

TOPICS IN THEORETICAL GRAVITY

A Dissertation

Presented to the Faculty of the Graduate School
of Cornell University

in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Andrew Paul Lundgren

August 2008

TOPICS IN THEORETICAL GRAVITY

Andrew Paul Lundgren, Ph.D.

Cornell University 2008

This dissertation presents four topics dealing with various aspects of gravitation, from theoretical matters to practical issues.

Chapter 2 is about the quasilocal energy, which is a mathematical tool for defining gravitational energy. We extend previous definitions so that they are valid within the event horizon of a black hole. We find that the energy at the center is finite rather than divergent, indicating that the nonlinearities of General Relativity cause a sort of renormalization. We explore a number of examples and point out a problem with some positivity theorems for this type of energy.

Chapter 3 considers the thermodynamics of a charged black hole in a canonical ensemble. We calculate the thermodynamic phase diagram of a black hole with a fixed temperature and charge that is confined in a cavity. We show that the phase diagrams possess the same features as an AdS black hole, suggesting that results such as the AdS/CFT conjecture are at least approximately valid for the more realistic scenario of a black hole in a cavity.

Chapter 4 is about the measurement of gravitational radiation. Laser interferometric observatories are now in operation that can in principle detect some likely astrophysical sources. We study the hyperboloidal family of light beam shapes that have the potential of reducing the thermal noise and therefore increasing the sensitivity to gravitational waves. We show that finite mirror effects are significant and show that small changes in the mirror shape can substantially decrease the thermal noise, increasing the detection range of such

observatories.

Chapter 5 explores a modified theory of gravity called $F(R)$ gravity which was proposed to solve the dark energy problem. We consider forms of $F(R)$ that are intended to mimic standard General Relativity at high densities, but have low-density behavior that can explain the observed acceleration of the cosmological expansion. We discuss the chameleon mechanism for suppressing deviations from standard GR and show that it requires a fine-tuning to function, hence generic $F(R)$ models without such fine-tuning are ruled out by Solar System and cosmological observations.

BIOGRAPHICAL SKETCH

Andrew Lundgren was born in Louisville, Kentucky in 1978, to Paul and LouAnne Lundgren. Two and a half years later, his sister Emily was born and the family moved to Syracuse, New York. After a few more years, the family ended up in Averill Park, NY, a small suburb of Troy. The local school system was quite good and he had a number of wonderful teachers. He learned to love math and science, and was fortunate to be given the opportunity to take two math classes at the Rensselaer Polytechnic Institute in Troy.

In 1997, Andrew graduated high school and became an undergraduate at the Massachusetts Institute of Technology. He studied physics and did his final undergraduate project with Professor Edward Farhi on adiabatic quantum computing. In 2001, he entered Cornell University as a graduate student. After initially wanting to do particle physics and working with Professor Csaba Csaki for a summer, he switched to General Relativity and worked first with Professor James York and later with Professor Eanna Flanagan.

To my family.

ACKNOWLEDGEMENTS

First of all, I must thank my advisor, Jimmy York, for taking me on as one of his last students. His classes and discussions have taught me a very idiosyncratic way of doing General Relativity. He gave me the freedom to work on nearly anything that interested me, a fact reflected in the somewhat random choice of topics in this thesis. He pointed me at a number of interesting and very deep problems which I hope to work on for years to come. I also would like to thank Eanna Flanagan for advising me after Jimmy retired from Cornell, and for having patience with me when I often wandered off track working on the last part of this thesis.

Thanks also to my other co-authors: Bjoern, Ruxandra, Dave, and Mihai. Without them, two of these papers wouldn't have been possible.

I'd like to Saul Teukolsky and Rich Galik for being on my special committee, and asking a lot of great questions at my B Exam. Ira Wasserman, Rachel Bean, Dave Chernoff, Dong Lai, and Larry Kidder also deserve thanks for making the sixth floor a great place to work.

I also want to acknowledge everybody at Astro Coffee Hour, who have given me a broader picture of astronomy. A number of Jimmy's former students and collaborators have been very kind to me, especially David Brown, and I have had very useful conversations with several of them. Of course, I have to thank all of my officemates past and present, including the office next door and the ones around the corner, for occasional help with work and constant help with avoiding work; a similar thanks goes to friends from MIT including Baris, Alyx, and Adnan.

Finally, I have to thank my family, first and foremost my parents and sister, and also my grandparents, aunts, uncles, and cousins. My parents have

been very supportive and have made it possible for me to get an amazing education that gives me the opportunity to do something I really love. My sister has brightened my life with her bizarre sense of humor and flair for writing. I couldn't ask for a more wonderful family.

I have been supported by several travel grants from Cornell, the National Science Foundation, the American Physical Society, and an IUPAP/ISGRG grant. My research was supported by NSF grants PHY-055216 and PHY-0757735.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vii
List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Quasilocal Energy	1
1.2 Black Hole Thermodynamics	3
1.3 Gravitational Wave Detection	4
1.4 Modifications of Gravity	8
2 Self-Renormalization of the Classical Quasilocal Energy	14
2.1 Introduction	14
2.2 The Brown-York Quasilocal Energy	17
2.3 Relationship between action and QLE	21
2.4 Coordinate Independence	25
2.5 deSitter Space and Black Holes	27
2.6 Reissner-Nordstrom	28
2.7 Conclusion	31
3 Charged Black Hole in a Canonical Ensemble	33
3.1 Introduction	33
3.2 The Geometry and Action	36
3.3 Equilibria and Stability	40
3.4 Uncharged Case	42
3.5 Charged Case	43
3.6 Canonical Ensemble	44
3.7 Conclusion	50
4 Finite Mirror Effects in Advanced Interferometric Gravitational Wave Detectors	55
4.1 Introduction	55
4.2 Construction of the Beams	59
4.3 Asymptotic Behavior of Wide Hyperboloidal Beams	64
4.4 Thermal Noises	65
4.5 Eigenvalues of the Propagator	67
4.5.1 Integral Form of the Propagator	69
4.5.2 Discrete Form of the Propagator	70
4.5.3 Meaning of Eigenvalues	71
4.6 Results	72

4.6.1	Finite Mirror Effects	73
4.6.2	Noises for fixed D	77
4.6.3	Hyperboloidal beams with 1 ppm Diffraction Loss	78
4.6.4	Correcting for Finite Mirror Effects	80
4.7	Conclusions	84
5	F(R) Modifications of Gravity	90
5.1	Introduction	90
5.1.1	The dark energy problem	90
5.1.2	The F(R) Idea	91
5.1.3	Original CDTT Model	92
5.1.4	The Chameleon Effect	93
5.1.5	Observational Constraints on F(R) Models	94
5.2	Equations of Motion	95
5.2.1	Exact Solutions	96
5.2.2	Cosmological Evolution	97
5.2.3	Perturbative Solution	98
5.3	The Chameleon Effect	101
5.4	Applying the Constraints	103
5.5	Fine-Tuning	105
5.6	Conclusion	107
	Bibliography	108

LIST OF TABLES

4.1	Advanced LIGO Parameters	58
4.2	The diffraction loss and phase separation for eigenvalues with losses less than 10 percent, in three different hyperboloidal configurations with 1 ppm loss in the fundamental mode. Only axisymmetric modes are shown.	68
4.3	The phase separation for the axisymmetric ($m = 0$) modes with diffraction losses less than 10%, both before and after the iteration scheme is applied. The phases do not change significantly as the mirror is iterated. The absolute value of the eigenvalues with nonzero phase (and hence the diffraction loss) increase upon iterating. Preliminary calculations show that non-axisymmetric modes have diffraction losses increased by the iteration process while the phases change by no more than 5%.	88
4.4	The coating, substrate Brownian and substrate thermoelastic noise are displayed after the iteration process. The diffraction loss is kept constant at 1 ppm. The noises are normalized to noises of the original ($\alpha = \pi$) Mesa (with $D = 9.62$ cm) which gives the 1 ppm diffraction loss. The iteration scheme lowers the noise by about 30% for $\alpha = \pi$ by allowing larger D for the same 1 ppm diffraction loss.	89

LIST OF FIGURES

2.1	QLE computed inside and outside the event horizon for a Schwarzschild black hole. Both axes are in units of the mass M , and the horizon is at $2M$	21
2.2	QLE of deSitter space (positive cosmological constant). Both axes are in the same units, with a mass scale proportional to $1/\sqrt{\Lambda}$	28
2.3	QLE of deSitter-Schwarzschild, a black hole in a spacetime with positive cosmological constant. The units are the same as the plot of the deSitter QLE, and the black hole has an unrealistically large mass so that details can be seen in the plot.	29
2.4	QLE of a Reissner-Nordstrom charged black hole. Both axes are in units of mass of the black hole, and the charge $e^2 = 0.8M^2$	31
3.1	Temperature versus x for (a) $q = 0.1$ and (b) $q = 0.4$	47
3.2	Phase diagram with charge q on the horizontal axis and temperature T on the vertical. The shaded region has three equilibria as opposed to only one outside.	48
3.3	Free energy versus temperature for (a) $q = 0.1$, (b) $q = 0.22 \approx q_{crit}$, and (c) $q = 0.4$	53
3.4	Free energy versus charge for (a) $\theta = 2.33$, (b) $\theta = 2.56$, and (c) $\theta = 3.57$	54
4.1	The cylindrical coordinate system that we use, with the ϕ coordinate suppressed. The dotted lines are the fiducial spheroids, while a mirror is shown schematically, with $h(r)$ indicated.	59
4.2	The generators of the minimal Gaussians that are superposed to produce the $\alpha = 0.8\pi$ hyperboloidal beam. For clarity, only the outermost set of generators are displayed.	60
4.3	(a) The beam intensity profile $ U_\alpha ^2$ and (b) corrections h_α are shown at fixed $D = 10$ cm for different twist parameters α	61
4.4	(a) The diffraction loss for an ($\alpha = \pi$) Mesa configuration ($\alpha = \pi$) is shown as a function of D on a logarithmic plot for several different mirror radii: $R = 16$ cm, $R = 17$ cm and $R = 18$ cm. The diffraction loss computed numerically using Eq. (4.27) (solid, dashed and dotted-dashed lines) exhibits local minima due to finite mirror effects. It can be seen that the minima get narrower as R increases and that they go below the values estimated using the clipping approximation (dotted lines). (b) The fractional difference $ U _{finite}^2/ U _{theory}^2 - 1$ between the theoretical infinite-mirror beam intensity profile and the actual profile given by the first eigenvector of the propagator in Eq. (4.24) is plotted for $D = 10.67$ cm and $R = 16, 17, 18$ and 20 cm. The deviation decreases with R as expected.	74

4.5	The diffraction losses of the fundamental mode and the first excited mode are plotted as a function of D for the $\alpha = \pi$ Mesa configuration. The two curves cross due to finite mirror effects causing an anomalous diffraction loss for the first excited mode.	76
4.6	The noise ratios $\text{Noise}_\alpha / \text{Noise}_{\text{Mesa}} - 1$ for three types of noise are shown as a function of α for fixed $D = 10$ cm. The minimal Gaussian $\alpha = \pi/2$ has the highest noise and Mesa ($\alpha = 0, \pi$) has the lowest noise.	77
4.7	Diffraction loss as a function of D is displayed for $\alpha = \pi, \alpha = 0.95\pi, \alpha = 0.90\pi$. It can be seen that as α decreases the minimum diffraction loss is lower and occurs for larger D .	78
4.8	The minimum diffraction loss in ppm (represented by crosses) and the corresponding D (dots) are shown as a function of α . The solid line represents the best exponential fit of the form $a + \exp(b + c \sin^2 \alpha)$ with $a = 0.094, b = -4.34$ and $c = 2.20$.	79
4.9	(a) The largest values of D giving a strict 1 ppm diffraction loss are shown as a function of α . The discontinuity is caused by finite mirror effects as explained in the text. (b) The corresponding values of the noise, normalized so that the noises for the 1 ppm Mesa ($\alpha = \pi, D = 9.62$ cm) are all equal to 1. The discontinuity in allowed maximum D leads to a sharp drop in noise at $\alpha = 0.91\pi$.	80
4.10	The phase of the fundamental eigenbeam as a function of radius is shown for iteration zero and 250. It can be seen that the iteration scheme drives it closer to zero as expected.	82
4.11	The diffraction loss in ppm calculated using the clipping approximation is compared to that using the propagator eigenvalues for the iterated and original mirrors as a function of D . As before, the configuration studied is $\alpha = \pi$ Mesa with $R = 17$ cm. The iteration process lowers the diffraction loss by a factor of 30 to 100.	83
4.12	The diffraction loss is shown as function of iteration number for a $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm. The original beam (not shown) has diffraction loss at 46.5 ppm, and it can be seen that after a few iterations the diffraction loss begins to converge to an exponential with lower bound ~ 1 ppm. The best fit exponential is given by $0.96 + 1.616 \exp(-0.013i)$ ppm, where i is the iteration number.	84
4.13	The intensity profile $ U_\pi ^2$ for the mirror with $R = \infty, R = 17$ cm uniterated and $R = 17$ cm at iteration 250 are compared for the $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm. The finite mirror effects induce oscillations in the intensity profile that do not disappear when the mirror is corrected. An inset shows the central 8 cm 'plateau' of the beam in detail.	85

4.14 The correction h_α to the mirror at iteration 0 and at iteration 250 are compared for the $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm, $R = 17$ cm. The iteration scheme introduces some bumps on the mirror of the size ~ 2 nm. The inset shows the central 8 cm of the mirror in more detail. 85

CHAPTER 1

INTRODUCTION

1.1 Quasilocal Energy

Despite the mathematical elegance of General Relativity, finding and interpreting solutions of the equations can be very difficult. The strong and electroweak forces are described by gauge fields on spacetime, while GR is a theory of spacetime itself. GR is a very nonlinear theory and general covariance also gives it a very large gauge group. Finally, the energy of GR cannot be localized due to the Equivalence Principle.

