{4\pi R^3}\theta(R - r) \ . \tag{3.96}$$

If the density is small enough the MOND regime prevails throughout, as in a low surface brightness galaxy. This means that Equation 3.91 can be integrated all the way down to $r = 0$ to give,

$$2\Phi' + \frac{a}{r} + \cdots = \frac{8\pi G}{c^4 r^2}\int_0^r dr' r'^2 \rho(r') \ . \tag{3.97}$$

We can also use Equation 3.89 to eliminate $b'(r)$ in $-r$ times Equation 3.92,

$$4\Phi' + \varphi' - \frac{a}{r} + \cdots = 0 \ . \tag{3.98}$$

Now eliminate $a(r)$ by adding Equation 3.97 and Equation 3.98, and then use Equation 3.81 to obtain an equation for the small potential,

$$\varphi'\left[1 + 6\mathcal{F}'\left(\frac{c^4\varphi'^2}{a_0^2}\right)\right] + \cdots = \frac{8\pi G}{c^4 r^2}\int_0^r dr' r'^2 \rho(r') \ . \tag{3.99}$$

For $r > R$ the mass integral is constant,

$$\varphi'\left[1 + 6\mathcal{F}'\left(\frac{c^4\varphi'^2}{a_0^2}\right)\right] + \cdots = \frac{2GM}{c^2 r^2} \qquad \forall r > R \ . \tag{3.100}$$

To get flat rotation curves we determined that MOND requires $\epsilon_2 = 2$. We have just computed explicitly that any model of the type Equation 3.40 must have

$\epsilon_1 = -\epsilon_2$ and therefore the weak-field limit Equation 3.89 for the small potential implies,

$$\varphi'(r) \longrightarrow -\frac{6}{r}\sqrt{\frac{a_0 GM}{c^4}} + \dots \tag{3.101}$$

It follows that the constant term within the square brackets of Equation 3.100 must exactly cancel, and that the next order term must involve one power of φ'. It is straightforward to compute our interpolating function,

$$\mathcal{F}'(x) = -\frac{1}{6} - \frac{\sqrt{x}}{108} + O(x) \implies \mathcal{F}(x) = -\frac{x}{6} - \frac{x^{\frac{3}{2}}}{162} + O(x^2) . \tag{3.102}$$

The associated weak-fields are,

$$a(r) \longrightarrow \frac{4GM}{3c^2 r} - 2\sqrt{\frac{a_0 GM}{c^4}} , \tag{3.103}$$

$$b(r) \longrightarrow -\frac{8GM}{3c^2 r} + 2\sqrt{\frac{a_0 GM}{c^4}} \ln\left(\frac{r}{R}\right) . \tag{3.104}$$

For the general weak-field Ansatz Equation 3.57 of section 2 we have just shown $-2\delta_1 = \delta_2 = -\frac{8}{3}$ and $-\epsilon_1 = \epsilon_2 = 2$.

We have a large amount of freedom in enforcing the MOND limit with regard to choosing the interpolating function $\mathcal{F}(x)$. The MOND limit is enforced by determining only the first two terms in the small x expansion of $\mathcal{F}(x)$. Therefore depending on the level of suppression desired its functional form is far from unique. Our only requirement is that the MOND corrections must be sufficiently small when entering the general relativity regime, i.e. $|x| \gg 1$. For example, we can make $\mathcal{F}(x) \longrightarrow -\frac{14}{3}|x|^{\frac{1}{2}}$ for large $|x|$ with the following extension,

$$\mathcal{F}'(x) = -\frac{\frac{7}{18}\mathrm{sgn}(x)}{1 + \frac{1}{6}|x|^{\frac{1}{2}}} + \frac{\frac{2}{9}\mathrm{sgn}(x)}{\left(1 + \frac{1}{6}|x|^{\frac{1}{2}}\right)^2} , \tag{3.105}$$

$$\mathcal{F}(x) = -\frac{\frac{22}{3}|x|^{\frac{1}{2}} + \frac{7}{9}|x|}{1 + \frac{1}{6}|x|^{\frac{1}{2}}} + 44\ln\left(1 + \frac{1}{6}|x|^{\frac{1}{2}}\right) . \tag{3.106}$$

For $|x| \gg 1$ this would typically suppress MOND corrections by some characteristic length of the system divided by $c^2/a_0 \sim 10^{27}$ m. If that is not sufficient one can always extend $\mathcal{F}(x)$ differently to obtain more suppression.

The FRW-MOND Solution

Even though we just discovered our relativistic formulation is unable to produce enough lensing, it is still instructive to see what it does for cosmology. Further, it serves as a potential gauge of how much a general formulation of MOND changes in passing from a static geometry to the time dependent one of cosmology. We begin with the usual Friedmann-Robertson-Walker metric for homogeneous and isotropic cosmologies,

$$ds^2 \equiv -c^2 dt^2 + a^2(t) d\vec{x} \cdot d\vec{x} \ . \tag{3.107}$$

In this geometry the Ricci scalar is,

$$c^2 R = 6\dot{H} + 12H^2 \qquad \text{where} \qquad H \equiv \frac{\dot{a}}{a} \ . \tag{3.108}$$

The small potential is defined by the equation,

$$\Box \varphi(t) = -a^{-3} \frac{d}{dct}\left(a^3 \frac{d\varphi}{dct} \right) = R(t) \ . \tag{3.109}$$

We define the initial values of φ and its first derivative to be zero, in which case the small potential becomes,

$$\varphi(t) = -\int_0^t dt' a^{-3}(t') \int_0^{t'} dt'' a^3(t'') \Big(6\dot{H}(t'') + 12H^2(t'') \Big) \ . \tag{3.110}$$

The large potential is defined by the differential equation,

$$\partial_\mu \Big(\sqrt{-g} g^{\mu\nu} \Phi_{,\nu} \Big) = \partial_\mu \Big(\sqrt{-g} g^{\mu\nu} \varphi_{,\nu} \mathcal{F}' \Big) \implies \frac{d}{dt}\Big(a^3 \dot{\Phi} \Big) = \frac{d}{dt}\Big(a^3 \dot{\varphi} \mathcal{F}' \Big) \ . \tag{3.111}$$

If we again assume null initial values the result is,

$$\Phi(t) = \int_0^t dt' \dot{\varphi}(t') \mathcal{F}'\Big(-c^2 a_0^{-2} \dot{\varphi}^2(t') \Big) \ . \tag{3.112}$$

The nonzero components of the second covariant derivative are,

$$\Phi_{;00} = c^{-2}\ddot{\Phi} \qquad , \qquad \Phi_{;ij} = -c^{-2}H\dot{\Phi}g_{ij} \ . \tag{3.113}$$

For a perfect fluid, the stress-energy tensor is,

$$T_{\mu\nu} = pg_{\mu\nu} + (p + \rho)u_\mu u_\nu \ . \tag{3.114}$$

Stress-energy conservation implies, $\dot{\rho} = -3H(\rho + p)$. The nonzero components of the Einstein tensor are,

$$c^2 G_{00} = 3H^2 \ , \tag{3.115}$$

$$c^2 G_{ij} = -(2\dot{H} + 3H^2)g_{ij} \ . \tag{3.116}$$

The two nontrivial equations of motion from Equation 3.70 in this geometry are,

$$8\pi G c^{-2}\rho = -6H\dot{\Phi} + 3H^2(1 - 2\Phi) + \frac{a_0^2}{2c^2}\mathcal{F} \ , \tag{3.117}$$

$$8\pi G c^{-2}p = 2\ddot{\Phi} + 4H\dot{\Phi} - (2\dot{H} + 3H^2)(1 - 2\Phi) - \dot{\varphi}\dot{\Phi} - \frac{a_0^2}{2c^2}\mathcal{F} \ . \tag{3.118}$$

Conservation tells us only one of these equations is independent.

In the MOND regime we can therefore express Equation 3.117 as,

$$-H\dot{\varphi} + 3H^2\left(1 - \frac{1}{3}\varphi\right) - \frac{1}{12}\dot{\varphi}^2 + \cdots = 8\pi G c^{-2}\rho \ . \tag{3.119}$$

For cosmology the argument $x = -(c\dot{\varphi}/a_0)^2$ is *negative* so the large potential has the same sign as the small potential,

$$\Phi(t) \longrightarrow \frac{1}{6}\varphi(t) \ . \tag{3.120}$$

In the MOND regime we can therefore express Equation 3.117 as,

$$-H\dot{\varphi} + 3H^2\left(1 - \frac{1}{3}\varphi\right) - \frac{1}{12}\dot{\varphi}^2 + \cdots = 8\pi G c^{-2}\rho \ . \tag{3.121}$$

Of special interest to cosmology is the case of a power-law scale factor,

$$a(t) = \left(1 + H_i t\right)^s .$$ (3.122)

Here H_i is $1/s$ times the Hubble parameter at $t = 0$. Substituting into Equation 5.13 gives the small potential,

$$\varphi(t) = -6s\left(\frac{2s - 1}{3s - 1}\right)\left\{\ln\left[1 + H_i t\right] - (1 - 3s)^{-1}\left[\left(1 + H_i t\right)^{1-3s} - 1\right]\right\} .$$ (3.123)

The logarithm term dominates in Equation 3.124 at late times. In this regime, we can express $\dot{\varphi}$ in terms of the Hubble parameter H,

$$\varphi(t) = -6s\left(\frac{2s - 1}{3s - 1}\right)\left\{\ln\left[1 + H_i t\right] - (1 - 3s)^{-1}\left[\left(1 + H_i t\right)^{1-3s} - 1\right]\right\} .$$ (3.124)

We can therefore write the MOND analog of the Friedman equation for power law expansion,

$$3\left\{1 + 2\sigma - \sigma^2 + 2s\sigma \ln\left[1 + H_i t\right]\right\}H^2(t) + \cdots = 8\pi G c^{-2}\rho(t) ,$$ (3.125)

where $\sigma \equiv (2s - 1)/(3s - 1)$. For $s > \frac{1}{2}$, the logarithm term serves to gradually slow the expansion — consistent with the MOND strengthening of the force of gravity in the weak-field regime.

For the case of radiation domination ($s = 1/2$ and $\sigma = 0$) we note that $\varphi(t) = 0$, and hence so too $\Phi(t) = 0$. The equations are therefore those of general relativity, but with the energy and pressure coming from ordinary matter. This is of course unacceptable in light of recent observations which show that nonbaryonic matter must predominate over baryonic matter by about a factor of six [1].

3.3   The MOND No-Go Statement for Purely Metric Approaches

In the previous section we discussed scalar-tensor theories of MOND. These models have had successes in complying with the phenomenological restrictions

currently available, but to date they have not shown to be completely satisfactory for the reasons already presented. We went on to develop systematically a relativistic version of MOND using a purely metric approach. The purely metric class of models suffered from far too little lensing. It is the purpose of this section to demonstrate that *any* purely metric theory of MOND will suffer the same phenomenological shortcoming.

3.3.1 Motivation

It was our intention in developing a phenomenologically viable theory of MOND which would satisfy the requirements of gravitational lensing and still reproduce the rotation curves. What we discovered, however, was that to leading order in the weak-fields in our class of models Equation 3.40, the MOND predicts *no* additional lensing to the prediction of general relativity. This is inconsistent without invoking the presence of dark matter. One is immediately tempted to consider different classes of models which can overcome the lensing "disaster". For example, one might replace the covariant d'Alembertian with the conformal one in defining a small potential,

$$\varphi_c[g] = \frac{1}{\Box_c} R \qquad \text{where} \Box_c = \Box - \frac{1}{6} R \ . \tag{3.126}$$

However, the distinction between $\Box$ and $\Box_c$ disappears in the weak-field regime since R scales as one power of the weak-fields times $1/r^2$. Therefore, this class of models would have no hope of having any success in acquiring a nonzero lensing contribution to leading order in the weak-fields.

Further, any MOND action which only contains the Ricci scalar as the source upon which the nonlocal operator acts will have no impact during the radiation phase of the universe — an entirely unacceptable feature. The next most

complicated scalar potential would seem to be,

$$\varphi_2[g] \equiv \frac{c^4}{a_0^2} \frac{1}{\Box}\left(R^{\mu\nu}R_{\mu\nu}\right) . \tag{3.127}$$

Because φ_2 has roughly two derivatives acting upon two powers of the weak-fields, one must also change the Lagrangian,

$$\mathcal{L}_2 = \frac{c^4}{16\pi G}\left[R + c^{-4}a_0^2\mathcal{F}_2\left(\varphi_2[g]\right)\right]\sqrt{-g} . \tag{3.128}$$

For this class of models $\mathcal{F}_2(x)$ would be linear in the MOND regime.

Instead of embarking on a program to discover a class of theories which are able to satisfy the lensing requirements, we propose to study the general features *any* purely metric formulation of MOND possesses. This approach is obviously the more powerful if a definitive result can be obtained (or at least if some firmer guidelines as to which classes of models can be considered in making MOND relativistic).

3.3.2   The Statement

We intend to show here that *no* phenomenologically viable, purely metric approach of MOND can be constructed within a set of given assumptions.

We assume that the gravitational force is mediated by the metric tensor $g_{\mu\nu}(x)$, and that its source is the usual stress-energy tensor $T_{\mu\nu}(x)$. In four space-time dimensions the metric is determined by the set of ten equations having the form,

$$\mathcal{G}_{\mu\nu}[g] = 8\pi G T_{\mu\nu} , \tag{3.129}$$

where the gravity tensor, $\mathcal{G}_{\mu\nu}$ is a functional of the metric (for ordinary general relativity it is simply the Einstein tensor $G_{\mu\nu}$). The stress-energy tensor is obtained as usual by varying the matter action.

We assume that the gravity tensor is covariant and is covariantly conserved,

$$g^{\rho\nu} D_\rho \mathcal{G}_{\mu\nu} = 0 \ . \tag{3.130}$$

At this point we have made no restrictions on the gravity tensor. In particular, we allow it to involve higher derivatives, and even to be a nonlocal functional of the metric.

Recall that MOND and Newtonian gravity are distinguishable only for very small accelerations. These accelerations are expressed as derivatives of the metric. It is well known that diffeomorphism invariance bestows the freedom to choose a coordinate frame in which the metric agrees with the Minkowski metric $\eta_{\mu\nu}$ at a single point. The observed fact that the gradients of the metric are small allow us to make the metric numerically quite close to $\eta_{\mu\nu}$ over a large region. Therefore, we are justified in expanding the gravity tensor in weak-field perturbation theory,

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) \ , \tag{3.131}$$

$$\mathcal{G}_{\mu\nu}[g] = \mathcal{G}_{\mu\nu}[\eta + h] \ . \tag{3.132}$$

The crucial observation now in specializing to MOND, is to notice that the MOND force law Equation 3.4 scales like the *square root* of the mass,

$$F_{\mathrm{MOND}} = \frac{\sqrt{a_0 G M}}{r} \ . \tag{3.133}$$

As a result, at least one component of $h_{\mu\nu}$ must scale like $\sqrt{GM}$ (for spherical distributions this would be the rr component but this does not matter). It is obvious that the right hand side of Equation 3.129 scales like GM, and therefore $\mathcal{G}_{\mu\nu}[\eta + h]$ must have at least one nonzero component whose lowest term is of order h^2.

If we further assume gravity to be absolutely stable then not all ten components of $\mathcal{G}_{\mu\nu}[\eta + h]$ can begin at quadratic order in the weak-field expansion. This

is due to the fact that the dynamical subset of field equations are obtained from varying the gravitational Hamiltonian. If its variation were quadratic then the Hamiltonian would be cubic, and this would be inconsistent with stability. The conclusion therefore is that only a *subset* of the ten components of $\mathcal{G}_{\mu\nu}[\eta + h]$ can begin at order h^2.

This subset must be distinguished in some covariant fashion. A symmetric, second rank tensor in four dimensions has two distinguished components: its covariant derivative and its trace. We can immediately see from Equation 3.130 that conservation occurs at all orders in perturbation theory, and therefore the covariant divergence cannot be responsible for the required h^2 term. We are left with the trace as the only remaining possibility,

$$g^{\mu\nu}\mathcal{G}_{\mu\nu}[\eta + h] = \mathcal{O}(h^2) \ . \tag{3.134}$$

Equation 3.134 implies asymptotic conformal invariance. Although we are able to reproduce the MOND rotation curves, the corrections to gravitational lensing of general relativity with no dark matter come in at quadratic order, and are therefore far too weak [37].

To see this, note that in the weak-field limit, one can perform a local, conformal rescaling of the metric,

$$g_{\mu\nu}(x) \longrightarrow \Omega^2(x)g_{\mu\nu}(x) \ , \tag{3.135}$$

and completely remove the corrections — the *linearized* MOND weak-fields are traceless. It has been known for some time that traceless metric field equations imply invariance under the conformal transformation in Equation 3.135 [48]. The full field equations are not traceless, and so neither is the full theory conformally invariant. This means that the linearized field equations only determine the metric up to a conformal factor (and a linearized diffeomorphism, but this is irrelevant

for the argument). The conformal part of the metric is fixed by the order h^2 term in the trace of the field equations, and this is how one component $h_{\mu\nu}$ contrives to scale like $\sqrt{GM}$.

This is a disaster for the phenomenology of gravitational lensing. Recall that for a general metric $g_{\mu\nu}$ the Lagrange density of electromagnetism is,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\rho\sigma}g^{\mu\rho}g^{\nu\sigma}\sqrt{-g} \ , \tag{3.136}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. Equation 3.136 is invariant under the metric rescaling Equation 3.135, and thus oblivious to MOND corrections to general relativity in the weak-field limit.

An Example

Here we will illustrate our no-go statement's using our nonlocal, purely metric model. The equations of motion Equation 3.70 imply the identification,

$$\mathcal{G}_{\mu\nu}[g] = 2[\Phi_{;\mu\nu} - g_{\mu\nu}\Box\Phi] + [1 - 2\Phi]G_{\mu\nu}$$
$$+ [g_{\mu\nu}\varphi^{,\rho}\Phi_{,\rho} - \varphi_{,\mu}\Phi_{,\nu} - \varphi_{,\nu}\Phi_{,\mu}] + \varphi_{,\mu}\varphi_{,\nu}\mathcal{F}' - \frac{a_0^2}{2c^4}g_{\mu\nu}\mathcal{F} \ . \tag{3.137}$$

We have already worked out the leading order terms in the small and large potentials, and the interpolating function as well. We may leave the large potential $\Phi[g] = \frac{1}{\Box}(\varphi^{,\rho}\mathcal{F}')_{;\rho}$ in terms of the small potential φ, in which case the necessary relations are,

$$\mathcal{F}(x) \longrightarrow -\frac{1}{6}x \ , \tag{3.138}$$

$$\Phi[g] \longrightarrow -\frac{1}{6}\varphi \ . \tag{3.139}$$

In taking the weak-field limit of $\mathcal{G}_{\mu\nu}$ we may neglect any products of $R_{\mu\nu}$, φ or Φ, such as $G_\mu\Phi$, $\varphi^{,\rho}\Phi_{,\rho}$, and $\varphi_{,\mu}\varphi_{,\nu}$. Henceforth, the weak-field limit of $\mathcal{G}_{\mu\nu}$ is

contained in the four terms,

$$\mathcal{G}_{\mu\nu} \longrightarrow \frac{1}{3}(g_{\mu\nu}\Box\varphi - \varphi_{;\mu\nu}) + R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R \ . \tag{3.140}$$

Of course these terms contain higher powers of $h_{\mu\nu}$, but more importantly they contain *all* the linear pieces. And the terms we have kept are exactly traceless,

$$g^{\mu\nu}\mathcal{G}_{\mu\nu} \longrightarrow g^{\mu\nu}\left(\frac{1}{3}g_{\mu\nu}\Box\varphi - \frac{1}{3}\varphi_{;\mu\nu} + R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) \ , \tag{3.141}$$

$$= \Box\varphi - R = 0 \ . \tag{3.142}$$

Note that tracelessness (and hence conformal invariance) is not a feature of the full field equations. In particular,

$$g^{\mu\nu}\mathcal{G}_{\mu\nu} = -6\Box\Phi - R[1 - 2\Phi] + 2\varphi^{,\mu}\Phi_{,\mu} + \varphi^{,\mu}\varphi_{,\mu}\mathcal{F}' - \frac{2a_0^2}{c^4}\mathcal{F} \ . \tag{3.143}$$

This scales like h^2 in the weak-field expansion.

3.3.3   A Connection with TeVeS

We have constructed MOND as a purely classical theory of gravity in the sense that no attempt at quantization has been made (that said, the class of models we derived can be thought of as originating from the effective action of gravity). It should be noted that Bekenstein's TeVeS model can be put in the form of a nonlocal, purely metric theory as long as the scalar and vector fields are not directly observed. This is done by integrating out those fields, leaving one with a nonlocal, purely metric action.