The main problem in solving GR is that we are solving for the metric of spacetime itself, so even defining the source is problematic. For instance, in a binary black hole system the separation between the two black holes is determined from the metric, but the metric is determined by the position of the black holes. Finding a metric representing a binary system with a given separation requires an iterative procedure to find a self-consistent solution. Even then, a different solution method may yield the same metric written in very different coordinates, and it is difficult to tell that the two metrics are physically the same. As we enter the age of numerical relativity we face the problem of finding gauge-independent quantities that can identify a metric's physical meaning regardless of the coordinate system in which it is written.

A related problem is the difficulty of interpreting the physical meaning of a metric in the absence of symmetries. A flat spacetime metric has ten symmetries: time translation, three space translations, three rotations, and three boosts.

These symmetries are expressed as Killing vector fields ξ_μ obeying Killing's equation

$$\mathcal{L}_\xi g_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0, \quad (1.1)$$

where \mathcal{L}_ξ is the Lie derivative along ξ and ∇ is the covariant derivative compatible with the metric. A more general metric will have fewer symmetries. The Schwarzschild metric of a non-spinning black hole has time translation and three rotational symmetries. As a result, the energy and angular momentum of a test particle can be unambiguously defined, but there is no clear definition of linear momentum. The Kerr metric of a rotating black hole only has an axial rotation symmetry, meaning that only one component of angular momentum can be defined. A general dynamical spacetime will not possess any symmetries of this type, and so it's not clear what how to define the energy or angular momentum of a particle, much less that of the gravitational field.

We will be most concerned with the difficulty of localizing gravitational energy. The root of the problem is the Equivalence Principle, which states that it is impossible to locally distinguish between a gravitational field or an accelerating frame. As a consequence, a freely-falling observer experiences no gravitational effects. To avoid this problem, we consider a quantity defined on a surface that represents the energy inside that surface. These quantities depend on the metric on the surface and derivatives of the metric normal to the surface. The particular form of the quasilocal energy we study was defined by Brown and York [1]. We are specifically interested in the question of what the energy of the gravitational field is at the center of a black hole. The energy of a point particle's field may be expected to diverge, but GR is non-linear and there is the complication of the event horizon.

Chapter 2 (in collaboration with Bjoern Schmekel and James York) extends the definition of the quasilocal energy to the interior of an event horizon. The field energy would be expected to diverge near a point particle, but in the simple case of a Schwarzschild black hole the quasilocal energy at the center is zero. Renormalization in quantum field theories is necessary to deal with the infinite self-energy problem. In GR, the nonlinear nature of the field seems to provide an effective renormalization. Black holes with a cosmological constant are shown to have the same feature, with diverging energy at large distances according to the subtraction procedure that we use. Finally, we show that the charged black hole has negative but finite energy at the center and argue that this is a physical effects by considering geodesics that approach the center. Although there is a positivity theorem [2] that would seem to contradict our result, we argue that the topological properties of the timelike singularity at the center of a charged black hole violate the conditions of the theorem. Since other spacetimes, such as the Kerr spinning black hole, also have a timelike singularity, this is an important flaw and it points to a need for improved theorems.

1.2 Black Hole Thermodynamics

In Chapter 3, we consider the thermodynamics of a charged black hole enclosed in a cavity. The charge in the cavity and the temperature at the walls are fixed, yielding a canonical ensemble. We derive the phase structure and stability of black hole equilibrium states. We compare our results to that of other work which uses asymptotically anti-deSitter boundary conditions to define the thermodynamics. The thermodynamic properties have extensive similarities which suggest that the idea of AdS holography is more dependent on the existence of

the boundary than on the exact details of asymptotically AdS metrics.

1.3 Gravitational Wave Detection

Gravitational waves are the ripples in spacetime caused by massive moving objects. Due to the weakness of gravity, they have not yet been detected directly, but hopefully will be in the next decade or so. This section will discuss the nature of gravitational waves, some of their sources, and some methods for detecting them. Chapter 4 will then focus on an improvement to the design of one part of the optical system of LIGO, the Laser Interferometric Gravitational-wave Observatory.

Gravitational waves are usually explained by analogy to the more familiar electromagnetic (EM) waves. Both propagate in vacuum at the speed of light, and carry energy, momentum, and angular momentum. Moving electric charges produce EM waves; any type of energy is a source of gravity and any movement of energy can be a source of gravitational waves. Moving masses, particularly astrophysical ones, are the only source we will be interested in. One difference now becomes apparent, because there are positive and negative electric charges, and so electromagnetic forces can repel or attract. Gravitation is universally attractive and the “gravitational” charge therefore only has one sign.

Electromagnetic waves have only two polarizations; in terms of field theory, the photon is a massless spin-one (vector) boson. EM waves with transverse electric fields are possible, while longitudinal fields (along the direction of propagation) are forbidden. As a result, the lowest multipole produced by an

isolated source is a dipole. To produce a monopole EM wave would require a changing total charge of the source, which due to current conservation does not happen for an isolated source. In the gravitational wave case, the lowest multipole is a quadrupole. The lack of monopole and dipole waves is a result of mass and momentum conservation laws, and the fact that gravitational “charge” only has one sign.

General Relativity is non-linear, but when linearized around flat spacetime the result is a field theory of a massless spin two boson called the graviton, which has two polarizations. The action of a gravitational wave on test particles is similar to the action of a tide on a planet. A wave moving along the z axis will squeeze particles along the x axis and stretch them along the y axis. Half a cycle later, it will stretch along the x axis and squeeze along the y. Waves of the other polarization are rotated 45 degrees, compared to 90 degrees for EM waves.

The strongest gravitational waves will be produced by sources with large masses and relatively small separations moving near the speed of light. Astrophysical sources produce the only gravitational waves that can conceivably be detected. There are several different sources; anything that involves motions of large amounts of mass may produce a detectable signal. Supernovae can produce bursts of gravity waves as long as the explosion is substantially non-spherical. Spinning neutron stars may lose angular momentum through gravitational radiation. Easier to detect are binary systems of white dwarfs, neutron stars, or black holes (in any combination). As these systems radiate angular momentum, the objects spiral closer until finally they merge into a single black hole. While no gravitational waves have been directly detected, the effects of gravitational wave emission have been measured indirectly. In 1974, Hulse

and Taylor reported timing observations of a pulsar in a binary that showed an inspiralling orbit which matched analytical post-Newtonian approximations.

Direct detection requires measurement of the squeezing and stretching of test masses caused by a passing gravitational wave. The natural shape of a gravitational wave antenna is a large L . When one arm of the detector is being squeezed the other arm is being stretched, and vice versa. Measurement of the relative length change of the two arms then directly yields the amplitude of the wave (rather than the much more difficult to measure intensity). The effect of a gravitational wave is quoted as a fractional strain on the apparatus. For equal-mass compact objects with mass M and separation d at a distance r from the detector, the strain h is on the order of

$$h \sim \frac{G^2 M^2}{c^4 d r} . \quad (1.2)$$

A typical example [4] is $M = 10 M_{\text{Sun}}$, $r = 100 \text{ Mpc}$, and a separation of 10 times the Schwarzschild radius of the objects, which yields a strain of $h \sim 10^{-21}$.

LIGO, the Laser Interferometric Gravitational-wave Observatory, has recently been built and has reached its design sensitivity. Each LIGO installation (there is one in Washington and one in Louisiana) essentially consists of four test masses, one at each end of two arms that are at a 90 degree angle. The relative length change of the two arms is monitored and a readout of the strain directly gives the gravitational wave amplitude. The arms are each four kilometers long, giving a typical displacement of 10^{-18} m or less than the width of a nucleus. To circumvent this difficulty, each arm of LIGO is a Fabry-Perot cavity whose mirrors are the test masses, and the entire device is a Michelson interferometer. These and other optical tricks allow LIGO to approach the required displacement sensitivity.

The displacement sensitivity is obviously limited by several sources of noise, each of which is frequency dependent. Seismic and gravitational gradient noise are dominant at low frequencies, below 50 Hz, and are so severe that this is effectively the lowest frequency that LIGO can observe. There are two types of quantum noise associated with the laser light. Shot noise is dominant at high frequencies, coming from fluctuations in the number of photons; radiation pressure noise is caused by the same fluctuations but is important at low frequencies. Finally, there is thermal noise which is caused by thermal fluctuations in the surface of the mirror and is most important at intermediate frequencies between the 50 Hz seismic floor and the effective limit at 1000 Hz set by shot noise.

Although LIGO reached its design sensitivity, no gravitational wave signals were detected in more than a year of data. The design sensitivity for Initial LIGO was not sufficient to detect sources at great distances, but it was mostly meant to be a proof of concept and design testbed. A minor upgrade is currently underway, and in a few years a much more ambitious upgrade called Advanced LIGO is planned. To decrease shot noise, the laser power is greatly increased, and the mass of the mirrors is increased to control the radiation pressure noise. Seismic isolation is improved lowering the seismic floor to 10 Hz. The thermal noise is now the most important contribution in the intermediate frequency range.

In Chapter 4 (in collaboration with Ruxandra Bondarescu, David Tsang, and Mihai Bondarescu), we consider an improved mirror design that can reduce noise in Advanced LIGO. In the current design, the laser beams have a Gaussian intensity profile which are quite narrow compared to the width of the mirror. The laser light is required to bounce within the Fabry-Perot cavities with very

low losses, so the beams are narrow in order to avoid losing too much energy to diffraction. A wider beam would average over more of the mirror surface and decrease the thermal noise. We study the hyperboloidal family of beams which are wider than Gaussians with equivalent diffraction loss. We calculate the thermal noises of this family, and also discover that finite mirror effects are substantial for the Advanced LIGO design. We demonstrate how finite mirror effects and the hyperboloidal beams can be taken into account to improve the mirror design and increase the sensitivity of the Advanced LIGO design.

1.4 Modifications of Gravity

Cosmology has undergone a revolution in the last decade, and we now have many lines of evidence telling us that most of the universe is unseen. Only about 4 percent of the energy density of the universe is in the form of galaxies or gas that we can detect directly from electromagnetic observations (this matter is called baryonic by cosmologists because baryons make up most of the mass). Relativistic particles such as photons and neutrinos make up only a negligible fraction of the energy density. The remaining 96 percent is only known through its gravitational influence, and is divided into two categories: dark matter and dark energy. We will see that dark matter is likely composed of exotic but not unnatural particles. Dark energy may be explained in GR by the addition of a cosmological constant, but the associated energy scale is unnaturally small and so this explanation is unsatisfying. In Chapter 5, we will discuss a theory of modified gravity that presents an alternative explanation of dark energy.

Cold dark matter (CDM) makes up approximately 26 percent of the energy

density of the universe. It is distinct from normal matter in that it does not interact electromagnetically at all (hence the name “dark” matter), so that it is only detectable through its gravitational influence. The presence of dark matter was originally inferred from the motions of galaxies in the Coma cluster [5]. The virial theorem shows that the velocities of galaxies in the cluster implied at least an order of magnitude more mass in the cluster than was observed. Further evidence comes from rotation curves of spiral galaxies [6], which give the picture of a disk of baryonic matter embedded in a larger, spherical, and more massive halo of dark matter (a particular example is the nearby Triangulum galaxy [7]).

Because the evidence for dark matter comes only from gravitational effects, it might seem that a modified gravity theory could eliminate the need for dark matter. In galaxies and clusters, dark matter and baryonic matter have different spatial structures but tend to share a common center. However, if two clusters collide, the gas in the clusters will interact and be slowed by ram pressure, while the dark matter will not be affected, so the two types of matter will be separated spatially. This is exactly the situation observed in the so-called “Bullet Cluster” [8]. X-ray observations located the gas in the cluster, while weak lensing maps the position of mass in the cluster by its bending of light passing through the cluster. The dark matter is clearly seen to pass through the collision while the gas stays behind.

The current paradigm is that dark matter is primarily composed of some massive non-baryonic particle that does not interact electromagnetically and has a small cross-section for weak interactions. Constraints from Big Bang nucleosynthesis require the particle to be non-baryonic, although some small fraction of the dark matter may simply be baryonic matter that is difficult to detect.

Structure formation requires that the dark matter must be cold, meaning that throughout most of the evolution of the universe the dark matter particles must be moving non-relativistically; this is most easily accomplished by making the particles massive.

From the relatively well-understood dark matter, we turn our attention to the more mysterious dark energy which makes up the remaining 70 percent of the universe. Like dark matter, dark energy does not interact electromagnetically and is only known through its gravitational effects. Also, no structure is seen; dark energy appears to have a constant density throughout space (and apparently constant in time), so it only has cosmological effects. The key piece of evidence for dark energy is the acceleration of the expansion of the universe. Dark matter, like baryonic matter and radiation, acts to slow the expansion of the universe, so dark energy must be something different.

To simplify our discussion of cosmology, we will assume that spatial sections of the universe are flat. This is supported by observations of fluctuations in the cosmic microwave background and of baryon acoustic oscillations. Zero spatial curvature is also a prediction of the inflationary model. On scales larger than 100 megaparsecs, the universe is approximately homogeneous and isotropic, and is described by the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric

$$ds^2 = -dt^2 + a(t)^2 (dx^2 + dy^2 + dz^2) . \quad (1.3)$$

Here t is the time coordinate and $a(t)$ is called the scale factor. The cosmological expansion means that a is increasing in time. The Hubble parameter $H = \dot{a}/a$, with an overdot denoting the time derivative, measures the rate of expansion of the universe. The deceleration parameter

$$q = -\frac{1}{H^2} \ddot{a}a \quad (1.4)$$

is positive if the expansion is decelerating, which is expected if matter or radiation is the dominant component of the universe.

The Hubble and deceleration parameters are measured by observations [9] of distant supernovae of type Ia. Measurement of spectral lines in a supernova yields the redshift of the object, which corresponds to a via $a = \lambda_{received}/\lambda_{emitted}$. Type Ia supernovae are also standard candles, meaning that their luminosities can be calibrated. In an expanding universe, the flux from a distant supernova does not fall off exactly with $1/r^2$ but instead an integral of H between the supernova and the observer is involved. Hundreds of supernovae have been measured and the deceleration parameter is negative.

Let us see now how to accomodate this result in standard GR using a cosmological constant. Substitution of the FLRW metric into Einstein's equation gives equations for the Hubble and deceleration parameters; later, our modified theory of gravity will alter these equations. The right hand side of Einstein's equation is the stress-energy tensor, which will have contributions from matter, radiation, and the cosmological constant. We can model each of these components as a fluid with density ρ_i and pressure p_i ; the total density and pressure are given by $\rho = \sum \rho_i$ and $p = \sum p_i$. The scale factor evolves according to:

$$H^2(t) = \frac{8\pi G}{3}\rho \quad (1.5)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) . \quad (1.6)$$

The flatness of the metric is assumed in deriving Eq. (1.5). The time evolution of the pressure and density can be simplified by assuming that the pressure of each component is related to the density by $p_i = w_i\rho_i$ and using energy conservation. Then the density depends on the scale factor as

$$\rho_i = \rho_{i,0}a^{-3(1+w_i)} . \quad (1.7)$$

Matter on cosmological scales can be modelled as pressureless dust and so it has $w = 0$; the density is proportional to a^{-3} because the volume of the universe increases while the number of particles stays constant. Radiation has $w = 1/3$ and $\rho \propto a^{-4}$. From Eq. (1.6), we see that any substance with $w < -1/3$ can cause accelerated expansion.

A cosmological constant can be included as a fluid whose density does not change with scale factor, so it has $w = -1$. This can clearly cause accelerated expansion. We now use cosmological measurements to determine the value of the cosmological constant. The most important measurement is the value of the Hubble parameter today, which is denoted H_0 . The value we will use is $H_0 = 72 \text{ km/s / Mpc}$ or in SI units, $H_0 = 2.3 \times 10^{-18} \text{ sec}^{-1}$. The current density of the universe is $\rho_0 = 3H_0^2/8\pi G = 9.7 \times 10^{-27} \text{ kg/m}^3$.

The natural value of the cosmological constant coming from quantum field theory arguments is about 60 orders of magnitude above the value that is actually measured. It would require a tremendous amount of fine-tuning to produce the actual value that we measure. An alternative explanation for the cosmic acceleration would be a modified theory of gravity. Because only cosmological observations probe gravity at very low curvatures, there could be some new dynamics not predicted by GR. This theory is interesting because it is an infrared (low energy) modification rather than the UV corrections that appear in quantum field theories.

Chapter 5 (in collaboration with Eanna Flanagan) explores F(R) theories of gravity as a solution to the dark energy problem. The chameleon mechanism that allows the theories to mimic GR at high curvatures is described. We argue that the chameleon effect must apply at scales at least as large as the galaxy,

while at the larger cosmological scales the $F(R)$ modifications are evident. We apply this constraint to several models that contain adjustable parameters, and find the values of the parameters. The conclusion is that these parameters must be fine-tuned for most models, and models without fine-tuning are contrived and are perhaps unlikely to be derivable from any fundamental physics. While the $F(R)$ theories cannot be ruled out on this basis, they do not provide a natural solution to the dark energy problem.

CHAPTER 2
SELF-RENORMALIZATION OF THE CLASSICAL QUASILOCAL
ENERGY

Originally published as A. Lundgren, B. Schmekel, and J. York, *Phys. Rev. D* **75**, 080426 (2007). Copyright 2007 by the American Physical Society.

2.1 Introduction

It is a fundamental fact of general relativity (GR) that there is no such concept as the local energy of the gravitational field. The local effects of gravity can be removed by transforming to a freely falling frame. A neutral object at the origin of a freely falling frame will not experience any gravitational acceleration. Tidal forces will remain, but they only act on particles that are separated by some distance. Since gravity has no local effect, there exists no local energy.

Various definitions of local energy densities can be made by making reference to special coordinate systems or background metrics. Heuristically, if we have an observer that we consider static we could use its acceleration as a measure of the local gravity. More sophisticated approaches yield a variety of quantities that are useful for certain applications. The structure of GR is such that local quantities representing an energy do not exist, therefore any attempt to define them must use concepts that are not a natural part of the theory, i.e., a special coordinate system. Studying the asymptotic behavior of the metric, as done in post-Newtonian approximations, or the behavior of the metric at spatial or null infinity [11, 12, 13, 10] leads to more useful and natural formulas for the

energy. From these ideas, we are led to the idea of finding the energy inside a given finite region rather than the energy at a point.

Quasilocal energy (QLE) is the energy inside a two-dimensional surface. The surface could be a sphere enclosing a star or a black hole, a small box enclosing some matter undergoing cosmological expansion, or a complicated, even disconnected, surface in the spacetime. In this paper, we follow the method of Brown and York [1, 14] which derives an energy from a Hamilton-Jacobi argument involving the canonical action. This QLE has many useful properties. For example, it agrees with the Newtonian limit for a spherical star, is applicable in thermodynamic problems [16, 15], and the asymptotic limit at Euclidean infinity is the ADM expression for energy. Furthermore, it can be directly obtained from the Hamiltonian of the same action principle (footnote 14 of [15]) without the need for any other geometric structures. There are many formulas for other quasilocal energies [17] (and many references given therein), derived using different methods and often having different properties.

We can define the quasilocal energy of the electric field in classical physics for comparison. The electric field of a point charge falls off with $1/r^2$, and the energy density equals the field strength squared. We have for the energy inside a surface of radius R

$$E(R) \propto \int_0^R \left(\frac{1}{r^2}\right)^2 r^2 dr \quad (2.1)$$

which diverges because the charge is pointlike. This problem remains in quantum electrodynamics and requires renormalization, where another infinite quantity is subtracted to leave a finite remainder.

In GR, the situation is somewhat different because objects of a given mass cannot have an arbitrarily small size. Once they become too small, they col-

lapse to form a black hole and an event horizon forms. The resulting object is effectively the size of the event horizon, and outside observers are shielded from the infinities at the center. The QLE for a Schwarzschild black hole of mass M has the large distance limit

$$E(R) = R \left(1 - \sqrt{1 - \frac{2M}{R}} \right) \approx M + \frac{M^2}{2R}. \quad (2.2)$$

The QLE becomes undefined at $R \leq 2M$, the radius of the event horizon. The large distance limit suggests that the energy will diverge at the center, although we are protected from seeing this behavior by the event horizon. However, this is still something of a problem because an observer can fall in through the horizon in finite proper time and survive to see the interior of the black hole. It may be useful to have a definition of energy for observers inside the event horizon.

A major issue that arises when defining QLE is which observers to use. We can imagine that the boundary is made of a fleet of observers, whose four-velocities we are free to choose. Energy is not an invariant, so boosting the velocities of the observers will change the energy that they measure. The observers that are stationary with respect to the boundary, i.e., their four-velocity is perpendicular to the normal, seem to be the most natural choice. We will show explicitly how this prescription leads to a QLE that does not depend on the time slicing. While there are other possible prescriptions, we prefer this one because it can be defined in terms of quantities on the boundary, and it has properties that are desirable from a physical perspective.

The boundaries that we use in this paper are spheres concentric with the black hole. The QLE will be a function of the size of the sphere. We express this quasilocally, in terms of quantities defined on the boundary, using the area of

the sphere. When we refer to the “radius”, we mean the quantity r for which $4\pi r^2$ is the area of the spherical boundary. We do not use the proper distance to the center of the coordinates because this depends on the time slice. In the metrics that we study, when the boundary is inside a horizon then the QLE is more than r in natural units. For instance, in the Schwarzschild metric the QLE for a boundary exactly on the horizon is $2M$.

The derivation in [1] involved the boundary term of the Hilbert action. In this paper, we write the action for a general spherically-symmetric and static metric and show how the QLE formula can be modified so that the boundary term is treated correctly in either case. Surprisingly, when not coupled to other fields, the energy of the singularity at the center of a black hole is *zero*. The energy climbs toward a maximum value at a radius inside the horizon, and at the horizon has an infinite downward slope. The charged black hole has a finite but *negative* energy at the singularity.

2.2 The Brown-York Quasilocal Energy

We now review the quasilocal energy defined in [1]. The basic idea is to use the Hamilton-Jacobi method in classical mechanics of expressing the energy as a variation of the action with respect to the endpoints. The generalization to curved spacetimes results in the following definition

$$E = \frac{1}{\kappa} \int_B d^2x \sqrt{\sigma} (k - k_0) \quad (2.3)$$

where $\sigma_{\mu\nu}$ is the induced metric on the boundary

$$\sigma_{\mu\nu} = g_{\mu\nu} + u_\mu u_\nu - n_\mu n_\nu \quad (2.4)$$

and σ is its determinant. In the last equation u^μ is a future pointing timelike unit normal for the spacelike hypersurface Σ whereas n^μ is an outward pointing spacelike normal to the boundary 3B which is also normal to B if $u \cdot n = 0$ which is assumed in this definition. The constant $\kappa = 8\pi G$ is just a constant of proportionality, and in natural units is just 8π .

The k in the above equation is the trace of the extrinsic curvature of the two-boundary's embedding into the spacelike hypersurface Σ . The k_0 term is the energy of the vacuum, which must be subtracted to obtain the physical energy. In our case, we take flat space as the vacuum so that k_0 is the trace of the extrinsic curvature for the same two-boundary embedded in flat space. It is natural to add such a term because otherwise the intrinsic geometry of the surface would contribute to the energy, even with no gravitational energy present.

$E(R)$ has been computed already in [1] for 4 dimensional spherically symmetric objects outside the event horizon. The metric of the spacelike slice can be written as

$$ds^2 = f(r)^{-2}dr^2 + r^2d\Omega^2 \quad (2.5)$$

and the unit normal to constant r surfaces is

$$n^\mu = (0, f(r), 0, 0). \quad (2.6)$$

The extrinsic curvature of the two-boundary is

$$k_{\mu\nu} = -\sigma_\mu^\alpha \nabla_\alpha n_\nu ; \quad k = -\sigma_\mu^\alpha \nabla_\alpha n^\mu \quad (2.7)$$

where the above covariant derivative is taken in the spacelike slice, and σ is serving as a projection operator; it can be found from

$$\sigma_{\mu\nu} = h_{\mu\nu} - n_\mu n_\nu \quad (2.8)$$

$$\sigma_{\mu}^{\nu} = h^{\nu\rho}\sigma_{\rho\mu} = \delta_{\mu}^{\nu} - n_{\mu}n^{\nu} \quad (2.9)$$

where $h_{\mu\nu}$ is the metric of the spacelike slice (2.5). The only connection coefficients that we need are

$$\Gamma_{r\theta}^{\theta} = \Gamma_{r\phi}^{\phi} = \frac{1}{r}. \quad (2.10)$$

and we obtain

$$k = -2\frac{f(r)}{r}. \quad (2.11)$$

The reference term k_0 comes from setting $f(r) = 1$ so that we are embedding the sphere in flat space. For a spherical star with energy density $\rho(r)$, we have [18]

$$f(r) = \sqrt{1 - 2m(r)/r} \quad (2.12)$$

$$m(r) = 4\pi \int_0^r r'^2 dr' \rho(r') \quad (2.13)$$

or for a black hole we simply have $m(r) = M$ and $f(r) = \sqrt{1 - 2M/r}$. With

$$\sqrt{\sigma} = r^2 \sin \theta, \quad (2.14)$$

the QLE becomes

$$E(R) = R \left[1 - \left(1 - \frac{2m(R)}{R} \right)^{1/2} \right] \quad (2.15)$$

as long as r is a spacelike coordinate.

Specializing to the Schwarzschild case where $m(r) = M$, there is a horizon at $r = 2M$. Inside the horizon the r coordinate becomes timelike as revealed by inspecting (2.5). It is not clear what normal to use once this happens. We will try one particular choice, investigate the properties of the resulting QLE, then justify our choice in the next two sections by investigating the behavior of the action on the boundary. Our guess for the correct normal to use is

$$n^{\mu} = (0, -(2M/r - 1)^{1/2}, 0, 0) \quad (2.16)$$

and the resulting QLE is

$$E(R) = R \left[1 + \left(\frac{2M}{R} - 1 \right)^{1/2} \right] \quad (2.17)$$

inside the horizon. The QLE of the entire Schwarzschild metric is plotted in Fig. (2.1) and shows three striking features.

First, the quasilocal energy at the singularity is zero. In Newtonian gravity, the energy of the gravitational field would diverge at the center for a point particle. So the nonlinearity of general relativity has removed this infinity, and gives us a picture where the singularity is not the most important feature of the black hole. We expected that the mass seen at infinity would reflect the mass of the singularity, but this seems not to be the case. The black hole looks like an extended object when we consider the second feature, that the QLE attains its maximum inside the horizon at a radius of $1 + \frac{\sqrt{2}}{2}M$. Most of the energy of the black hole seems to be “stored” just inside the horizon.

The third striking fact is that the derivative of the QLE matches across the horizon, but is infinite there. This is not simply a coordinate effect, because the r coordinate is measuring a real physical quantity (the area of the sphere). The energy should be continuous on physical grounds, but the derivative might not be. In fact, if we had chosen the opposite sign of the normal, there would be a cusp in the QLE at the horizon. We will show that this is the correct choice in the next section.