A second and extremely important distinction between the TeVeS theory and the purely metric theory is the coupling of gravity to matter. The TeVeS theory possesses a physical metric and an Einstein metric. One simple and direct consequence of this fact is that gravitational and electromagnetic radiation traveling from distant sources should possess disparate travel times. Recall that in the TeVeS theory, the physical ($\tilde{g}_{\mu\nu}$) and Einstein metric ($g_{\mu\nu}$) are related via the scalar field ϕ

and the vector field U_μ,

$$\tilde{g}_{\mu\nu} = e^{-2\phi}g_{\mu\nu} - 2U_\mu U_\nu \sinh(2\phi) \ . \tag{3.144}$$

Suppose we were to observe a "significant" astrophysical event (such as a Supernova) from a source such as the Large Magellanic Cloud (LMC)[7] . Assuming a Minkowski background (which is a reasonable first approximation) we would expect gravitational waves to take a time $T = D/c$ to reach us, while ultrarelativistic neutrinos would take a time,

$$\tilde{T} = \frac{1}{c} \int_0^D e^{-2\phi} dr \ , \tag{3.145}$$

$$\simeq T(1 - 2\phi) \ . \tag{3.146}$$

In obtaining Equation 3.146 we made two assumptions. First, that the vector of Equation 3.144 is directed in the direction of the cosmological evolution parameter, i.e. $U_\mu = \delta^0{}_\mu$. Secondly, we assume that the value of the scalar field is approximately constant and has a magnitude $\phi << 1$. This last assumption can be understood in far greater detail in [18]. However, a reasonable choice would be $\phi \sim 10^{-6}$. The distance to the LMC is $D \approx 10^5$lyrs; and therefore this value of ϕ would correspond to a delay of $\Delta T \approx 10$ min. Consequently, the parameter space of theories like TeVeS should ostensibly be constrained, and perhaps either be bolstered or falsified.

---

[7] For example, an observer at the Laser Interferometer Gravitational-Wave Observatory (LIGO) and an observer at SuperKamiokande (SuperK) would be capable of detecting the gravitational waves and neutrino flux from the event in the LMC.

3.3.4 Revisiting the No-Go Statement

This section examines the no-go statement in more detail. Specifically we consider circumventing the gravitational lensing disaster by relaxing the assumption of gravitational stability.

Let us revisit the assumptions which led to our no-go result:

1. The gravitational force is carried by the metric with its source being the usual stress-energy tensor.

2. Gravity is described by a covariant theory.

3. The MOND force law can be realized in weak-field perturbation theory.

4. The theory of gravity is absolutely stable.

5. Electromagnetism couples conformally to gravity.

The third and fifth assumptions are the most rigid. The third, if untrue, would inhibit us from working with any relativistic theory of MOND; if there is no region for which the MOND force is weak (or at least as weak as the Newtonian gravitational force), then there is no hope in passing even solar system experiments.

The first assumption may be violated if one makes a distinction between a "physical" and "gravitational" metric. In such a case test particles would follow geodesics of the former while gravity would behave according to dynamics of the latter. This is obviously a violation of the *strong* equivalence principle, but it is worth noting that to date there has been no conclusive data forbidding this possibility (See [36] for a detailed consideration). This old idea has been explored with many more modern theories such as Brans-Dicke [49], Dirac's variable gravitational constant [50], and string theories [51] to name a few.

The second assumption is easily foregone if one specifies a preferred-frame as we have seen in a previous section. However, by losing covariance or perturbability would certainly seem to be counter to the spirit in which we set upon in constructing a purely metric theory. Our fundamental prescription makes strong use of

the *strong* equivalence principle whereby gravity and matter are described by one metric, not two.

Relaxing the fourth assumption, that of gravitational stability, is seemingly the most reasonable choice to overcome the no-go statement. Given the choice between a stable and unstable theory when examining an observably stable system, the physicist will always choose the former. However, when phenomenologically driven, the latter may be the choice of greater utility. When the instability manifests itself at scales outside or nearly outside the physical scale, or at least in regimes where perturbative predictions no longer hold, the phenomenologist may cautiously accept (or at least consider accepting) the unstable solution as a candidate explanation. If the gravitational stability is on the super-cluster scale or larger, we may consider the possibility of all ten of the linearized MOND equations vanishing — the trace component is no longer distinguished.

How does this relaxation affect the no-go statement? Imagine that *all* of the linearized MOND weak-fields vanish in the field equations, in which case there would no longer be a linearized theory – a sufficient bending of light could be realized. We have already discussed the fact that if the MOND weak-fields begin at order h^2, then the gravitational Hamiltonian begins at cubic order. This signals an instability, but not necessarily a fatal one. There are *two* weak-field regimes – the weak-field (or Newtonian) and ultra-weak-field (or deep MOND). In regions such as the solar system it would be the Newtonian regime which dominates and thus we experience no deviation from well established physics. At larger scales (galactic and/or cosmological) we expect the deep MOND regime to enter the fold. At these scales, the unstable solution would proceed to decay into large wavelength particles diffusing as the universe expands. The result is that a return to the Newtonian regime could occur as decay products build a sufficiently large gravitational potential. The instability would, in essence, turn itself off just as it

becomes too large to become quantitatively reliable. We would no longer have a the tracelessness of the linearized equations — there would be no linearized theory, and hence gravitational lensing could be affected by MOND corrections.

CHAPTER 4
DARK ENERGY: THE MISSING ENERGY

4.1 Introduction

One of the greatest surprises to astrophysicists and particle physicists alike in the last 20 years is the recent observation from Type Ia Supernovae that the universe is entering a phase of acceleration. The Standard Cosmology is characterized by an early period of accelerated expansion (inflation) leading to a flat universe, a process supported by the large-scale isotropy observed in the Cosmic Microwave Background (CMB) [1]. The matter we see today is the result of gravitational collapse over 13.76 Gyrs after the initial singularity — which in turn is a manifestation of the density perturbations created by quantum fluctuations at the end of inflation. The sum of the critical energy fractions is very nearly one, with its current decomposition consisting of nearly 30% from matter (of which only approximately 4% is ordinary), and over 70% from an unknown source. The first year's data of WMAP [1] gives us (in terms of the Friedmann equation of cosmology in the present epoch) the critical fractions,

$$\Omega_{tot} = \Omega_k + \Omega_m + \Omega_r + \Omega_X = 1.02 \pm 0.02 \ , \tag{4.1}$$

$$\Omega_k = \Omega_r = 0 \ (95\% \ \text{CL}) \ , \tag{4.2}$$

$$\Omega_m = 0.27 \pm 0.04 \ , \tag{4.3}$$

$$\Omega_X = 0.73 \pm 0.04 \ . \tag{4.4}$$

Ω_X of course represents the source of critical energy responsible for the acceleration of the universe. This "dark" energy has commanded the attention of a wide variety of researchers, both theoretical and experimental.

The term dark energy, much like dark matter, is a rather broad encompassment of theoretical ideas – essentially referring to some added component to the right hand side of the Einstein equation which represents a "substance" which exerts a negative pressure and therefore induces expansion. Interestingly, Einstein's greatest blunder, the cosmological constant introduced in his research to ensure a static universe, has now almost impishly reintroduced itself into the theoretical arena.

4.2   The Many Faces of Dark Energy

Here we review the fundamental physics behind dark energy. Simple observational ideas — coupled with theory — allows one to easily understand the nature of dark energy. We will survey the modern landscape of theoretical ideas with a broad brush attempting to capture the more important themes and identify the common properties each must share.

On large scales the universe is homogeneous and isotropic, allowing one to use the Friedman-Robertson-Walker (FRW) metric to define the invariant length element in natural units,

$$ds^2 \equiv g_{\mu\nu}(x)dx^\mu dx^\nu = -dt^2 + a^2(t)d\vec{x} \cdot d\vec{x} \ . \tag{4.5}$$

Reading off the metric components, assuming the perfect-fluid form (with energy and pressure densities $\rho(t)$ and $p(t)$, respectively), and inserting them into the Einstein equation yields the two independent equations,

$$3\left(\frac{\dot{a}}{a}\right)^2 = 8\pi G\rho \ , \tag{4.6}$$

$$2\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 = -8\pi G p \ . \tag{4.7}$$

$$\tag{4.8}$$

Taking the linear combination of Equations 4.6 and 4.7 gives,

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\big[\rho(t) + 3p(t)\big], \tag{4.9}$$

then it is clear that in order for the universe to undergo accelerated expansion (i.e. $\ddot{a} > 0$), we must have,

$$p < 0, \qquad \text{and} \qquad |p| \sim |\rho| \; . \tag{4.10}$$

A useful but unfortunately named quantity is q, the deceleration parameter defined by,

$$q \equiv -\frac{\ddot{a}}{\dot{a}^2}a \; . \tag{4.11}$$

Clearly, $\dot{a}^2$ and a are positive definite, and therefore an accelerating universe demands $q < 0$.

Observationally, one can understand late-time expansion using a Hubble plot. The cosmological redshift, z, can be equivalently defined using the ratios of photon wavelengths or the ratios of scale factors at different times,

$$z \equiv \frac{\lambda_{now}}{\lambda_{then}} - 1 \quad , \quad z \equiv \frac{a_0}{a(t)} - 1 \; , \tag{4.12}$$

where a_0 is the value of the scale factor now and time $t = 0$ is the present and all values of $t > 0$ involve the past. Physical distances are determined via the relation,

$$d_{phys} = a(t)d_{co-moving} \; . \tag{4.13}$$

Supernovae have the desirable feature of having well-determined luminosities, and thus are good for distance and velocity measurements. The flux $\mathcal{F}$ one measures is related to the luminosity L thusly:

$$\mathcal{F} = \frac{L}{4\pi d_L^2} \; , \tag{4.14}$$

where d_L is the luminosity distance (physical distance) to the supernova. It is simple to show using Equation 4.12 and the Hubble parameter $H = \dot{a}/a$ that the

luminosity distance can be calculated with the integral,

$$d_L = (1 + z) \int_0^z \frac{dz'}{H(z')} \ . \tag{4.15}$$

Substituting into Equation 4.15 the expansion of $H(z)$ for small z gives,

$$d_L = \frac{(1+z)}{H_0} \int_0^z dz' \left(1 - \frac{1}{H_0} H_0' z' + \dots \right) \ . \tag{4.16}$$

Using Equation 4.11 and the chain rule allows us to make the identification $H_0' = (1 + q)H_0$. Integrating Equation 4.16 term by term and collecting powers results in the following power series in z,

$$d_L = \frac{z}{H_0} \left[1 + \frac{1}{2}(1 - q_0)z + \dots \right] \ . \tag{4.17}$$

The first term in Equation 4.17 represents Hubble's law — namely, $v = H_0 d$. The second term is the first deviation of Hubble's law. Therefore, by measuring d_L and z and plotting them one infers the curvature (or deviation from linearity). For $z \gtrsim 1$, this expansion breaks down; in which case numerical integration can be performed using the energy density one assumes to be present in Equation 4.6 (this obviously introduces some model-dependent effects).

To determine the evolution of the missing energy component, we define the parameter w which relates the energy and pressure densities at any given time by the equation of state,

$$p = w\rho \tag{4.18}$$

Equations 4.9 and 4.18, along with the requirement that the universe accelerate forces the inequality,

$$\rho(1 + 3w) < 0 \ . \tag{4.19}$$

Since $\rho \geq 0$, dark energy must give rise to $w < -\frac{1}{3}$. Recently, w has become slightly more constrained as measurements have improved. In order for the structure formation we currently observe to exist from the density perturbations indicated

by CMB anisotropy measurements, we must have $w < -\frac{1}{2}$ [52]. Additionally, the absence of any intragalactic physics due to dark energy leads one to believe that its distribution be smooth and homogeneous on large-scales.