We note in passing that in three-dimensional spacetime the QLE is constant everywhere, so these features would be missing. In $2 + 1$ spacetime, there is no horizon and the metric only possesses a conical defect at the center. This is

because the Schwarzschild metric in n dimensions is

$$ds^2 = - \left(1 - \frac{2m}{r^{n-3}}\right) dt^2 + \left(1 - \frac{2m}{r^{n-3}}\right)^{-1} dr^2 + r^2 d\Omega^2 \quad (2.18)$$

In three dimensions with $d\Omega^2 = d\theta^2$ and $\sqrt{\sigma} = r$ the QLE is constant everywhere.

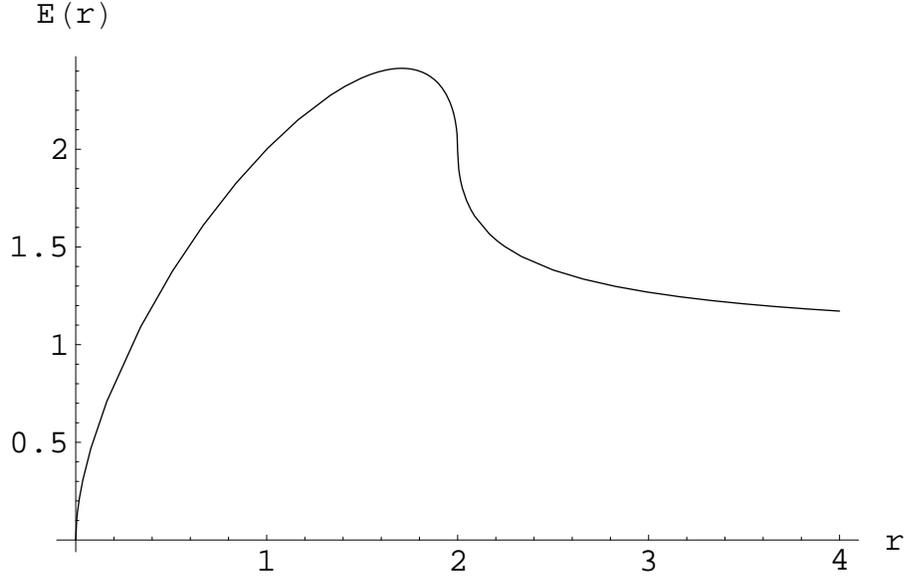


Figure 2.1: QLE computed inside and outside the event horizon for a Schwarzschild black hole. Both axes are in units of the mass M , and the horizon is at $2M$.

2.3 Relationship between action and QLE

The above calculation is not convincing because it is not clear that we have chosen the correct unit normal. A different choice would change the sign and give a much different result. To fix this ambiguity, we will go back to the derivation of QLE from the action. The extrinsic curvature formula for the QLE arises

from the boundary term in the action. From the behavior of this term inside and outside the horizon, we can determine how to modify the QLE.

We begin by considering a region of spacetime M . The spacetime is foliated by spacelike hypersurfaces Σ , which can play the role of moments of time. The region of spacetime will be bounded by hypersurfaces that we will simply refer to as t_1 and t_2 . Each Σ has a spatial boundary B , which in this paper will always have the intrinsic geometry of a sphere. The time history of the boundaries B will be called the three-boundary 3B .

The three-boundary 3B is the important boundary for the derivation of the quasilocal energy. In Brown and York's derivation, this boundary must be perpendicular to the time slices Σ . Besides simplifying the calculation, this is also an important physical point: the observers whose velocities are normal to the time slice are the observers that will measure the QLE. These observers should be at rest with respect to the boundary, and therefore the time slices should be perpendicular to the boundary. If they are not, then the observers will be boosted, and it is to be expected that they will measure a different value for the QLE. In this paper, we will relax the condition slightly but measure the same energy.

To make clear how the boundary term gives rise to the QLE, we will restrict our attention to only a simple class of metrics. Spherical symmetry is imposed so that we can easily embed a sphere with a given surface area in the four-dimensional metric. The metrics we study are static, by which we mean that they do not depend on the time coordinate, so the only nonzero derivatives are radial. Although the time coordinate becomes spacelike inside the horizon, for convenience we will call the entire metric static. While we would like to

generalize this derivation at a later date, this restricted version gives several interesting results and has the benefit of being easily understood.

A convenient form of the general spherically symmetric and static metric is

$$ds^2 = -\epsilon N(r)^2 dt^2 + \epsilon f(r)^{-2} dr^2 + r^2(d\theta^2 + \sin(\theta)^2 d\phi^2) \quad (2.19)$$

where ϵ is either 1 or -1 . When $\epsilon = 1$, g_{tt} is negative and t is a timelike coordinate. At a horizon, g_{tt} and g_{rr} exchange signs and so $\epsilon = -1$. N and f will be chosen to always be positive. This particular form of the metric is intended to make clear how the boundary terms behave at a horizon. The metric components $g_{tt} = -\epsilon N^2$ and $g_{rr} = \epsilon f^{-2}$ are the usual expressions of Schwarzschild, Reissner-Nordstrom, etc. in static coordinates.

To investigate the properties of the action, we will impose boundary conditions at some fixed r . The boundary term that we add is the one suitable to fix the induced metric on the three-boundary 3B . The induced metric in this simple case is

$$\gamma_{ij} dx^i dx^j = -\epsilon N(r)^2 dt^2 + r^2(d\theta^2 + \sin(\theta)^2 d\phi^2). \quad (2.20)$$

The action, with a boundary term added to fix the metric on the boundary, is

$$S = \frac{1}{2\kappa} \int_M d^4x \sqrt{-g} R - \frac{\epsilon}{\kappa} \int_{\partial M} d^3x \sqrt{|\gamma|} \Theta \quad (2.21)$$

where the extrinsic curvature is

$$\Theta_{ij} = -\frac{1}{2} \mathcal{L}_n \gamma_{ij} \quad (2.22)$$

and the trace Θ simplifies to

$$\Theta = \gamma^{ij} \Theta_{ij} = -\frac{f}{\sqrt{|\gamma|}} \partial_r (\sqrt{|\gamma|}). \quad (2.23)$$

The bulk term for the action using this ansatz for the metric is

$$S_{\text{bulk}} = \frac{1}{\kappa} \int d^4x \sin \theta \quad (2.24)$$

$$\times [N/f - \epsilon(Nf + 2rNf' + 2rfN' + r^2f'N' + r^2fN'')]]$$

where primes denote r derivatives. The boundary term is

$$S_{\text{BT}} = -\frac{\epsilon}{\kappa} \int d^3x \sin \theta (-f(Nr^2)'). \quad (2.25)$$

The boundary term can be converted into an integral over all 4 dimensions by also fixing the metric at $r = 0$ and integrating the derivative of the boundary term.

$$S_{\text{BT}} = \int d^3x (BT) = \int d^3x dr (BT)' + \int d^3x (BT) \Big|_{r=0}. \quad (2.26)$$

Adding the two terms shows that in this case, the action has a very simple form.

$$S = \frac{1}{\kappa} \int d^4x \sin(\theta) \left[\left(\frac{1}{f} + \epsilon f \right) N + (2\epsilon r f) N' \right]. \quad (2.27)$$

We have dropped the constant that comes from the second term in (2.26), because it will not affect the final result. Varying N yields

$$(2\epsilon r f)' = \left(\frac{1}{f} + \epsilon f \right), \quad (2.28)$$

and substituting back into the action shows that the action is an integral of a total derivative. Also doing the angular integrations (trivial because of spherical symmetry) gives

$$S = \frac{8\pi}{\kappa} \int (N dt) (\epsilon r f). \quad (2.29)$$

Following [1], we define the quasilocal energy as minus the second term in parentheses, so

$$E(r) = -(\epsilon r f). \quad (2.30)$$

The values for the Schwarzschild metric are

$$n(r) = f(r) = \sqrt{\epsilon(1 - 2M/r)}. \quad (2.31)$$

We also need to subtract the energy of flat space, which does not depend on ϵ ; we are embedding a sphere in flat space where there is no horizon. The subtraction term just has $k_0 = -\frac{2}{r}$. This will be the subtraction term used in the entire paper. The result for the QLE for any metric of the form considered is

$$E(r) = r(1 - \epsilon f(r)) \quad (2.32)$$

which reproduces the result in the previous section.

2.4 Coordinate Independence

We can relax the restriction on the form of the metric slightly, and consider what happens when the time coordinate is given an r dependence. The Brown-York derivation requires that the t and r coordinates be perpendicular at the boundary, a condition which is violated by this transformation. More general derivations of the QLE have been considered [19, 20, 21] where this condition is eliminated. We will not consider this issue in depth, but simply use the transformation to show that our version of the QLE is not coordinate dependent, for coordinate transformations of this type.

If we make the transformation $t = t(T, r)$, then we can write $dt = t_T dT + t_r dr$, where subscripts denote derivatives. The metric becomes

$$\begin{aligned} ds^2 &= -\epsilon N^2 (t_T dT + t_r dr)^2 + \epsilon f^{-2} dr + r^2 d\Omega^2 \\ &= -\epsilon (t_T N)^2 \left(dT + \frac{t_r}{t_T} dr \right)^2 + \epsilon f^{-2} dr + r^2 d\Omega^2 \end{aligned} \quad (2.33)$$

written in a 3+1 form where the foliations are hypersurfaces of constant r , which is appropriate for finding the induced metric on the three-boundary. Two good examples of this form are the Eddington-Finkelstein and Painlevé-Gullstrand coordinates. The action becomes

$$S = \frac{8\pi}{\kappa} \int_{\Gamma} (Nt_T dT + Nt_r dr)(\epsilon r f). \quad (2.34)$$

The integral is taken over a contour Γ which holds r fixed, so the final result is that the proper time Ndt has been transformed to $Nt_T dT$, the proper time written in the new coordinates. The QLE, $-\epsilon r f$, is *not* changed. We will not relate this quantity to the extrinsic curvature because it is not necessary to our point here.

As an example, consider a transformation $t(T, r) = T - r_*(r)$ where r_* is the tortoise coordinate defined by

$$\frac{dr_*(r)}{dr} = \frac{\epsilon}{Nf}. \quad (2.35)$$

This puts the metric into the form

$$ds^2 = -\epsilon N^2 dT^2 + 2\frac{N}{f} dT dr + r^2 d\Omega^2. \quad (2.36)$$

The appearance of ϵ in the tortoise coordinate may seem odd, but it reproduces the standard definition since N and f are always positive. In the case of the Schwarzschild metric, the result of this transformation is ingoing Eddington-Finkelstein coordinates [18] where T is the null coordinate. The hypersurfaces $T = \text{constant}$ are quite different from $t = \text{constant}$, but as shown above the QLE does not change. The choice we have made for the observers does not depend on the hypersurface.

The definition of the QLE is made of simple components. We embed a three-boundary in the space in such a way that at some moment of time, it has the

intrinsic geometry of a sphere with a specified area. The geometry of the sphere does not change when moving along the time coordinate of the three-boundary. The QLE is the change in the action with proper time, both of which are invariants. The invariance of the result under this particular type of coordinate change is therefore not surprising.

2.5 deSitter Space and Black Holes

The details of the derivation are not changed if a cosmological constant is added. The term added to the action is

$$S_{CC} = \int d^4x \sqrt{-g} (-2\Lambda) \quad (2.37)$$

which is proportional to N . When the equation of motion obtained by varying N is substituted back into the action, the same result as before is obtained but now $f(r)$ is different.

The metric is

$$N(r) = f(r) = \sqrt{\epsilon \left(1 - \frac{\Lambda r^2}{3} \right)}. \quad (2.38)$$

There is a cosmological horizon at $r = \sqrt{\frac{3}{\Lambda}}$. As expected, the energy continually grows with increasing r . The horizon forms when the energy inside the surface grows larger than

$$E(R) = \frac{c^4}{G} R. \quad (2.39)$$

This is the usual Schwarzschild radius expressed in a different form.

The deSitter-Schwarzschild solution has both a black hole horizon and a cosmological horizon. The black hole has a large amount of QLE inside a certain radius, but outside this radius the gravitational binding energy provides enough

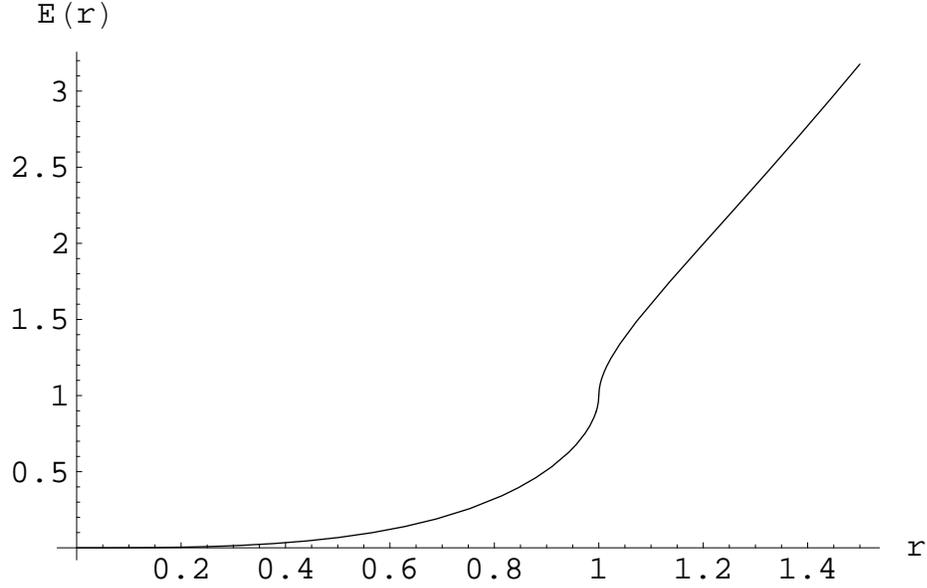


Figure 2.2: QLE of deSitter space (positive cosmological constant). Both axes are in the same units, with a mass scale proportional to $1/\sqrt{\Lambda}$.

shielding to bring the energy below $c^4 R/G$. The inner horizon forms at this radius. As one gets farther away, the cosmological constant begins to contribute noticeably. The cosmological horizon forms where the energy has once again climbed above the necessary value.