The Cosmological Constant

The history of the cosmological constant is now so well known it needs little development. Einstein introduced a constant to his general relativity equations to balance the collapsing effect that matter alone would exert on the cosmic fluid. By doing this it imposed what he felt at the time to be the natural state of the universe — static.

Of course, the observed redshift of distant galaxies quickly did away with the notion of a static universe; however, the cosmological constant would undergo a conceptual "revolution" soon after, when particle theorists were forced to incorporate the quantum fluctuations of the vacuum which persist in gravity even after renormalization. For example, consider the Hamiltonian of the quantum harmonic oscillator with N degrees of freedom in terms of the raising and lowering operators $a^\dagger$ and a, respectively:

$$H = \hbar\omega \sum_{i=1}^{N}\left[a_i^\dagger a_i + \frac{1}{2}\right] = \hbar\omega \sum_{i=1}^{N} a_i^\dagger a_i + \frac{N\hbar\omega}{2} \ . \tag{4.20}$$

The transition to field theory takes the number of degrees of freedom to infinity,

$$H = \sum_{\vec{k}}\left[a^\dagger(\vec{k})a(\vec{k}) + \frac{1}{2}\right]\hbar\omega(\vec{k}) \ . \tag{4.21}$$

Clearly, the ground state contributes an infinity to Equation 4.21. The usual practice is to redefine the Hamiltonian by shifting the energy by an infinite amount as only energy differences are observable quantities. This procedure, however, cannot be employed with gravity. Theories like QED, QCD, and the EW force all possess dimensionless expansion parameters. Thus, one may always find enough counter terms in the renormalization scheme at all energy scales.

The expansion parameter of gravity is Newton's constant G_N, which in natural units has dimensions M^{-2}. Thus, as one increases in energy (i.e. probing the ultraviolet) it takes more and more counter terms to renormalize to a finite value — an infinite such counter terms for higher order terms and thus gravity in this sense is nonperturbatively finite. Admitting our ignorance we may insert an *ad hoc* cutoff,

$$H_{\text{VAC}} \sim \sum_{\vec{k}}^{\Lambda} \hbar\omega(\vec{k}) = \Lambda \ . \tag{4.22}$$

The cutoff scale is often chosen to be the Planck mass, $\Lambda \sim M_P \sim 10^{19}\text{GeV}$, at which point new physics is needed to make predictions as to how gravity and spacetime behave. Further, the Casimir effect, which in QED is the force registered by two neutral, conducting plates as a result of quantum vacuum fluctuations lends credence that the vacuum plays a definite role at certain scales [53].

Vacuum energy is naturally homogeneous, isotropic, and of course must enter Einstein's equation covariantly,

$$T_{\mu\nu}^{\text{VAC}} = \rho_{\text{VAC}} g_{\mu\nu} \sim M_P^4 g_{\mu\nu} \ , \tag{4.23}$$

Vacuum energy has the inherent properties that ρ_{VAC} is uniform throughout spacetime and that $p_{\text{VAC}} = -\rho_{\text{VAC}}$ (i.e. $w = -1$).

The proposition of a cutoff introduces an awkward problem which we must face. Currently, the value of the constant can be grossly estimated by assuming,

$$\rho \sim \rho_\Lambda \sim \frac{3H_0^2}{8\pi G} \sim 10^{-48}\text{GeV}^4 \ . \tag{4.24}$$

If we take a cutoff seriously, then a bare cosmological constant would have a value,

$$\rho_{\Lambda \ bare} \sim \Lambda^4 \sim 10^{76}\text{GeV}^4 \ . \tag{4.25}$$

Thus, we are forced to account for a discrepancy of 120 orders of magnitude between the expected and the observed. One may do away with many orders of

magnitude if supersymmetry is included (with a cutoff scale $M_SUSY \sim$ TeV), or if the cutoff is not the Planck scale but rather the electroweak scale of 100 GeV; however, it does not do nearly enough and we are left with essentially the same questions, if but perhaps in a slightly less embarrassing form.

The observation of a small but non-zero cosmological constant which leads to the so-called coincidence problem: namely, why has it only recently achieved relative dominance [54, 55]?

There have been many attempts at understanding these critical problems [56, 57, 58, 59, 60, 61, 62], none of which can be deemed satisfactory solutions, else we would certainly have something far more profound to say about dark energy. Introducing a homogeneous scalar field which possesses dynamics will work [63, 64, 65, 66], but one must understand why it is homogeneous [67] and again why it has achieved dominance now. This approach, named quintessence, works as a tracker solution, whereby the energy density of the scalar field follows the energy density of the universe in such a way as to produce late-time acceleration.

Long-range forces have been suggested [68] whereby one introduces a charged scalar field with a long-range, self-interacting force mediated by vector gauge boson. If the gauge boson mass were to vanish at the minimum of the scalar potential, the field would be unable to relax to its minimum, and cosmic acceleration could be achieved [68]. Unlike quintessence, this model predicts an oscillating equation of state [68] which can ostensibly be observed by high-z Supernovae; and therefore, this model is distinguishable.

CHAPTER 5
LATE TIME ACCELERATION WITH A MODIFIED EINSTEIN-HILBERT ACTION

As was the case with dark matter, dark energy plays the role of an added and hitherto unknown component to the right hand side of the Einstein equation. Endowing it with the special property that it exerts a negative pressure on the cosmological fluid provides us with a somewhat natural mechanism with which to explain late-time acceleration. And just as MOND announces itself as an alternative to the dark matter hypothesis, modified Einstein-Hilbert gravities position themselves as alternative candidate explanations.

This chapter illustrates how adding a term proportional to an inverse power of the Ricci scalar gives rise to an accelerating universe in late-time cosmology, i.e. post big-bang inflation. We then examine the effect an added inverse Ricci term in the action has on the resulting force of gravity.

5.1   Late-time Acceleration

Carroll, Duvvuri, Trodden, and Turner proposed a purely gravitational approach [69]. Late time acceleration is achieved by considering a subset of nonlinear gravity theories in which a function of the Ricci scalar is added to the usual Einstein-Hilbert action,

$$S_g[g] = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[R + f(R)\right] , \tag{5.1}$$

where,

$$f(R) = -\mu^{2(p+1)} R^{-p} \qquad \forall \qquad p > 0 . \tag{5.2}$$

From dimensionality we see that μ is an *a priori* unknown parameter of mass dimension one. Some connections to braneworlds have been proposed in which terms with inverse powers of the Ricci scalar are exhibited [70].

For simplicity we will consider the case $p = 1$ at no loss of qualitative understanding. We also include the matter action for completeness, in which case the action is,

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left(R - \frac{\mu^4}{R} \right) + \int d^4x \, \mathcal{L}_M \ . \tag{5.3}$$

The equations of motion follow directly from Equation 5.28 via the variation,

$$8\pi G T^M_{\mu\nu} = \frac{-2}{\sqrt{-g}} \frac{\delta S_g}{\delta g^{\mu\nu}} \ , \tag{5.4}$$

$$= \left(1 + \frac{\mu^4}{R} \right) R_{\mu\nu} - \frac{1}{2} \left(1 - \frac{\mu^4}{R^2} \right) R g_{\mu\nu} + (g_{\mu\nu} \Box - D_\mu D_\nu) \frac{\mu^4}{R^2} \ , \tag{5.5}$$

where $T^M_{\mu\nu}$ is the matter energy-momentum tensor. It is quite evident that the limit $\mu \to 0$ in Equation 5.29 takes us back to the usual Einstein equation of motion. Equation 5.29 can be trivially solved for R if one is considering the constant-curvature vacuum solution (i.e. $D_\mu R = 0$ and $T^M_{\mu\nu} = 0$). Interestingly, they are non-zero,

$$R_{\text{vac}} = \pm\sqrt{3}\mu^2 \ , \tag{5.6}$$

unlike their Minkowski counterpart. Of course, a (negative) positive constant-curvature solution *is* precisely (anti) de Sitter space, and we therefore see immediately how our equation of motion Equation 5.29 is capable of providing a purely gravitational mechanism for explaining cosmological acceleration.

We wish to consider cosmological scenarios. Thus, on the grounds of large-scale isotropy and homogeneity of the cosmological fluid, we restrict ourselves to the perfect fluid form of the energy-momentum tensor,

$$T^M_{\mu\nu} = (\rho_M + P_M)U_\mu U_\nu + P_M g_{\mu\nu} \ , \tag{5.7}$$

where U^α is the fluid rest-frame four velocity, ρ_M is the energy density of matter and radiation, P_M is the pressure of matter and radiation which is related to the energy density via the equation of state $P_M = w\rho_M$. In a matter dominated universe, $w = 0$; and in a radiation dominated universe, $w = 1/3$.

Homogeneity and isotropy allows also to limit our analysis to metrics of the Robertson-Walker form,

$$ds^2 = -g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + a^2(t)d\vec{x} \cdot d\vec{x} \, . \tag{5.8}$$

It is straightforward to compute the Ricci scalar in terms of the scale factor $a(t)$ from Equation 5.8,

$$R = 6\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2}\right) = 6(\dot{H} + 2H^2) \, , \tag{5.9}$$

where H is the Hubble parameter,

$$H \equiv \frac{\dot{a}}{a} \, . \tag{5.10}$$

With Equation 5.9 and Equation 5.8, we obtain the two time-time and space-space equations of motion from Equation 5.29,

$$3H^2 - \frac{\mu^4}{12(\dot{H} + 2H^2)^3}(2H\ddot{H} + 15H^2\dot{H} + 2\dot{H}^2 + 6H^4) = 8\pi G\rho_M \, , \tag{5.11}$$

$$\dot{H} + \frac{3}{2}H^2 - \frac{\mu^4}{72(\dot{H} + 2H^2)^2}\left(4\dot{H} + 9H^2 + 2\frac{\ddot{R}}{R} - 6\frac{\dot{R}^2}{R^2} + 4H\frac{\dot{R}}{R}\right) = -4\pi GP_M \, ,$$
$$\tag{5.12}$$

respectively.