2.6 Reissner-Nordstrom

The Reissner-Nordstrom metric for a charged black hole behaves quite differently at the center from the uncharged case. The form of N and f is now

$$N(r) = f(r) = \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right)^{1/2} \quad (2.40)$$

where e is the charge of the black hole in natural units. There are now two horizons at $r_{\pm} = M \pm \sqrt{M^2 - e^2}$. The outer horizon is the same type as the

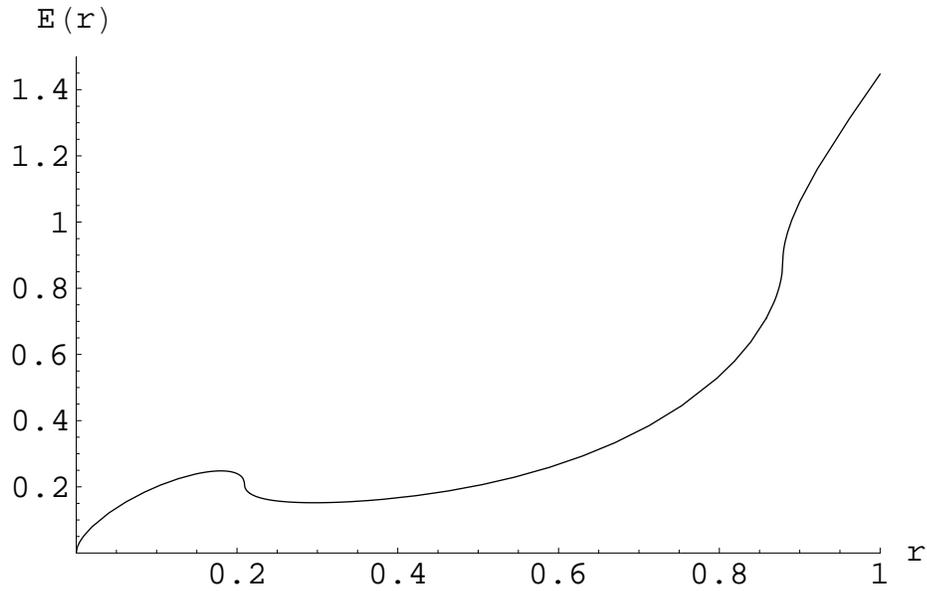


Figure 2.3: QLE of deSitter-Schwarzschild, a black hole in a spacetime with positive cosmological constant. The units are the same as the plot of the deSitter QLE, and the black hole has an unrealistically large mass so that details can be seen in the plot.

Schwarzschild horizon. The inner horizon exchanges the signature of the t and r coordinates again, such that t is a timelike direction. The consequence is that the singularity is now avoidable. The inward radial direction is spacelike and not timelike, and so particles are not inexorably drawn into the singularity.

Adding a new field into the theory will not change the definition of the QLE. The QLE only measures the gravitational energy, and so only the gravitational action is important. Of course, the addition of a new field changes the metric. One of the most important characteristics of the gravitational energy is universality. All mass-energy contributes to gravity, and so the QLE measures the energy of everything inside the surface (including the contributions purely from gravity).

The striking feature of this case is that the energy becomes negative within a certain radius. The QLE in either region where the time coordinate is timelike is

$$E(R) = R \left(1 - \sqrt{1 - \frac{2M}{R} + \frac{e^2}{R^2}} \right) \quad (2.41)$$

and so the energy becomes negative for $R < e^2/2M$. This is always inside the inner horizon. The energy at the singularity is $E(0) = -|e|$. The singularity has the electric field of a point charge, and so using just classical electromagnetism, the energy should diverge for small radius. However, the gravitational binding energy is negative, and while the cancellation is not perfect it seems that the binding energy at least makes the energy at the center finite.

The geodesics of massive neutral particles in the spacetime offer a probe of the effects of negative gravitational energy. The radial geodesics obey the equation

$$\dot{r}^2 + V(r) = p_0^2 - 1 ; \quad V(r) = (e^2 - 2Mr)/r^2 \quad (2.42)$$

where p_0 is the conserved energy per unit mass of the particle, and \dot{r} is the proper time derivative of r . A particle that starts from $r = \infty$ with zero velocity will not reach the center, but turn around at $r = e^2/2M$. Particles with higher energies will penetrate farther toward the center, but massive particles of all energies are repelled. This result is well known [22]. The turnaround radius agrees with the radius where the quasilocal energy becomes negative, so it seems that the two effects are very likely connected. Negative energy densities are expected to possess repulsive gravitational fields, and negative gravitational energy itself should be no exception.

Negative quasilocal energy has been ruled out by positivity theorems [2] which should apply in this case because the spacetime is static and the time coordinate timelike at the radius we are interested in. The energy condition which

is assumed for the theorem holds true throughout the Reissner-Nordstrom spacetime. However, the spacelike slice is assumed to be compact within the boundary where the QLE is defined. We conjecture that this is the condition that breaks down and causes the QLE to become negative.

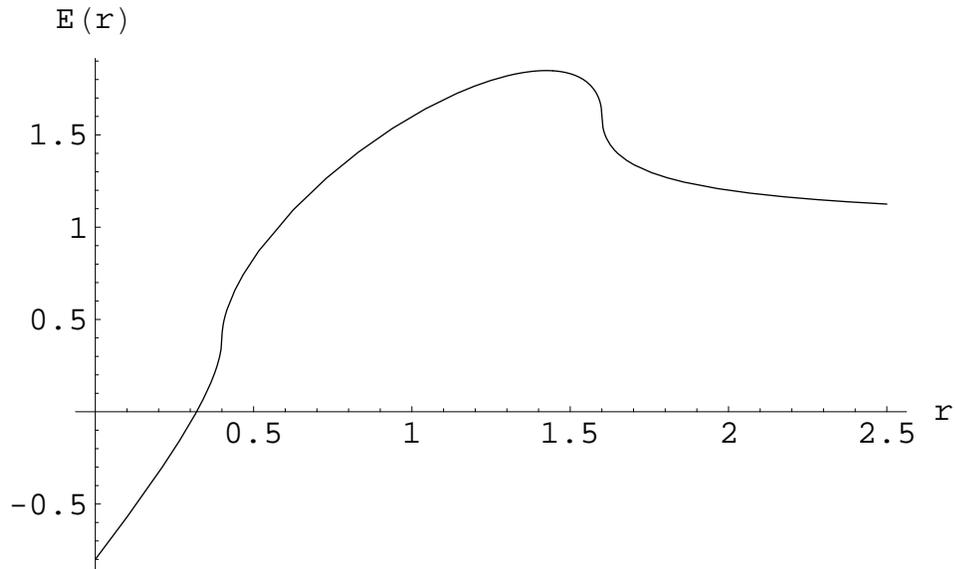


Figure 2.4: QLE of a Reissner-Nordstrom charged black hole. Both axes are in units of mass of the black hole, and the charge $e^2 = 0.8M^2$

2.7 Conclusion

We have shown that there is a sensible way to extend the definition of the quasilocal energy to surfaces inside an event horizon. The Schwarzschild singularity has zero energy, and the energy of the black hole mostly resides in a region just inside the horizon. The addition of a positive cosmological constant does not change these features but adds a cosmological horizon. In these cases, when the energy inside a given radius is less than r (in natural units), the space and

time coordinates play their usual roles. When the energy exceeds this quantity, a horizon forms and space and time switch roles. The derivative of the energy with respect to r at a horizon always seems to be infinite. These two features make it easy to locate the horizons on a plot of the QLE.

In a Reissner-Nordstrom black hole, the singularity at the center behaves like a point charge, and so there should be a divergence from the positive electric field energy. However, the contribution from the gravitational binding energy is negative and apparently cancels the divergence, rendering a finite energy at the singularity. The QLE is negative inside the radius $e^2/2M$, which is always inside the inner horizon of the Reissner-Nordstrom metric. A massive neutral particle released from rest at infinity will fall to the radius where the energy becomes negative, then reverse direction and be repelled. This provides the justification for the QLE's negativity, which is also related to the effect that clocks inside this radius run faster than those at asymptotic infinity.

In this paper, we have used a specific preferred set of observers to define the quasilocal energy. A direction for future research is to remove this restriction as in other work [19, 21, 20, 2] to define a more invariant quantity. We would like to extend this work to non-spherical boundaries and more general spacetimes. Two of interest are the spinning black hole metric and the metric for a star collapsing to form a black hole. There may also be applications to semiclassical and quantum gravity. Also, the issue of under what conditions the quasilocal energy is negative and what this means physically requires careful attention.

CHAPTER 3

CHARGED BLACK HOLE IN A CANONICAL ENSEMBLE

Originally published as A. Lundgren, *Phys. Rev. D* **77**, 042003 (2008). Copyright 2008 by the American Physical Society.

3.1 Introduction

A black hole in asymptotically flat space is thermodynamically unstable when the temperature at infinity is fixed. To solve this problem we can place the black hole inside a finite spherical cavity. The temperature is fixed at the surface of the cavity, which could be physically realized by placing a heat bath around the cavity. The black hole can now be thermodynamically stable, which is partly due to the fact that the horizon of the black hole can be near the point where the temperature is fixed. In this paper we are specifically interested in the thermodynamic ensemble where the cavity also contains a fixed amount of charge; this is an example of a canonical ensemble. The grand canonical ensemble, where the electric potential is fixed rather than the charge, was considered in [15], and we will use many of the methods developed there.

Many papers use the alternate method [23] of adding a negative cosmological constant to put the black hole in anti-de Sitter (AdS) space. The properties of AdS space stabilize the black hole by acting as a reflecting box. There has been much interest in the thermodynamics of anti-de Sitter space because of the conjectured AdS holography [24, 25, 26]. It appears that there is a duality between the thermodynamics of an AdS black hole and a field theory in one fewer

dimension. It is not clear whether the duality is a result specifically of the properties of anti-de Sitter space, or whether it is simply a result of the confinement provided by the reflecting walls.

We are therefore motivated to compare the results of our analysis with the results obtained using the AdS method, in particular the work by Chamblin et.al. [27, 28]. They compared the AdS charged black hole with the liquid-gas transition in classical thermodynamics, and demonstrated that the influence of the conserved electric charge yields a distinctive phase structure. This paper shows that the same phase structure exists when the black hole is in a finite box with no cosmological constant, which is a strong suggestion that the AdS holography is a result of confinement rather than the properties of AdS space. As we were finishing this paper, we discovered a paper by Carlip and Vaidya [29] with the same motivation. That work focused on the critical exponent of the second-order phase transition; we will focus more on the detailed phase structure.

The cavity we are considering has surface area $4\pi r_B^2$ and temperature T . The simplest case is when the charge inside the cavity is fixed to be zero. Then there is a minimum temperature $kT = \frac{3\sqrt{3}\hbar}{8\pi r_B}$ below which no black hole can be in equilibrium inside the cavity. Above this temperature, there are two states of the black hole that are possible. The larger one is quite close in size to the cavity; the cavity is within the $3M$ photon orbit of the black hole, where $2M$ is the Schwarzschild radius. This state is locally but not globally stable, and given enough time it will tunnel to the state with only radiation inside the cavity. The other state which is possible at the same temperature is a smaller, unstable black hole which corresponds to a maximum of the free energy. Therefore, it also sets

the height of the barrier that prevents tunneling to the state with only radiation and no black hole (hot flat space).

Above $kT = \frac{27\hbar}{32\pi r_B}$, the larger radius solution becomes globally stable, and now hot flat space can “decay” to form a black hole. There is still an unstable state of smaller radius that controls the height of the barrier between the stable states. As the temperature rises, the height of the barrier decreases, and the radius of the unstable black hole shrinks. At the same time, the size of the stable black hole solution is increasing until at infinite temperature it merges with the surface of the cavity.

In the canonical ensemble the charge of the black hole is fixed and the only variable allowed to fluctuate is the radius of the horizon. With zero charge, there were no equilibrium states other than flat space at low temperature, but there were two black hole states in addition to flat space at higher temperature. The structure becomes more complicated for non-zero charge. For charges less than $(\sqrt{5} - 2)r_B$ in natural units, there is a range of temperatures for which there are three possible black hole states. The largest and smallest are locally stable, and the one between must be unstable. For values of the temperature outside the range, or for any temperature when the charge is above the critical value, there is only one solution, which is always at least locally stable. One special case is when the temperature is very low, and so the black hole should have a very small radius. The danger is that the inner and outer horizons can merge, and yield an extremal black hole or a naked singularity (the extremal case is the dividing line between having a horizon or not). For any finite temperature, this does not happen, and the horizons are always separated.

We will begin by discussing the physical situation we are considering and

the boundary conditions necessary to implement the canonical ensemble. Then we derive the action and use it to find the possible equilibria given a certain temperature and charge. The uncharged and grand canonical ensembles are reviewed before looking in detail at the canonical ensemble. We find the number and stability of equilibria over the entire temperature - charge phase space, and compare to the AdS result. Some slices of the free energy function are displayed which are identical in structure to those of the AdS result. We end with a discussion of the meaning of these similarities and some directions for future research.

3.2 The Geometry and Action

We start with the usual form of the static spherically symmetric spacetime. The only free functions are b and α , which are both functions only of r . We analytically continue the time coordinate by defining $\tau = it$ to give a positive-definite metric. The metric takes the form

$$ds^2 = b^2 d\tau^2 + \alpha^2 dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad (3.1)$$

which is the same as in [15], except simplified slightly. Derivatives with respect to r will be denoted by primes. Throughout this paper, we will refer to a value of r as the radius, although the physical meaning of r is not the distance to the origin but instead gives the area of a sphere concentric with the origin. This is a quantity that can be measured without leaving the sphere, and is not dependent on the function α . Also we use natural units $k = c = G = \hbar = \frac{1}{4\pi\epsilon_0} = 1$, where mass, charge, and inverse temperature have units of distance.

The “Euclideanized” metric forces the τ coordinate to be periodic because

$b(r)$ shrinks to zero at the horizon and the $t - r$ plane behaves like the center of a disk there. The periodic imaginary time coordinate thermalizes the space (see [30, 31] for details). For convenience, we choose the period of τ to be 2π . In the Euclideanized action, inverse temperature at some radius is the period in the τ direction. The inverse temperature at a radius r is $2\pi b(r)$, which depends on r as a result of the gravitational redshift.

We define the thermodynamic ensemble by fixing the temperature at the outer boundary, and also fixing the electric field which serves to fix the charge inside. The coordinate position of the boundary will be $r = r_B$ and the area is $4\pi r_B^2$. The outer boundary is the full three-dimensional metric on the hypersurface $r = r_B$; we need the value of b to completely specify this metric. In terms of the reciprocal temperature $\beta = T^{-1}$, the boundary condition is

$$\int b(r_B) d\tau = 2\pi b(r_B) = \beta. \quad (3.2)$$

Note that b is analogous to the lapse N , which in the Schwarzschild metric is $\sqrt{1 - 2M/r}$. Approaching the horizon, this quantity decreases toward zero.

The inner boundary of the system is at the event horizon of the Euclidean black hole. The position will be denoted $r = r_+$. Since this is the horizon, we must have

$$b(r_+) = 0 \quad (3.3)$$

and since the $\tau - r$ part of the metric looks like a disc, we must avoid a conical singularity by requiring

$$(\alpha^{-1}b')|_{r_+} = 1. \quad (3.4)$$

The gravitational action is

$$I_g = -\frac{1}{16\pi} \int_M d^4x \sqrt{g} R + \frac{1}{8\pi} \int_{\partial M} d^3x \sqrt{\gamma} (K - K_0) \quad (3.5)$$

which is the usual Hilbert action with the metric fixed on the spatial boundary of the system ∂M . K is the extrinsic curvature of the timelike boundary of the system, which is the surface $r = \text{constant}$. K_0 is a subtraction term so that the action for flat space will be zero. It is calculated by embedding the same surface in flat spacetime. The results are

$$\sqrt{-g}R = -\frac{2b}{\alpha} + 2b\alpha - \frac{4rb'}{\alpha} - 4rb \left(\frac{1}{\alpha}\right)' - 2r^2 \left(\frac{b'}{\alpha}\right)' \quad (3.6)$$

$$K = -\frac{1}{\alpha br^2}(br^2)' \quad ; \quad K_0 = -\frac{2}{r}. \quad (3.7)$$

The K and K_0 terms are only present at the outer boundary, because that is where the metric is fixed. The quantities that are fixed are r and b ; α is free to vary because it involves the direction normal to the boundary. We need to integrate by parts any terms with second derivatives of b or first derivatives of α . The integration by parts in the bulk term yields total derivatives that cancel the boundary term involving K while depositing some terms on the inner boundary. The action, after integrating over angles and performing the integration by parts, is

$$I_g = -\pi \int_{r_+}^{r_B} \left(\left(\frac{2r}{\alpha}\right) b' + \left(\alpha + \frac{1}{\alpha}\right) b \right) dr + 2\pi br \Big|_{r_B} - 2\pi \frac{rb}{\alpha} \Big|_{r_+} - \pi r^2 \frac{b'}{\alpha} \Big|_{r_+}. \quad (3.8)$$

Of the three terms after the integral, the first is the remnant of the K_0 subtraction which makes the energy zero when $r_+ = 0$ which corresponds to flat space. The middle term is zero because $b(r_+) = 0$. Using the regularity condition (3.4) the last term becomes one-fourth the horizon area, which is the standard result for the entropy of a black hole. The entropy term is a direct result of the periodicity of the τ coordinate.

The electromagnetic action is very simple because we are only interested in a spherically-symmetric static electric field, so the only potential we need is

$A_\tau(r)$. The variation of A_τ in the Maxwell action gives the curved space version of Gauss's Law. As in [15], we can use this and the analytic continuation to a Euclidean metric to show that the charge e in the cavity is defined by

$$\frac{r^2}{b\alpha} A'_\tau = -ie. \quad (3.9)$$

The Maxwell action simplifies to (integrating over angles and τ)

$$I_{EM} = \pi \int dr \left(\frac{r^2}{\alpha b} A'_\tau \right) A'_\tau - 2\pi \left(\frac{r^2}{\alpha b} A'_\tau \right) A_\tau \Big|_{r_B} \quad (3.10)$$

where the second term is included so that we are fixing the quantity (3.9) on the boundary rather than the potential. On the inner boundary, we have another regularity condition. The potential in an orthonormal basis is $A_{\hat{\tau}} = b^{-1} A_\tau$. At the inner boundary, b goes to zero so we also fix $A_\tau = 0$ there so that the physical potential does not become unbounded. The form of the action is already suitable for fixing the potential on the inner boundary so we can now proceed to deriving the thermodynamics from the action.

We vary the action to obtain the equations of motion for b , α , and A_τ . The solutions depend on r_+ and e , and give the usual metric of a charged black hole. The reduced action I_* is defined as the value of the action evaluated for the solution we have obtained, which is

$$I_*(r_+; \beta, e) = \beta r_B \left(1 - \sqrt{\left(1 - \frac{r_+}{r_B} \right) \left(1 - \frac{e^2}{r_+ r_B} \right)} \right) - \pi r_+^2. \quad (3.11)$$

The last term is the entropy of the black hole, as was mentioned earlier. The first term is β times the quasilocal energy of the black hole [1] evaluated as the surface of the cavity. This gives the nice result that the free energy (which is I_*/β) is $E - TS$ with the quasilocal energy playing the role of E .

3.3 Equilibria and Stability

It is convenient to define non-dimensionalized variables using the radius of the boundary as a standard length:

$$I \equiv \frac{I_*}{4\pi r_B^2}, \quad x \equiv \frac{r_+}{r_B}, \quad q \equiv \frac{e}{r_B}, \quad \sigma \equiv \frac{\beta}{4\pi r_B}, \quad \Theta \equiv 4\pi r_B T. \quad (3.12)$$

Now the conditions for a physical solution are easy to write. We must have $0 < x < 1$, and $0 < q < x$, so that the horizon exists but is smaller than the outer boundary and larger than the inner horizon. The charged black hole solution has two horizons; when the Schwarzschild radius is equal to the charge the horizons merge to give an unphysical extremal Reissner-Nordström solution.

The action written in these new variables is

$$I(x; \sigma, q) = \sigma \left(1 - \sqrt{(1-x)\left(1 - \frac{q^2}{x}\right)} \right) - \frac{x^2}{4}. \quad (3.13)$$

Note that the physical action scales linearly with the area of the boundary.

The stationary points of the action will dominate the path integral. These stationary points are the equilibria of the system at the given temperature and charge. The only free variable is x , so stationary points are defined by

$$\frac{\partial I}{\partial x} = \frac{1}{2} \left[\frac{\sigma(x^2 - q^2)}{x^2 \sqrt{(1-x)\left(1 - \frac{q^2}{x}\right)}} - x \right] = 0. \quad (3.14)$$

The stability of an equilibrium point is determined by the curvature of the action. A negative second derivative indicates that the equilibrium is unstable; small fluctuations of x will tend to grow. For instance, if the black hole absorbs a small amount of excess heat it will grow and the amount of energy it radiates to the boundary will decrease causing the black hole to continue to absorb energy and grow, stopping only when it reaches a stable equilibrium or grows to the size of the boundary ($x = 1$).

A local minimum of the action is a stable equilibrium, for which small fluctuations do not grow. States like this may only be metastable if there are other states with lower action. There are two possibilities. The state with lowest action may also be an equilibrium, which we will then call globally stable. It is also possible that one of the edge cases has the lowest action, which most likely means that the the actual globally stable equilibrium is not a member of the ensemble we have defined. We will see that in the canonical action this does not happen; the edge cases never have the lowest action. Since we are only considering a subset of the possible metrics and neglecting the action of the radiation that should be present in the cavity, there may still be some other state with lower action.

We can make a single graph that shows the solutions and their stability at once. First we introduce the free energy

$$F = (E - TS) = I/(\beta). \quad (3.15)$$

The condition for a solution is that the derivative should be zero. This means that the temperature of a solution is the standard result

$$T(x_s) = \frac{E'}{S'}. \quad (3.16)$$

Now, treating the temperature as a function of x , we have

$$T' = \frac{E''}{S'} - \frac{E'S'}{(S')^2}. \quad (3.17)$$

and examining the second derivative of F yields

$$F'' = E'' - TS'' = S'T'. \quad (3.18)$$

T is actually fixed, so T' should not be thought of as the derivative of the actual temperature, but is merely the slope of the temperature graph with respect to

x . However, since S' is always positive in this case, we can simply plot the temperature as a function of x and determine the stability by looking at the graph.

3.4 Uncharged Case

Setting $q = 0$ and solving for the zeros of (3.14), we obtain the cubic equation

$$x^3 - x^2 + \sigma^2 = 0. \quad (3.19)$$

We can solve this for σ and substitute back into the equations for the action and its second derivative to determine the local and global stability of solutions. If we let x_e denote a value of x that is an equilibrium, we have:

$$\sigma = x_e \sqrt{1 - x_e} \quad (3.20)$$

$$I_s = x_e \left(\sqrt{1 - x_e} - 1 + \frac{3}{4} x_e \right) \quad (3.21)$$

$$\frac{\partial^2 I}{\partial x^2} = \frac{(3x_e - 2)}{4(1 - x_e)} \quad (3.22)$$

The lowest temperature (largest sigma) for which the action has a local minimum is $\sigma = \frac{2\sqrt{3}}{9}$. The size of the black hole is $x_s = \frac{2}{3}$ and so the cavity that encloses the black hole is located at the $3M$ photon orbit. A local minimum means that the black hole will be at least metastable. The solution will be globally stable if the action is a global minimum. There are no other local minima, but we need to check the action at the extreme values of x , which are $I(x = 0) = 0$ and $I(x = 1) = \sigma - \frac{1}{4}$. The solution we are considering has positive action, although it is less than that of the $x = 1$ solution. As a result, the black hole is stable, but eventually it will tunnel to hot flat space. It is clear that there must be a state

with maximum action between the local minimum and the $x = 0$ solution. This state is an unstable black hole with a smaller radius than the stable solution at the same temperature.

For global stability, the action has to be negative, which occurs at $x = \frac{8}{9}$ and $\sigma = \frac{8}{27}$. As the temperature is increased from this point, the radius of the black hole increases but the action of the black hole solution remains less than the action at the outer boundary ($x = 1$). At any temperature above this point, flat space is unstable and can tunnel to the black hole solution. There is always another solution with a smaller radius representing an unstable black hole, just as above. A detailed analysis of stability and phase transitions, including changes in topology, is given in [32].

3.5 Charged Case

The Reissner-Nordström metric of a charged black hole is

$$ds^2 = - \left(1 - \frac{2M}{r} + \frac{e^2}{r^2} \right) dt^2 + \left(1 - \frac{2M}{r} + \frac{e^2}{r^2} \right)^{-1} dr^2 + r^2 d\Omega^2 \quad (3.23)$$

with Lorentzian signature; with the $\tau = it$ substitution this is the metric that we would find from our equations of motion. The main new feature is the existence of two horizons, at $r_{\pm} = M \pm \sqrt{M^2 - e^2}$. It is convenient to make r_+ and e the independent variables because r_+ determines the entropy and e determines the electromagnetic field. If the charge grows large enough that $r_+ = e$, or $x = q$ in rescaled variables, the two horizons merge. This is the extremal Reissner-Nordström solution; if the charge is increased any more, the horizon disappears and the singularity is visible to the external universe. This super-extremal Reissner-Nordström solution is not desirable from a physical stand-

point, and we will show that the electric potential required to reach this solution is unrealistically high.

In the grand canonical ensemble, the equations giving a solution are:

$$(1 - \phi^2)x^3 - x^2 + (1 - \phi^2)^2\sigma^2 = 0 \quad (3.24)$$

$$q = \frac{\phi x^2}{\sigma(1 - \phi^2)}. \quad (3.25)$$

We now introduce a trick for solving these equations, which can also be somewhat useful in the canonical ensemble. We look for a solution with a given ratio of q and σ to x . Define

$$q \equiv \epsilon x, \quad \sigma \equiv bx. \quad (3.26)$$

The result is

$$x = 1 - \frac{1 - \epsilon^2}{\epsilon^2} \frac{\phi^2}{1 - \phi^2} \quad (3.27)$$

In the GCE, q only stands for the expectation value of the charge, and the potential ϕ is the quantity that is fixed. This result shows that when the expectation value of the charge coincides with the event horizon, then both must also coincide with the boundary, which means this is not a good physical solution. For $\phi < 1$, the non-extremal solutions have $x < 1$; they are physical and super-extremal solutions are not. The situation reverses for $\phi > 1$, but this corresponds to about 10^{27} volts (the Planck voltage). So for any reasonable values of the potential, the black hole is sub-extremal.