These fourth-order equations are cumbersome and therefore extracting their cosmological implications not an easy task in their present form. Instead, Carroll *et al.* [69] performed a specific conformal transformation on the original degrees of

freedom,

$$\tilde{g}_{\mu\nu} = p(\phi)g_{\mu\nu} \,, \qquad p \equiv \exp\left(\sqrt{\frac{16\pi G}{3}}\phi\right) = 1 + \frac{\mu^4}{R^2} \,, \tag{5.13}$$

$$d\tilde{t} = \sqrt{p}dt \,, \qquad \tilde{a}(t) = \sqrt{p}a(t) \,, \tag{5.14}$$

$$\tilde{\rho}_M = p^{-2}\rho_M \,, \qquad \tilde{P}_M = p^{-2}P_M \,. \tag{5.15}$$

where ϕ is a real scalar function on space-time. This transformation has been extensively treated [47], and involves representing metric degrees of freedom in terms of a fictitious scalar field. The transformation leads to the following equation of motion for the transformed expansion parameter,

$$\tilde{H}^2 = \frac{8\pi G}{3}\left(\rho_\phi + \tilde{\rho}_M\right) \,, \tag{5.16}$$

and scalar equation of motion,

$$\phi'' + 3\tilde{H}\phi' + \frac{dV(\phi)}{d\phi} - \frac{(1-3w)}{\sqrt{6}}\rho_{\tilde{M}} \,, \tag{5.17}$$

where a prime denotes differentiation with respect to $\tilde{t}$. We have introduced the potential,

$$V(\phi) = \frac{\mu^2}{8\pi G}\frac{\sqrt{p-1}}{p^2} \,, \tag{5.18}$$

and here we identify the transformed energy density and scalar energy density,

$$\tilde{\rho}_M = \frac{K}{\tilde{a}^{3(1+w)}}\exp\left[-\sqrt{\frac{4\pi G}{3}}(1-3w)\phi\right] \,, \tag{5.19}$$

$$\rho_\phi = \frac{1}{2}\phi'^2 + V(\phi) \,, \tag{5.20}$$

respectively. Carroll *et al.* [69] considered three qualitatively distinct cases, assuming an initial value for the scalar field to satisfy,

$$\phi_0 \ll M_P \equiv \frac{1}{\sqrt{16\pi G}} \,. \tag{5.21}$$

From Equation 5.18, we see that the potential vanishes when $\phi \to 0$ and $\phi \to \infty$. The limit $\phi \to 0$ would normally correspond to the Minkowski vacuum, but from Equation 5.13 it is clear that instead a curvature singularity exists in this limit. Although $\phi \to \infty$ corresponds to $R \to 0$ and seems like a possible Minkowski vacuum solution. However, from Equation 5.17 and Equation 5.13 we see that the solution is oscillatory at asymptotically large values of ϕ and therefore unphysical.

When the initial condition, $\phi'_0 = \phi'_C$, where ϕ'_C is the critical value for which the scalar field comes to rest at the peak of the potential, the scalar field energy density becomes constant. Therefore,

$$\widetilde{H}[\phi'_0 = \phi'_C] = \text{constant} . \tag{5.22}$$

This of course is the hallmark of a de Sitter expansion, albeit under unstable conditions since any perturbations in the scalar field will have it exhibit one of the alternative qualitative possibilities.

For the scenario $\phi'_0 < \phi'_C$, the scalar field never reaches the maximum but rolls back toward $\phi = 0$ and the universe collapses upon itself. As $V \to 0$ and $\widetilde{H}$ goes to a constant, the deceleration parameter and the Ricci scalar, both of which depend upon $\dot{H}$ or $\widetilde{H}'$, are singular since $\dot{H} \sim V' \sim \frac{1}{\sqrt{\phi}} \to \infty$.

Alternatively, the scalar field can be endowed with $\phi'_0 > \phi'_C$ in which case the scalar field becomes quite large with time and the potential behaves as,

$$V(\phi) \to \mu^2 M_P^2 p^{-3/2} = \mu^2 M_P^2 \exp\left(-\sqrt{\frac{3}{2}}\frac{\phi}{M_P}\right) . \tag{5.23}$$

If we seek a power law solution for the scale factor,

$$\tilde{a} \sim \tilde{t}^n \propto p \longrightarrow \widetilde{H} \sim \frac{1}{\tilde{t}} , \tag{5.24}$$

then this implies,

$$\phi' \sim \frac{1}{\tilde{t}} \longrightarrow p \sim \tilde{t}^{4/3} . \tag{5.25}$$

Thus, the scale factors behave as,

$$\tilde{a} \sim \tilde{t}^{4/3} \, , \tag{5.26}$$

$$a \sim t^{2/3} \, . \tag{5.27}$$

It is possible to consider the above situations for the more realistic case of $\tilde{\rho}_M \neq 0$, which was considered in [69]. However the results are no more instructive, and we therefore direct to the aforementioned article for a more thorough discussion.

To this point we have said nothing of the μ parameter. Although the $\phi_0' = \phi_C'$ scenario is unstable, one may argue that this theory holds phenomenological relevance. This eternal de Sitter inflation is not too absurd if the decay rate of the phase is on the order of $\tau^{-1} \sim (14 \text{ Gyrs})^{-1}$ — the inverse age of the universe. In terms of a mass scale, this corresponds to $\mu \sim 10^{-33}$eV. Therefore, one can argue on phenomenological grounds that this theory is worthy of consideration since it clearly gives rise to late-time acceleration. Of course, μ is no better than a tuned parameter serving the function of giving credence to the above statement. Nevertheless, it is a viable alternative to the dark energy mechanism, and as such merits further investigation.

5.2   The Gravitational Response

We have shown in the previous section that with an inverse Ricci scalar term in the gravity action, it is possible to explain late-time acceleration. However, we have yet to see what this theory says about the force of gravity on cosmological and local scales. That is the task of this section, and we will restrict ourselves to the de Sitter solution which was discussed to be unstable, but with a slow enough decay rate to justify its study.

Before we embark on calculating the force of gravity with an inverse Ricci term added to the gravitational action, we should comment on the inherent features of such actions. As was apparent from Equation 5.11 and Equation

5.12, our equations of motions of motion are of the higher derivative variety (that is, they possess more than two time derivatives on one of the degrees of freedom). Typically, higher derivatives bring negative energy degrees of freedom; however, endowing the Lagrangian with nonlinear functions of the Ricci scalar can sometimes be permitted [71]. This will only give rise to a single, spin zero higher derivative degree of freedom. But since the lower derivative spin zero is a constrained, negative energy degree of freedom (the Newtonian potential), its higher derivative counterpart can occasionally carry positive energy.

There have been several recent articles which examine aspects of this model. Dick considered the Newtonian limit in perturbation theory about a maximally symmetric background [72]; while Dolgov and Kawasaki discovered and discussed an instability in the interior of a matter distribution [73]. However, Nojiri and Odintsov have shown than R^2 can be added to the action without changing the cosmological solution, and that the coefficient of this term can be chosen to enormously increase the time constant of the interior instability [74]. Meng and Wang have explored perturbative corrections to cosmology [75]; and others have drawn connections with a special class of scalar-tensor theories [76, 77].

What we wish to consider here is the gravitational response to a diffuse matter source after the epoch of acceleration has set in. The procedure will be to solve for the perturbed Ricci scalar, whence we determine the gravitational force carried by the trace of the metric perturbation. We will constrain the matter distribution to have the property that its rate of gravitational collapse is identical to the rate of spacetime expansion, thereby fixing the *physical* radius of the distribution to a constant value. Further, we impose the condition that inside the matter distribution the density is low enough to justifiably employ a locally de Sitter background, in which case the Ricci scalar can be solved exactly and remains constant.

The Calculation

We shall consider a gravitational action parameterized by $p > 0$,

$$S[g] = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[R - \mu^{2(p+1)} R^{-p} \right].$$

(5.28)

(We employ a space-like metric with Ricci tensor $R_{\mu\nu} \equiv \Gamma^\rho{}_{\nu\mu,\rho} - \Gamma^\rho{}_{\rho\mu,\nu} + \Gamma^\rho{}_{\rho\sigma}\Gamma^\sigma{}_{\nu\mu} - \Gamma^\rho{}_{\nu\sigma}\Gamma^\sigma{}_{\rho\mu}$.) Functionally varying with respect to the metric and setting it equal to the matter stress energy tensor leads to the equations of motion,

$$\left[1 + p\mu^{2(p+1)} R^{-(p+1)} \right] R_{\mu\nu} - \frac{1}{2} \left[1 - \mu^{2(p+1)} R^{-(p+1)} \right] R g_{\mu\nu}$$
$$+ p\mu^{2(p+1)}(g_{\mu\nu}\Box - D_\mu D_\nu)R^{-(p+1)} = 8\pi G T_{\mu\nu}. \quad (5.29)$$

D_μ is the covariant derivative and $\Box \equiv (-g)^{-1/2}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu)$ is the covariant d'Alembertian.

Although one must really solve all components of the field equations Equation 5.29 we can get an important part of the gravitational response by simply taking the trace. We shall also restrict to $p = 1$ for simplicity. Inside the matter distribution the trace equation is,

$$- R + 3\frac{\mu^4}{R} + 3\mu^4\Box\frac{1}{R^2} = 8\pi G g^{\mu\nu}T_{\mu\nu} \equiv \mathcal{T}.$$

(5.30)

(Note that $\mathcal{T}$ is negative.) Normally, one would expect the matter stress energy to be redshifted by powers of the scale factor in an expanding universe. However, recall that this matter distribution possesses a rate of gravitational collapse equal to the rate of universal expansion, and thus $\mathcal{T}$ remains *constant.* Since our matter source is also diffuse, we may perturb around a locally de Sitter background. For the interior solution, we are able to solve for R exactly using Equation 5.30 for the case $\mathcal{T}$ is constant and $D_\mu R = 0$,

$$R_{\text{in}} = -\frac{\mathcal{T}}{2}\left[1 \mp \sqrt{1 + \frac{12\mu^4}{\mathcal{T}^2}} \right].$$

(5.31)

Obtaining de Sitter background obviously selects the negative root. Further, we concentrate on the situation $|\mathcal{T}| \ll \mu^2$,

$$R_{\text{in}} = \sqrt{3}\mu^2 - \frac{\mathcal{T}}{2} + \cdots. \tag{5.32}$$

Outside the matter source we perturb around the de Sitter vacuum solution,

$$R_{\text{out}} = \sqrt{3}\mu^2 + \delta R. \tag{5.33}$$

Substituting Equation 5.33 into Equation 5.30 and expanding to first order in δR yields the equation defining the Ricci scalar correction,

$$\Box \delta R(x) + \sqrt{3}\mu^2 \delta R(x) = 0. \tag{5.34}$$

In our locally de Sitter background the invariant length element is,

$$ds^2 \equiv -dt^2 + a^2(t)d\vec{x} \cdot d\vec{x}, \tag{5.35}$$

with $a(t)$ having the property,

$$H \equiv \frac{\dot{a}}{a} = \text{constant}. \tag{5.36}$$

We can relate the Hubble constant H to the parameter μ via the vacuum Ricci scalar,

$$R = 12H^2 = \sqrt{3}\mu^2. \tag{5.37}$$

Identifying $\Box = a^{-3}\partial_\rho(a^3 g^{\rho\sigma}\partial_\sigma)$, we expand Equation 5.34,

$$\left[\partial^2 - 3H\partial_0 + 12H^2\right]\delta R(t, \vec{x}) = 0, \tag{5.38}$$

where $\partial^2 \equiv -\partial_0^2 + a^{-2}\nabla^2$. It is evident from Equation 5.38 that the frequency term has the wrong sign for stability [69]. However, since the decay time is proportional to $1/H$, we may safely ignore this issue.

Seeking a solution of the form $\delta R = \delta R(Ha\|\vec{x}\|)$ allows us to convert Equation 5.38 into an ordinary differential equation,

$$\left[(1-y^2)\frac{d^2}{dy^2} + \frac{2}{y}(1-2y^2)\frac{d}{dy} + 12\right]\delta R = 0, \tag{5.39}$$

where $y \equiv Ha\|\vec{x}\|$. To solve this equation we try a series of the form,

$$f_\alpha(y) = \sum_{n=0}^{\infty} f_n y^{\alpha+n}. \tag{5.40}$$

Substituting this series into Equation 5.38 yields a solution with $\alpha = 0$,

$$f_0(y) \equiv \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{3}{4}-\frac{\sqrt{57}}{4})\Gamma(n+\frac{3}{4}+\frac{\sqrt{57}}{4})}{\Gamma(\frac{3}{4}-\frac{\sqrt{57}}{4})\Gamma(\frac{3}{4}+\frac{\sqrt{57}}{4})}\frac{(2y)^{2n}}{(2n+1)!}, \tag{5.41}$$

and a solution with $\alpha = -1$,

$$f_{-1}(y) \equiv \frac{1}{y}\sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{1}{4}-\frac{\sqrt{57}}{4})\Gamma(n+\frac{1}{4}+\frac{\sqrt{57}}{4})}{\Gamma(\frac{1}{4}-\frac{\sqrt{57}}{4})\Gamma(\frac{1}{4}+\frac{\sqrt{57}}{4})}\frac{(2y)^{2n}}{(2n)!}. \tag{5.42}$$

Both solutions converge for $0 < y < 1$. Both also have a logarithmic singularity at $y = 1$, which corresponds to the Hubble radius. We can therefore employ them quite reliably within the visible universe.

The solution we seek is a linear combination,

$$\delta R(y) = \beta_1 f_0(y) + \beta_2 f_{-1}(y), \tag{5.43}$$

whose coefficients are determined by the requirements that $\delta R(y)$ and its first derivative are continuous at the boundary of the matter distribution. We employ a spherically symmetric distribution of matter, centered on the comoving origin. If the matter distribution collapses at the same rate as the expansion of the universe, its physical radius is a constant we call ρ. (This means that the comoving coordinate radius is $\rho/a(t)$.) If the total mass of the distribution is M we can

identify $\mathcal{T}$ as the constant,

$$\mathcal{T} = -\frac{8\pi GM}{\frac{4}{3}\pi\rho^3} = -\frac{6GM}{\rho^3}. \tag{5.44}$$

In terms of our variable $y = Ha(t)\|\vec{x}\|$, the boundary of the matter distribution is at $y_0 = H\rho$. Demanding continuity of the Ricci scalar and its first derivative at y_0 gives the following result for the combination coefficients of the exterior solution Equation 5.43,

$$\beta_1 = \frac{3MG}{\rho^3}\left[f_0(y_0) - \frac{f_0'(y_0)}{f_{-1}'(y_0)}f_{-1}(y_0)\right]^{-1}, \tag{5.45}$$

$$\beta_2 = \frac{3MG}{\rho^3}\left[f_{-1}(y_0) - \frac{f_{-1}'(y_0)}{f_0'(y_0)}f_0(y_0)\right]^{-1}, \tag{5.46}$$

where a prime represents the derivative with respect to the argument.

We are now in a position to calculate the gravitational force carried by the trace of the graviton field. The metric perturbation modifies the invariant length element as follows,

$$ds^2 = -(1-h_{00})dt^2 + 2a(t)h_{0i}dtdx^i + a^2(t)(\delta_{ij} + h_{ij})dx^i dx^j. \tag{5.47}$$

Further defining $h \equiv -h_{00} + h_{ii}$ and imposing the gauge condition,

$$h_{\mu\nu}{}^{;\nu} - \frac{1}{2}h_{,\mu} + 3h^\nu{}_\mu(\ln a)_{,\nu} = 0, \tag{5.48}$$

allows us to express the Ricci scalar in terms of h,

$$\delta R = \frac{1}{2}\left(-\partial^2 h + 4H\partial_0 h\right). \tag{5.49}$$

(Recall that we define $\partial_\mu \equiv (\partial_t, a^{-1}\vec{\nabla})$.) Assuming $h = h(y)$ as we did for δR gives the equation for the gravitational force carried by h,

$$\left[(y^2-1)\frac{d}{dy} + \frac{1}{y}(5y^2-2)\right]h'(y) = \frac{2\delta R(y)}{H^2}. \tag{5.50}$$

The solution to Equation 5.50 is,

$$h'(y) = -\frac{2}{y^2(1-y^2)^{3/2}} \int_0^y dy' y'^2 (1-y'^2)^{1/2} \frac{\delta R(y')}{H^2}.$$ (5.51)

At this point it is useful to consider the y values which are relevant. The Hubble radius corresponds to $y = 1$, whereas the typical distance between galaxies corresponds to about $y = 10^{-4}$, and a typical galaxy radius would be about $y = 10^{-6}$. We are therefore quite justified in assuming that $y_0 \ll 1$, and in specializing to the case of $y_0 \ll y \ll 1$. Now consider the series expansions,

$$f_0(y) = 1 - 2y^2 + \frac{1}{5}y^4 + \mathcal{O}(y^6) \qquad , \beta_1 = \frac{3MG}{\rho^3} + \mathcal{O}(y_0^2),$$ (5.52)

$$f_{-1}(y) = \frac{1}{y}\left[1 - 7y^2 + \frac{14}{3}y^4 + \mathcal{O}(y^6)\right], \beta_2 = -\frac{12MGy_0^3}{\rho^3} + \mathcal{O}(y_0^5).$$ (5.53)

We see first that $|\beta_2| \ll \beta_1$ — which means $\delta R(y) \approx \beta_1 f_0(y)$ — and second, that $f_0(y) \sim 1$ — which implies $\delta R(y) \approx -\mathcal{T}/2$. This means that the integrand in Equation 5.51 fails to fall off for $y > y_0$, so the integral continues to grow outside the boundary of the matter distribution. For small $y \gg y_0$ we have,

$$h'(y) = -\frac{2GM}{H^2\rho^3}y + \mathcal{O}(y^3).$$ (5.54)

To see that this linear growth is a phenomenological disaster it suffices to compare Equation 5.54 with the result that would follow for the same matter distribution, in the same locally de Sitter background, if the theory of gravity had been general relativity with a positive cosmological constant $\Lambda = 3H^2$. In that case $\delta R(y) = -\mathcal{T}\theta(y_0 - y)$ and, for $y > y_0$, the integral in Equation 5.51 gives,

$$h'(y)\Big|_{GR} = \frac{\mathcal{T}}{4H^2y^2(1-y^2)^{\frac{3}{2}}}\left\{\arcsin(y_0) - y_0(1-2y_0^2)\sqrt{1-y_0^2}\right\}.$$ (5.55)

$$= -\frac{4GMH}{y^2} + \mathcal{O}(1).$$ (5.56)

The linear force law Equation 5.54 of modified gravity is stronger by a factor of $\frac{1}{2}(\frac{y}{y_0})^3$. For the force between two galaxies this factor would be about a million.

5.3   Remarks on our Calculation and Future Work

We have determined the gravitational response to a diffuse matter source in a locally de Sitter background. Our result is the leading order result in the expansion variable y, the fractional Hubble distance. Equation 5.54 clearly forces us to disregard the class of theories considered here Equation 5.28 when compared to GR with a cosmological constant (for example, the correction to the gravitational force between the Milky Way and Andromeda increases by six orders of magnitude).

The two assumptions made in our analysis were:

- the matter distribution is gravitationally bound,
- the matter distribution has a mean stress energy $|\mathcal{T}| \lesssim \mu^2$.

The second of these assumptions can be viewed rather flexibly if interested only in phenomenological implications. Regardless of whether it is satisfied, we still would expect a linearly growing response far from the source. To see this, recall that the dominant piece of the solution, $f_0(y)$, from equation Equation 5.43 remains constant and approximately equal to one for many orders of magnitude (for instance, $f_0(10^{-8}) - f_0(10^{-3}) \approx 10^{-6}$). Therefore, although the exterior solution would not be very reliable near the matter source, we can be confident that at cosmic or even intergalactic scales perturbing about de Sitter becomes appropriate and a growing solution would still be observed.

This analysis was performed for $p = 1$, but of course nothing restricts us from considering arbitrary powers of the inverse Ricci scalar. To no surprise, however, varying the power only changes the coefficient of the gravitational force leaving its qualitative behavior alone. The instability found by Dolgov and Kawasaki [73] and the growing solution calculated in this work seem to preclude all such theories phenomenologically. The two problems seem to complement one another because

either problem could be avoided by the addition of an R^2 term, which would not alter the cosmological solution [74]. However, avoiding the interior instability seems to require the R^2 term to have a large coefficient, whereas avoiding the exterior growth requires a smaller value [77].

None of these issues diminishes the importance that should be placed on considering novel approaches to understanding the dark energy problem. It is the responsibility of both theorists and experimentalists to construct and constrain candidate theories, and it is truly an exciting epoch of human investigation for which we are just beginning to acquire these capabilities. Greater freedom can be obtained by adding different powers of R. (Note that this generally alters the cosmological solution.) Although such models seem epicyclic when considered as modifications of gravity, the same would not be true if they were to arise from fundamental theory. For example, it can be shown that the braneworld scenario of Dvali, Gabadadze and Shifman [78] avoids both the interior instability and the linearly growing force law [79].

CHAPTER 6
CONCLUSIONS

This thesis has examined alternative explanations to the dark matter and dark energy problems. Each problem has been presented with an alternative that modifies the Einstein-Hilbert action of gravity in four dimensions with a function of the Ricci scalar,

$$S_g[g] = S_{EH}[g] + \frac{1}{16\pi G} \int d^4x \sqrt{-g} f(R) . \tag{6.1}$$

The subsequent phenomenology has been discussed and used to make definitive statements as to the standing of these theories and prospects for future investigation.

Dark matter's successes — particularly a Λ-CDM scenario — leaves many with the impression that its role in galactic rotation curves is a necessary feature. CDM is able to explain galaxy formation by providing enough gravitational presence to ensure luminous matter clumping on the scales we see today. If one takes seriously the Peccei-Quinn mechanism as a solution to the strong-CP problem of QCD, then the axion is a real particle and thus a prime candidate for dark matter. Big bang nucleosynthesis also cannot do without dark matter. Baryonic matter alone is unable to account for the density required to allow BBN to occur. We clearly see that dark matter's connection to the entire cosmology of the universe is too intertwined for its existence to not be taken as a possible reality.