3.6 Canonical Ensemble

Turning to the canonical ensemble, the charge is fixed so we only vary x . We have only one equation to consider, which simply states that at a solution, de-

noted x_s , the action should be a local extremum. The condition for an equilibrium is therefore

$$\frac{\sigma(x_e^2 - q^2)}{x_e^2 \sqrt{(1 - x_e)(1 - \frac{q^2}{x_e})}} - x_e = 0 . \quad (3.28)$$

We start with the high-temperature behavior where σ goes to zero. The equilibria in that case are determined from $x_e^2(1 - x_e)(x_e - q^2) = 0$, which gives $x = 0, q^2, 1$. Of these, only $x = 1$ is a physically meaningful solution, and it indicates that the event horizon will tend to merge with the outer boundary at extremely high temperature.

Now we turn to the low-temperature behavior, where $\sigma \gg 1$. Without loss of generality we assume that q is positive. The first term in (3.28) will dominate, so the only solutions are near $x = q$. The second term must be negative, so actually we have $x > q$. This shows that the solution stays non-extremal for any non-vanishing temperature. Both the canonical and grand canonical ensembles will avoid the extremal case given physically reasonable conditions. Since the charge is fixed in this ensemble, we might have worried that as the temperature is lowered, the black hole will shrink without losing any charge, and the horizons will merge to give a naked singularity. However, the conserved quantity tends to improve the stability of the ensemble, and seems to prevent the black hole from decaying to a smaller radius.

At slightly higher temperatures, the entropy of the black hole makes more of a contribution to the free energy. It is now possible for larger black holes to form, because the energy that they have to gain from the heat bath is accompanied by an increase in the entropy. In fact, it is possible for the interplay between these effects to produce multiple solutions for a given temperature and charge. If there are multiple solutions, we must consider which ones are locally stable or

unstable.

The first case is when the charge is larger than a critical value. For any temperature there is a single solution, and it is always stable. The size of the black hole grows with the temperature, and it does not grow to the size of the outer boundary for any finite temperature. The solution is globally stable relative to the states we are considering, because it has lower free energy than the endpoints. It is possible that there are other states with the same charge and temperature but perhaps a different topology or non-spherical metric fluctuations. If these states have a lower free energy, than the charged black hole solution would be only metastable and could decay to them.

When the charge drops below the critical value, there exists a band of temperatures for which three states can coexist. At a given temperature, the smallest and largest radius equilibria are stable, and since the temperature graph must change slope between them, the intermediate radius solution is always unstable. Whichever solution has the lower free energy will be globally stable, because once again the endpoints have higher free energy. The intermediate unstable solution is a local maximum and so it sets the height of the barrier between the two stable states. As we approach the critical charge, the barrier becomes lower and the three states approach the same radius until they merge into one solution which must be marginally stable.

Outside the band of temperatures where there are three solutions, the behavior is the same as for larger charge. The low temperatures have nearly extremal black holes and the high temperatures have very large ones, and both are stable. Figure 3.1a shows the behavior below the critical charge, including the band of temperatures with three coexisting states. Figure 3.1b contrasts the behavior

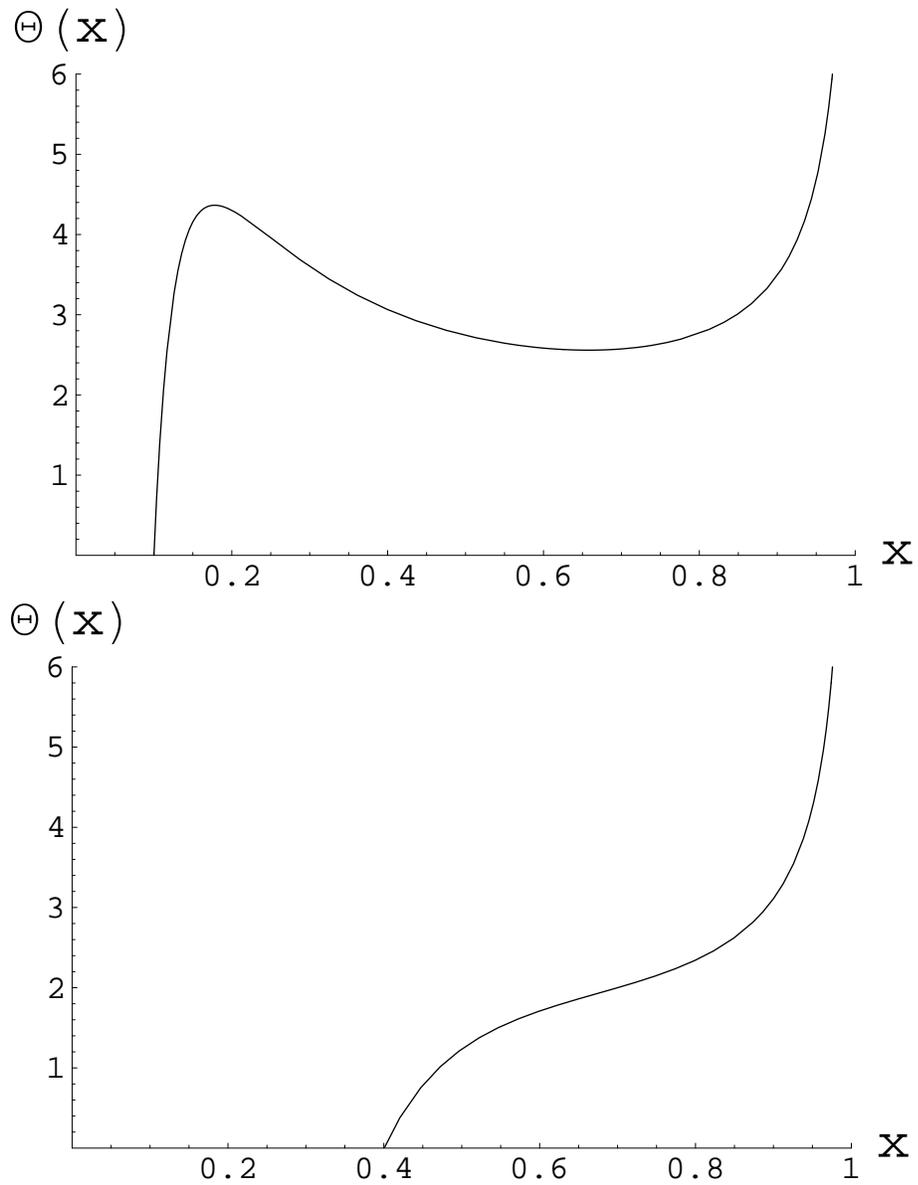


Figure 3.1: Temperature versus x for (a) $q = 0.1$ and (b) $q = 0.4$.

above the critical charge.

The temperature - charge phase space is divided into the region with one solution and the region with three coexisting solutions. The dividing line is formed by the states that are marginally stable, that is, the second derivative

of the free energy is zero. The transition from three solutions happens because the unstable solution merges with one of the stable solutions and annihilates it, leaving only a single solution. The dividing line can be found by solving for the position where the temperature graph has zero slope. It is easiest to find the charge q as a function of x for which $\partial\Theta/\partial x = 0$ which yields

$$q_{transition} = \left[\frac{x}{6x-5} \left(x^2 + 3x - 3 + \sqrt{(x-9)(x-1)^3} \right) \right]^{1/2}. \quad (3.29)$$

In Figure 3.2 the two regions in the $q - \Theta$ plane are plotted. The critical charge above which there is never more one solution is $q_{crit} = \sqrt{5} - 2$.

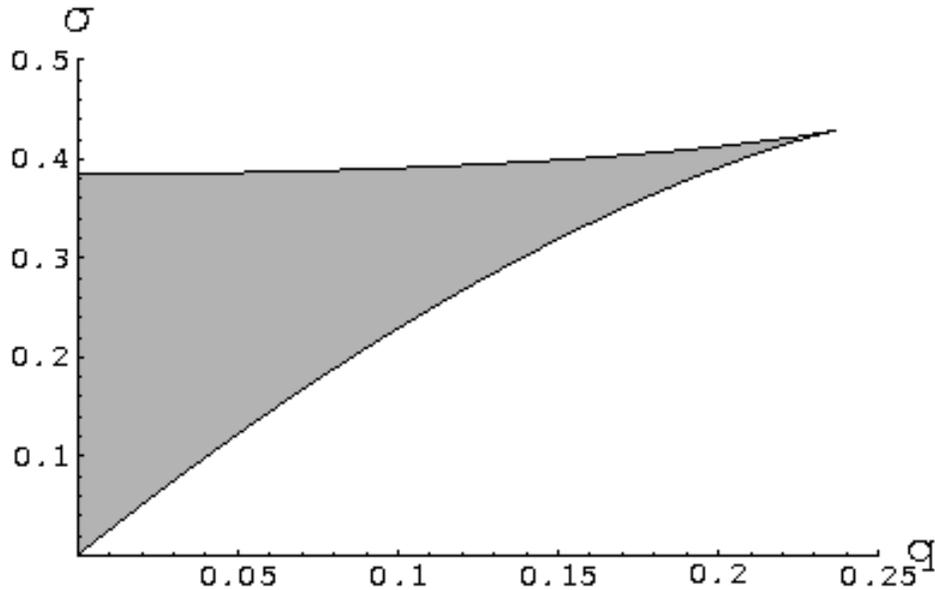


Figure 3.2: Phase diagram with charge q on the horizontal axis and temperature T on the vertical. The shaded region has three equilibria as opposed to only one outside.

We can now address the issue of global stability which is simpler here than in the uncharged case. The most stable configuration could be any of the local minima, or one of the endpoints of the range where the action is defined. When $x = q$ the horizons merge and we have a different topology, and when $x = 1$ the

horizon merges with the boundary of the system. In either case the action is no longer defined. However, one of the stable equilibria must be the global minimum for all values of temperature and charge. Throughout the phase space, there are always either two minima of the action with a maximum between them, or a unique minimum. The continuity of the action then guarantees that one of the equilibria will always be the global minimum of the action, over the range $q < x < 1$ where the action is considered to be physically meaningful. There are likely to be other states outside those which we have considered in our ensemble, particularly those where a gas of charged particles has replaced the black hole. We will discuss this further in the conclusion.

To make the phase structure more clear, we can plot the free energy F against the temperature with the charge fixed. When there are three equilibria, this will allow us to see which one is globally stable. Figure 3.3 shows that above the critical charge, there is only one branch of the free energy. As we raise the temperature, the black hole smoothly grows from a small near-extremal state to a large black hole that nearly fills the cavity. Below the critical charge, the smooth transition from small to large black holes is interrupted in the band of intermediate temperatures where there are three equilibria. As the temperature increases in this band, the free energy of the small black hole increases while that of the large one decreases. They cross each other and the larger black hole becomes the globally stable one. As the temperature increases past this point, the smaller black hole is still locally stable, so it could persist for a long time. However, the free energy of this state and the unstable state are growing closer in value, so as the temperature increases there is less of a barrier preventing a transition to the globally stable state. At some point, the branch of the free energy that we are on merges with the branch for the unstable state, and ends. There is now only the

larger black hole state, and this will continue to grow and be the only possible equilibrium as the temperature increases to infinity.

At the temperature where the free energies of the two stable equilibria are equal, the two phases (small black hole and large black hole) can coexist. This is a slightly dangerous viewpoint, because we are only considering a single black hole concentric with the cavity; nothing in the analysis shows that two different-size black holes could coexist. We can think of the black hole of being in a superposition of the two states, but they are macroscopic and of very different masses so interactions with the environment would cause decoherence. The best viewpoint to take is that the two states are equally probable.

Figure 3.4 graphs the free energy versus the charge for fixed values of the temperature. The shapes, but not the values, can be compared with Figure 5 in [28]. Figure 3.3 can be compared with their Figure 4. These are simply a more visual way of demonstrating that the detailed phase structure of the AdS charged black hole and the charged black hole in a cavity are essentially identical.

3.7 Conclusion

We have derived the thermodynamics of a charged black hole in a finite spherical cavity. The charge governs the number of possible equilibria. Above the critical value $q = (\sqrt{5} - 2)r_B$, there is always a single equilibrium state, which is a small black hole with a Schwarzschild radius only slightly larger than its charge in natural units. As the temperature increases the size of the black hole increases, but for any finite temperature above zero, we always have physically reasonable behavior; the horizon never disappears as in the case of an extremal

Reissner-Nordström solution nor does the horizon grow to touch the walls of the enclosing cavity.

For charges below the critical value, there is a band of temperatures where there are three different equilibria. Figure 3.2 shows the region of the temperature-charge phase space where this is the case. One of the equilibria is always unstable and so it is not directly relevant to the thermodynamics. The behavior in this temperature band is like a phase transition, from the smaller black hole state to the larger one. At the critical temperature and charge, there is a second-order phase transition beyond which there is only one phase (one stable equilibrium rather than two stable and one unstable). In [28], the critical exponent was derived for this transition in the AdS case, and in [29] it was shown to be the same for the finite cavity. The current paper has shown in more detail that the phase structure is the same. It seems likely, therefore, that AdS holography does not depend on specific properties of AdS space but instead simply on the confinement of a reflecting box or cavity at a fixed temperature.

We have not discussed the issue of hot flat space in this paper, or the possibility of other topologies that may contribute to the ensemble [32]. When we showed that one of the equilibria is always a global minimum of the action, that only counts the states that we considered in deriving the action. More exotic topologies could possibly have a smaller value of the action, which would allow our “globally stable” state to decay to the more exotic topology.

A more important and possibly more tractable concern is the Hawking-Page [23] transition to hot flat space. With the charge fixed to zero, this is likely to happen because flat space has a lower value of the action than the black hole unless the temperature is high enough. We have the problem that there is no

solution for flat space with a nonzero charge. In reality, the black hole can emit charged Hawking radiation, and possibly evaporate so that the box is filled with a charged gas of electrons or other particles. More work is needed to understand charged thermal gasses in this context. One direction for future research is to add charged fields to the action and determine how they affect the thermodynamics. It seems that fixing the potential at the boundary (grand canonical ensemble) rather than the charge inside makes more physical sense when the black hole can emit charged particles, but the canonical ensemble may still have a role to play.