Nevertheless, new gravitational physics which occurs at different scales is certainly not an impossibility. The evidence stated above is only gravitational in nature. That is, it only serves to identify dark matter via its gravitational interactions. Observing a particle gravitationally is an insufficient method of

detection. This only serves to determine the metric in a fixed gauge, whence one may then construct the Einstein tensor and then define the matter-stress energy tensor so as to make the Einstein equation true. Therefore, flipping the solution on its head: there currently exists modifications due to gravity which are capable of accounting for the observed cosmology; and these modifications can be interpreted as the presence of matter stress-energy we call dark matter (much like the organizing principle of perturbation theory in classical general relativity).

The current dark matter profiles which have been studied suffer definite problems which we have discussed in this work to some length. They are unable to explain: the Tully-Fisher relation — the proportionality of the absolute luminosity to the quartic power of the maximum rotational velocity; and Milgrom's law — the fact that dark matter needs to be evoked when satellites possess an acceleration $a \lesssim a_0 \sim 10^{-10}\mathrm{m/s^2}$. Additionally, their fine-tuning features (by virtue of their three parameter fitting) all leave one to conclude that the rotation curves alone — an amazingly consistent phenomenon — cannot presently enable one to say much about the fundamental nature of dark matter.

MOND is purely gravitational at the nonrelativistic level, and by design is constructed to satisfy Milgrom's law. Therefore, at the empirical level, it is vastly superior to dark matter halos. It is at the fundamental level where one properly displays reservations as to its viability in light of the successes of dark matter in several key physical processes. The need for a covariant metric formulation of MOND becomes immediately evident — one which can be directly measured alongside its dark matter competitor.

MOND's relativistic extension has been treated in this thesis by considering the two predominant approaches: the scalar-tensor theories of Milgrom, Bekenstein, and Sanders, and the purely metric approach of Soussa and Woodard. At present, it can be said conclusively that of the scalar-tensor varieties, the TeVeS theory of

Bekenstein is the most viable candidate. Naturally, all approaches are constructed to be able to reproduce the nonrelativistic version of MOND. However, the key issues in extending MOND have been the lack of sufficient gravitational lensing of light, the acausal propagation of dynamical fields, and inherent ambiguities in regard to its cosmological impact.

Bekenstein's TeVeS is the first relativistic version which is able to resolve the first two of these three issues (under appropriate assumptions) and not be a preferred-frame theory. As it is fully relativistic, one may see what it says about cosmology. Presently, no definitive conclusions can be made and is the subject for future work. TeVeS, however, it not without problems. Namely, the large parameter space creates ambiguity and it is not overly clear which observables can set or constrain these parameters. One possibility would be to take advantage of the disparate travel times that gravitons and neutrinos would possess from distant astrophysical sources. We found after a simple computation that the delay in arrival times of a gravitational wave and a pulse of neutrinos could be on the order of a few minutes under reasonable assumptions of the parameters. Therefore one would ostensibly be able to constrain their values (or ratios thereof). Solar system tests serve as good constraining tools in scalar-tensor theories (e.g. Brans Dicke gravity), and therefore it is certainly not unreasonable to have a degree of optimism in the falsifiability inherent to a theory like TeVeS.

The purely metric theory gains the advantage in overall "naturalness" – that is, the purely metric degrees of freedom, if sufficient to describe MOND in all regimes consistent with astrophysical observations, would follow Ockam's razor. Avoiding philosophical vicissitudes, we shall unabashedly assume the preferential treatment of such a feature in the purely metric theory. The avoidance of scalar and vector degrees of freedom results in fewer parameters. What we discovered, unfortunately from the model builder's perspective, is that, under

conservative assumptions, *any* purely metric theory will never give enough lensing due to the conformal invariance of the linearized MOND equations. However, this result should be viewed in a positive light, as any predictive and ultra-restricting statement in physics should. We may conclude that the most plausible way to avoid the lensing disaster in a purely metric formulation of MOND is by foregoing the notion of gravitational stability, a less pleasant but not unprecedented nor unfathomable situation.

In similar fashion, we have surveyed the current landscape of the dark energy problem. Like dark matter, many of the approaches have centered on adding a new component to the universe such as a constant scalar field, a dynamical scalar field designed to turn on at the appropriate time, charged scalars that exhibit long-range forces, etc. By construction, all these models serve their purpose — they give rise to late-time acceleration in the universe. Each, however, begs the question to their detection at the level of new particles and fields.

Contrarily, modifications of the Einstein-Hilbert action interact with *all* matter and energy, and there signatures are in the evolutions and dynamics of the universe's constituents. This thesis has considered specifically the modification of Carroll et al. [69] in which an inverse power of the Ricci scalar is added to the action. This type of term is shown to give rise to late-time acceleration under appropriate assumptions.

The work of Carroll et al. [69] did not consider the effect this kind of term would have on the force of gravity. This work presents this very calculation in a locally de Sitter background for the case of a diffuse matter source. The result clearly shows that this type of term can in no way be phenomenologically viable. The solution possesses a term which grows linearly with distance, and therefore even physics on the cluster scale completely rules out such a model. Further, the lack of a Newtonian limit which surfaces as an instability in the inner regions

of matter sources in Dolgov and Kawasaki's work [73] seemingly dooms such a proposal.

One may consider adding terms proportional to R^2, and R^3, and/or using the Palatini formalism to remove the inherent instabilities of only having a $1/R$ term in the action. Presently, it does not seem clear at all that one may both remove the instability found by Dolgov and Kawasaki *and* the linearly growing solution discussed here. Adding more and more terms to the action in the epicyclic spirit seems counter to how we should seek solutions. However, upon doing so a larger theory or a more fundamental gravitational principle may emerge — we may find these terms to naturally arise from some larger theory, either as an effective field theory, or perhaps from a string theory. Measurement, phenomenology, and consistency are our guides to this end.

Dark energy and dark matter are without question the consensus — the currently orthodox approach to explaining 96% of the universe's energy. It is, undoubtedly, extremely peculiar that we have not directly detected *any* of this 96% — never once. The only means at our disposal to say anything empirically, the sole fashion we may claim to have observed either of these two phenomena, is via gravity. Simply stated: Einstein's theory works. Therefore, changes to it at any scale should and will meet resistance from the wealth of data that exists — not to mention the theoretical challenges which must be overcome. That said, there is serious reasons to believe that general relativity even at the classical level is unable to account for all of the observed universe. The processes discussed here have all been gravitational and their orthodox explanations can all be recast into the form of a purely gravitational solution. This *fact* serves not only as an incentive to search for alternatives, but almost obliges the physicist, in conformity with the scientific spirit, to allow its possibilities.

REFERENCES

[1] H. V. Peiris, E. Komatsu, L. Verde, D. N. Spergel, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. Meyer, L. Page, G. S. Tucker, E. Wollack, and E. L. Wrightland, "First year wilkinson microwave anisotropy probe (wmap) observations: Implications for inflation," *Astrophys. J. Suppl.*, vol. 148, pp. 213, 2003.

[2] Jaiyul Yoo, Julio Chaname, and Andrew Gould, "The end of the macho era: Limits on halo dark matter from stellar halo wide binaries," *Astrophys. J.*, vol. 601, pp. 311–318, 2004.

[3] Marco Battaglia, Ian Hinchliffe, and Daniel Tovey, "Cold dark matter and the lhc," *J. Phys.*, vol. G30, pp. R217–R244, 2004.

[4] Steven Weinberg, "A new light boson?," *Phys. Rev. Lett.*, vol. 40, pp. 223–226, 1978.

[5] R. D. Peccei and Helen R. Quinn, "Cp conservation in the presence of instantons," *Phys. Rev. Lett.*, vol. 38, pp. 1440–1443, 1977.

[6] S. J. Asztalos, E. Daw, H. Peng, L. J. Rosenberg, C. Hagmann, D. Kinion, W. Stoeffl, K. van Bibber, J. LaVeigne, P. Sikivie, N. S. Sullivan, D. B. Tanner, F. Nezrick, M. S. Turner, and D. M. Moltz, "Experimental constraints on the axion dark matter halo density," *Astrophys. J.*, vol. 571, pp. L27–L30, 2002.

[7] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, "First year wilkinson microwave anisotropy probe (wmap) observations: Determination of cosmological parameters," *Astrophys. J. Suppl.*, vol. 148, pp. 175, 2003.

[8] Marc Davis, F. J. Summers, and David Schlegel, "Large scale structure in a universe with mixed hot and cold dark matter," *Nature*, vol. 359, pp. 393–396, 1992.

[9] Edmund Bertschinger, "Simulations of structure formation in the universe," *Ann. Rev. Astron. Astrophys.*, vol. 36, pp. 599–654, 1998.

[10] James E. Gunn and III Gott, J. Richard, "On the infall of matter into cluster of galaxies and some effects on their evolution," *Astrophys. J.*, vol. 176, pp. 1–19, 1972.

[11] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White, "The structure of cold dark matter halos," *Astrophys. J.*, vol. 462, pp. 563–575, 1996.

[12] Julio F. Navarro, Carols S. Frenk, and Simon D. M. White, "A universal density profile from hierarchical clustering," *Astrophys. J.*, vol. 490, pp. 493–508, 1997.

[13] F. J. Sanchez-Salcedo, "Unstable cold dark matter and the cuspy halo problem in dwarf galaxies," *Astrophys. J.*, vol. 591, pp. L107–L110, 2003.

[14] Roelof Bottema, "Simulations of normal spiral galaxies," *Mon. Not. Roy. Astron. Soc.*, vol. 344, pp. 358, 2003.

[15] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White, "Simulations of x-ray clusters," *Mon. Not. Roy. Astron. Soc.*, vol. 275, pp. 720–740, 1995.

[16] Vincent R. Eke, Julio F. Navarro, and Carlos S. Frenk, "The evolution of x-ray clusters in low density universes," *Astrophys. J.*, vol. 503, pp. 569, 1998.

[17] Stacy S. McGaugh and W. J. G. de Blok, "Testing the hypothesis of modified dynamics with low surface brightness galaxies and other evidence," *Astrophys. J.*, vol. 499, pp. 66–81, 1998.

[18] Jacob D. Bekenstein, "Relativistic gravitation theory for the mond paradigm," *Phys. Rev.*, vol. D70, pp. 083509, 2004.

[19] Keith A. Olive, "Big bang nucleosynthesis," *Nucl. Phys. Proc. Suppl.*, vol. 70, pp. 521–528, 1999.

[20] Subir Sarkar, "Big bang nucleosynthesis and physics beyond the standard model," *Rept. Prog. Phys.*, vol. 59, pp. 1493–1610, 1996.

[21] Mario G. Abadi, Julio F. Navarro, Matthias Steinmetz, and Vincent R. Eke, "Simulations of galaxy formation in a lambda cdm universe ii: The fine structure of simulated galactic disks," *Astrophys. J.*, vol. 597, pp. 21–34, 2003.

[22] Wayne Hu, "Structure formation with generalized dark matter," *Astrophys. J.*, vol. 506, pp. 485–494, 1998.

[23] Wayne Hu, "Mapping the dark matter through the cmb damping tail s," *Astrophys. J.*, vol. 557, pp. L79–L83, 2001.

[24] B. Ratra, K. Ganga, N. Sugiyama, G. S. Tucker, G. S. Griffin, H. T. Nguyen, and J. B. Peterson, "Using white dish cmb anisotropy data to probe open and flat-lambda cdm cosmogonies," *Astrophys. J.*, vol. 505, pp. 8–11, 1998.

[25] M. Milgrom, "A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis," *Astrophys. J.*, vol. 270, pp. 365–370, 1983.

[26] J. D. Bekenstein and M. Milgrom, "Does the missing mass problem signal the breakdown of newtonian gravity?," *Astrophys. J.*, vol. 286, pp. 7–14, 1984.

[27] K. G. Begeman, A. H. Broeils, and R. H. Sanders, "Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics," *Mon. Not. Roy. Astron. Soc.*, vol. 249, pp. 523, 1991.

[28] Mordehai Milgrom, "Dynamics with a nonstandard inertia acceleration relation: An alternative to dark matter in galactic systems," *Ann. Phys.*, vol. 229, pp. 384–415, 1994.

[29] R. B. Tully and J. R. Fisher, "A new method of determing distances to galaxies," *Astron. Astrophys.*, vol. 54, pp. 661–673, 1977.

[30] R. H. Sanders and J. R. Fisher, ," *A. Rev. Astron. Astrophys.*, vol. 40, pp. 263, 2002.

[31] W. J. G. de Blok and S. S. McGaugh, "Testing modified newtonian dynamics with low surface brightness galaxies –rotation curve fits-," *Astrophys. J.*, vol. 508, pp. 66–81, 1998.

[32] E. F. Bell and R. de Jong, "Stellar mass-to-light ratios and the tully-fisher relation," *Astrophys. J.*, vol. 550, pp. 212–229, 2001.

[33] Anthony Aguirre, Joop Schaye, and Eliot Quataert, "Problems for mond in clusters and the ly-alpha forest," *Astrophys. J.*, vol. 561, pp. 550, 2001.

[34] Fransisco Prada, Mayrita Vitvitska, Anatoly Klypin, Jon A. Holtzman, David J. Schlegel, Eva K. Grebel, H. W. Rix, J. Brinkmann, T. A. McKay, and I. Csabai, "Observing the dark matter density profile of isolated galaxies," *Astrophys. J.*, vol. 598, pp. 260–271, 2003.

[35] J. Penarrubia and A. J. Benson, "Effects of dynamical evolution on the distribution of substructures," December 2004, Submitted to Mon. Not. Astron. Soc.

[36] J. Bekenstein and M. Milgrom, "Does the missing mass problem signal the breakdown of newtonian gravity?," *Astrophys. J.*, vol. 286, pp. 7, 1984.

[37] Daniel J. Mortlock and Edwin L. Turner, "Empirical constraints on alternative gravity theories from gravitational lensing s," *Mon. Not. Roy. Astron. Soc.*, vol. 327, pp. 552, ,327.

[38] Daniel J. Mortlock and Edwin L. Turner, "Gravitational lensing in modified newtonian dynamics s," *Mon. Not. Roy. Astron. Soc.*, vol. 327, pp. 557, ,327.

[39] J. D. Bekenstein, "Phase coupling gravitation: Symmetries and gauge fields," *Phys. Lett.*, vol. B202, pp. 497–500, 1988.

[40] J. D. Bekenstein, ," *Developments in General Relativity, Astrophysics and Quantum Theory*, p. 156, 1990, IOP Publishing, Bristol.

[41] J. D. Bekenstein, ," *Proceedings of the Sixth Marcel Grossman Meeting on General Relativity*, p. 905, 1992, World Scientific, Singapore.

[42] Jacob D. Bekenstein and Robert H. Sanders, "Gravitational lenses and unconventional gravity theories," *Astrophys. J.*, vol. 429, pp. 480, 1994.

[43] R. H. Sanders, "A stratified framework for scalar-tensor theories of modified dynamics," *Astrophys. J.*, vol. 480, pp. 492–502, 1997.

[44] T. Jacobson and D. Mattingly, "Einstein-aether waves," *Phys. Rev.*, vol. D70, pp. 024003, 2004.

[45] R. P. Woodard, "The ostrogradskian instability of lagrangians with nonlocality of finite extent," *Phys. Rev.*, vol. A67, pp. 016102, 2003.

[46] N. C. Tsamis and R. P. Woodard, "Nonperturbative models for the quantum gravitational back- reaction on inflation," *Annals Phys.*, vol. 267, pp. 145–192, 1998.

[47] Guido Magnano and Leszek M. Sokolowski, "On physical equivalence between nonlinear gravity theories and a general relativistic selfgravitating scalar field," *Phys. Rev.*, vol. D50, pp. 5039–5059, 1994.

[48] N. C. Tsamis and R. P. Woodard, "No new physics in conformal scalar - metric theory," *Ann. Phys.*, vol. 168, pp. 457, 1986.

[49] C. Brans and R. H. Dicke, "Mach's principle and a relativistic theory of gravitation," *Phys. Rev.*, vol. 124, pp. 925–935, 1961.

[50] Paul A. M. Dirac, "Long range forces and broken symmetries," *Proc. Roy. Soc. Lond.*, vol. A333, pp. 403–418, 1973.

[51] Jr. Callan, Curtis G., Robert C. Myers, and M. J. Perry, "Black holes in string theory," *Nucl. Phys.*, vol. B311, pp. 673, 1989.

[52] Michael S. Turner and Martin J. White, "Cdm models with a smooth component," *Phys. Rev.*, vol. D56, pp. 4439–4443, 1997.

[53] H. B. G. Casimir, "On the attraction between two perfectly conducting plates," *Kon. Ned. Akad. Wetensch. Proc.*, vol. 51, pp. 793–795, 1948.

[54] Steven Weinberg, "The cosmological constant problem," *Rev. Mod. Phys.*, vol. 61, pp. 1–23, 1989.

[55] Sean M. Carroll, "The cosmological constant," *Living Rev. Rel.*, vol. 4, pp. 1, 2001.

[56] Lawrence M. Krauss and Michael S. Turner, "The cosmological constant is back," *Gen. Rel. Grav.*, vol. 27, pp. 1137–1144, 1995.

[57] J. P. Ostriker and Paul J. Steinhardt, "The observational case for a low density universe with a nonzero cosmological constant," *Nature*, vol. 377, pp. 600–602, 1995.

[58] Andrew R. Liddle, David H. Lyth, Pedro T. P. Viana, and Martin J. White, "Cold dark matter models with a cosmological constant," *Mon. Not. Roy. Astron. Soc.*, vol. 282, pp. 281, 1996.

[59] Bharat Ratra and P. J. E. Peebles, "Cosmological consequences of a rolling homogeneous scalar field," *Phys. Rev.*, vol. D37, pp. 3406, 1988.

[60] C. Wetterich, "Cosmology and the fate of dilatation symmetry," *Nucl. Phys.*, vol. B302, pp. 668, 1988.

[61] David Wands, Edmund J. Copeland, and Andrew R. Liddle, "Exponential potentials, scaling solutions and inflation," Presented at 16th Texas Symp. on Relativistic Astrophysics and 3rd Particles, (PASCOS '92), Berkeley, CA, Dec 13-18, 1992.

[62] Pedro G. Ferreira and Michael Joyce, "Structure formation with a self-tuning scalar field," *Phys. Rev. Lett.*, vol. 79, pp. 4740–4743, 1997.

[63] Edmund J. Copeland, Andrew R Liddle, and David Wands, "Exponential potentials and cosmological scaling solutions," *Phys. Rev.*, vol. D57, pp. 4686–4690, 1998.

[64] Ivaylo Zlatev, Li-Min Wang, and Paul J. Steinhardt, "Quintessence, cosmic coincidence, and the cosmological constant," *Phys. Rev. Lett.*, vol. 82, pp. 896–899, 1999.

[65] Philippe Brax and Jerome Martin, "The robustness of quintessence," *Phys. Rev.*, vol. D61, pp. 103502, 2000.

[66] T. Barreiro, Edmund J. Copeland, and N. J. Nunes, "Quintessence arising from exponential potentials," *Phys. Rev.*, vol. D61, pp. 127301, 2000.

[67] Tanmay Vachaspati and Mark Trodden, "Causality and cosmic inflation," *Phys. Rev.*, vol. D61, pp. 023502, 2000.

[68] Martina M. Brisudova, William H. Kinney, and Richard Woodard, "Cosmology with a long range repulsive force," *Phys. Rev.*, vol. D65, pp. 103513, 2002.

[69] Sean M. Carroll, Vikram Duvvuri, Mark Trodden, and Michael S. Turner, "Is cosmic speed-up due to new gravitational physics?," *Phys. Rev.*, vol. D70, pp. 043528, 2004.

[70] Shin'ichi Nojiri and Sergei D. Odintsov, "Where new gravitational physics comes from: M-theory," *Phys. Lett.*, vol. B576, pp. 5–11, 2003.

[71] Andrew Strominger, "Positive energy theorem for r+r**2 gravity," *Phys. Rev.*, vol. D30, pp. 2257, 1984.

[72] Rainer Dick, "On the newtonian limit in gravity models with inverse powers of r," *Gen. Rel. Grav.*, vol. 36, pp. 217–224, 2004.

[73] A. D. Dolgov and M. Kawasaki, "Can modified gravity explain accelerated cosmic expansion?," *Phys. Lett.*, vol. B573, pp. 1–4, 2003.

[74] Shin'ichi Nojiri and Sergei D. Odintsov, "Modified gravity with negative and positive powers of the curvature: Unification of the inflation and of the cosmic acceleration," *Phys. Rev.*, vol. D68, pp. 123512, 2003.

[75] Xinhe Meng and Peng Wang, "Modified friedmann equations in r^{-1}-modified gravity," *Class. Quant. Grav.*, vol. 20, pp. 4949–4962, 2003.

[76] Takeshi Chiba, "1/r gravity and scalar-tensor gravity," *Phys. Lett.*, vol. B575, pp. 1–3, 2003.

[77] Shin'ichi Nojiri and Sergei D. Odintsov, "Modified gravity with ln r terms and cosmic acceleration," *Gen. Rel. Grav.*, vol. 36, pp. 1765–1780, 2004.

[78] Gia Dvali, Gregory Gabadadze, and M. Shifman, "Diluting cosmological constant in infinite volume extra dimensions," *Phys. Rev.*, vol. D67, pp. 044020, 2003.

[79] Arthur Lue and Glenn Starkman, "Gravitational leakage into extra dimensions: Probing dark energy using local gravity," *Phys. Rev.*, vol. D67, pp. 064002, 2003.

BIOGRAPHICAL SKETCH

Marc Soussa was born in York, Pennsylvania. He received a B.A. from Cornell University in the field of biochemistry and chemistry. He went on to study high-energy theory in the Physics Department at the University of Florida under the supervision of Richard Woodard.