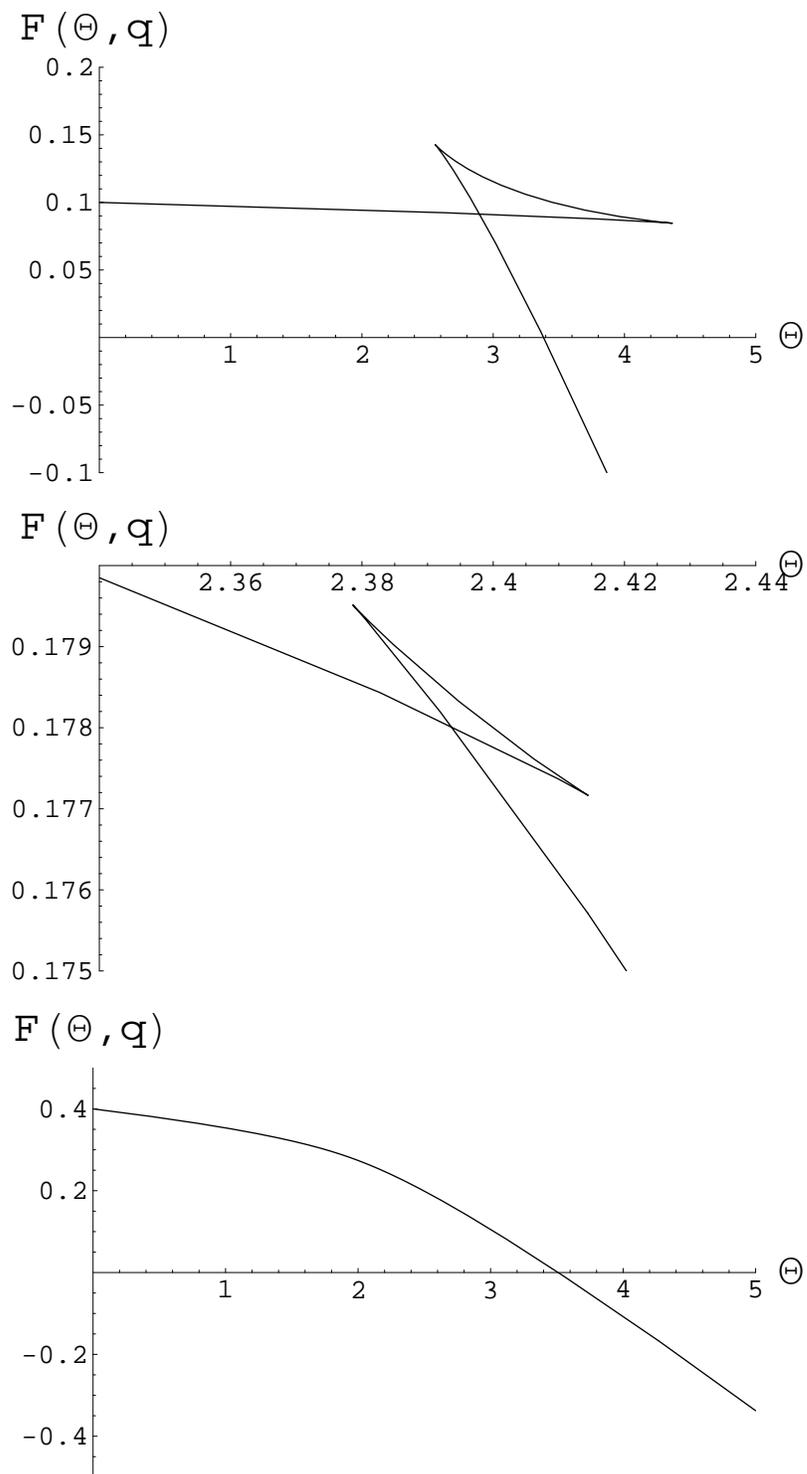


Figure 3.3: Free energy versus temperature for (a) $q = 0.1$, (b) $q = 0.22 \approx q_{crit}$, and (c) $q = 0.4$.

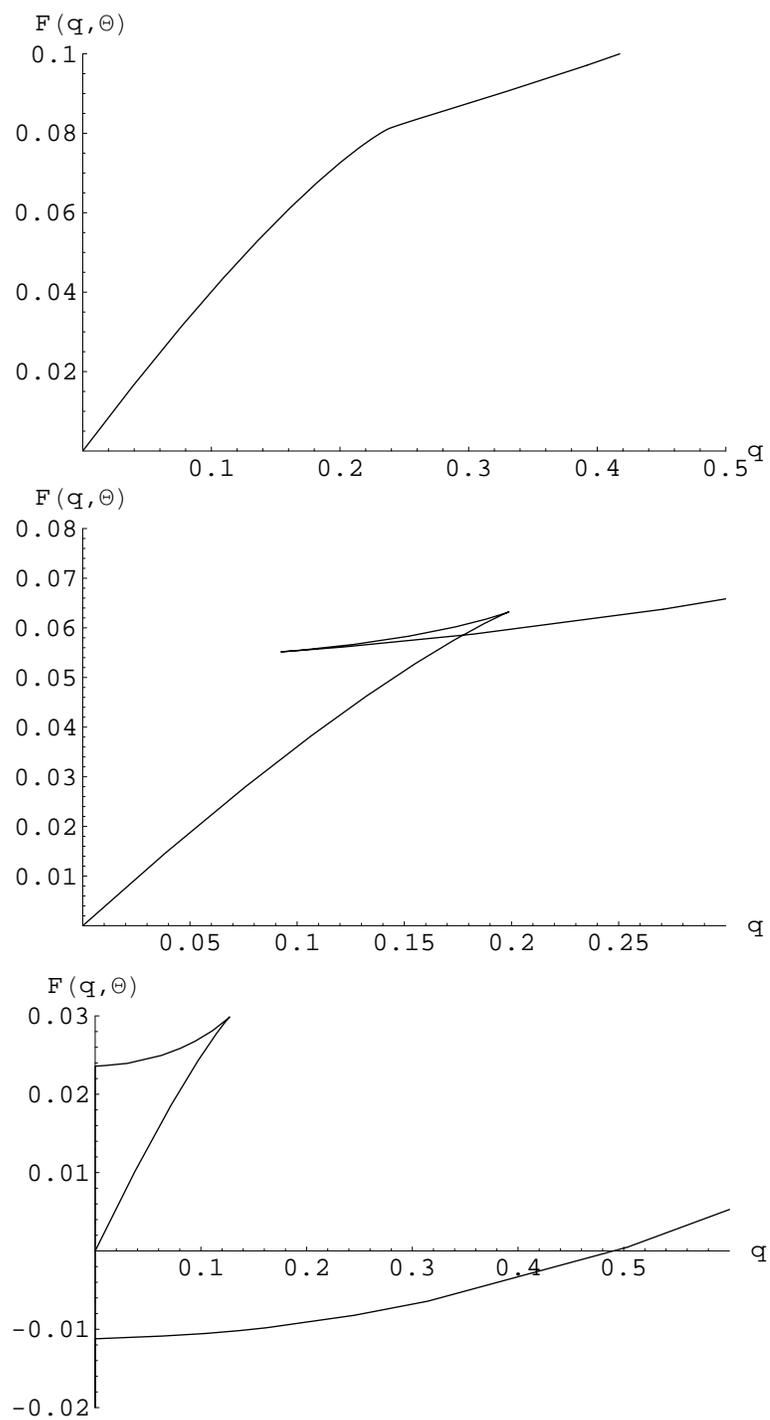


Figure 3.4: Free energy versus charge for (a) $\theta = 2.33$, (b) $\theta = 2.56$, and (c) $\theta = 3.57$.

CHAPTER 4
FINITE MIRROR EFFECTS IN ADVANCED INTERFEROMETRIC
GRAVITATIONAL WAVE DETECTORS

Originally published as A. Lundgren, R. Bondarescu, D. Tsang, and M. Bondarescu, Phys. Rev. D **77**, 042003 (2008). Copyright 2008 by the American Physical Society.

4.1 Introduction

The initial baseline design for the Advanced LIGO gravitational wave detectors [34, 33] employs Gaussian beams in the arm cavities. The leading noise source in the most sensitive frequency band of the instruments ($\sim 30 - 300$ Hz) is the thermal noise in the substrate and reflective coating of the mirror test masses. Lowering thermal noise is therefore of paramount importance for achieving a higher event rate in LIGO. There are a number of other detectors that are being built or upgraded to similar specifications. While we will choose to study Advanced LIGO for definiteness, our general conclusions should be more widely applicable to any interferometric detector that needs to limit thermal noise. Some of the important parameters that we use are summarized in Table 4.1.

LIGO is a Fabry-Perot interferometer with four mirrored test masses. The resonant beams in the cavity measure the position of the test masses, averaging over the mirrored surface, with the average weighted by the power distribution of the beam. Thus, the highly illuminated central area is weighted more than the mirror boundary that is left nearly dark. One way of decreasing the thermal

noise is to flatten the beam so that a larger fraction of the mirror is in use. Motivated by this intuitive reasoning, O’Shaughnessy et. al. [35, 36] proposed the flat topped Mesa beams, which were subsequently explored in detail by them and others [37, 38, 40, 39]. These beams would lower the thermal noise by a factor of approximately 2.5 compared to the baseline design. The original Mesa beam supported by nearly-flat Mexican Hat mirrors was found to be susceptible to a tilt instability [41]. This triggered the proposal of a Mesa beam supported by nearly-concentric mirrors [42]. In the same paper, a family of hyperboloidal beams that include all Mesa and Gaussian beams previously considered was introduced. Mesa is currently the leading alternative beam design for Advanced LIGO, and is being studied experimentally [43, 44].

In this paper we first discuss the formulation of hyperboloidal beams. The “nearly-flat” Mesa is created by superposing minimal Gaussians with generators uniformly distributed inside a cylinder, and the “nearly-concentric” by generators falling inside a cone, and passing through the center of the cavity. These two choices have the same intensity distribution on the mirrors, but the second has a much smaller susceptibility to tilt instability. The hyperboloidal beams smoothly interpolate between these two cases by twisting the generators of the minimal Gaussians by an angle α . After discussing some geometric properties of the beams, we present a proof that Gaussian beams are a special case of the hyperboloidal beams, confirming a conjecture in [42].

We then compute the three types of mirror thermal noise for a variety of hyperboloidal beam shapes, using a set of simple scaling laws developed in [45, 46] that simplify previous work [47, 48, 49]. The first is substrate Brownian noise, occurring due to mechanical dissipation in the material; this is the least

significant source of thermal noise. The substrate thermoelastic noise is caused by random thermal expansion. The coating also has both Brownian and thermoelastic noise, but these follow the same scaling laws so we consider them as a single source. The coating noise is the most severe of the three types for the fused-silica substrates currently planned for Advanced LIGO. The substrate thermoelastic noise would dominate in a material like sapphire which has a higher thermal expansion coefficient. We show that the noises decrease with increasing width of the beam, as expected, and that the hyperboloidal beams have larger noise than the relevant Mesa beams.

The constraint on our ability to lower the noise comes from the need to keep the diffraction loss small. Gravitational-wave interferometers must keep a large circulating power in the cavity, and so cannot allow significant amounts of light to escape past the edge of the mirrors. The current design constraint used in the most recent papers [50, 51, 40] is a diffraction loss of 1 part per million (ppm) per bounce for 17 cm fused-silica mirrors. The clipping approximation Eq. (4.19) indicates that the desired Mesa width is approximately $D = 10$ cm. Previous work [52, 39] have shown that the clipping approximation is not accurate for Gaussian beams of finite mirrors, and have used Fast Fourier transform simulations for accurate calculations.

We calculate the diffraction losses accurately from eigenvalues of the cavity propagator using an exponentially convergent code that enables us to study the structure of Mesa and other hyperboloidal beams as a function of D , mirror radius, and twist angle α in detail. We find that the diffraction loss is not a monotonic function of D , but due to finite mirror effects has anomalous local minima where the loss is significantly below what is expected from the clip-

ping approximation. These minima are observed to become more shallow and eventually disappear as the radius of the mirror is increased. However, for the mirror radii and beam widths relevant for Advanced LIGO the finite mirror effects are important. We show that they can be used to increase the width of the beam, lowering the noise even further than previous work.

Finally, we develop an iteration scheme to redesign the mirror, explicitly accounting for finite mirror effects. The iterated mirror is altered to match the phasefront of the primary eigenbeam of the finite mirror cavity, reducing the diffraction loss of this mode, thus allowing even larger beam widths to satisfy the diffraction loss constraint.

The mathematical construction of the hyperboloidal beams are discussed in Sec. 4.2, while the asymptotic limit of the hyperboloidal beams are derived in Sec. 4.3. The thermal noise scaling laws are described in Sec. 4.4. The cavity propagator construction and eigenmode decomposition are presented in Sec. 4.5, with the results, including finite mirror effects, discussed in Sec 4.6. We then summarize our work in Sec. 4.7.

Table 4.1: Advanced LIGO Parameters

L	3999.01 m	Length of LIGO cavity
λ_0	1.064×10^{-6} m	Laser wavelength
w_o	$\sqrt{\lambda_0 L / 2\pi} = 2.6023$ cm	Minimal Gaussian width
R	17 cm	Mirror radius

4.2 Construction of the Beams

The beams we study are supported by two identical mirrors facing each other, forming a cavity of length L . The mirrors are cylindrically symmetric around the optical axis, which runs along the length of the cavity and will be called the z axis. The center of the cavity, equidistant between the mirrors, is $z = 0$; the mirrors are located at $z = -z_R$ and $z = z_R$, where $z_R = L/2$. The transverse distance from the z axis will be denoted by r , and the angular coordinate by ϕ . The geometry is shown in Fig. 4.1.

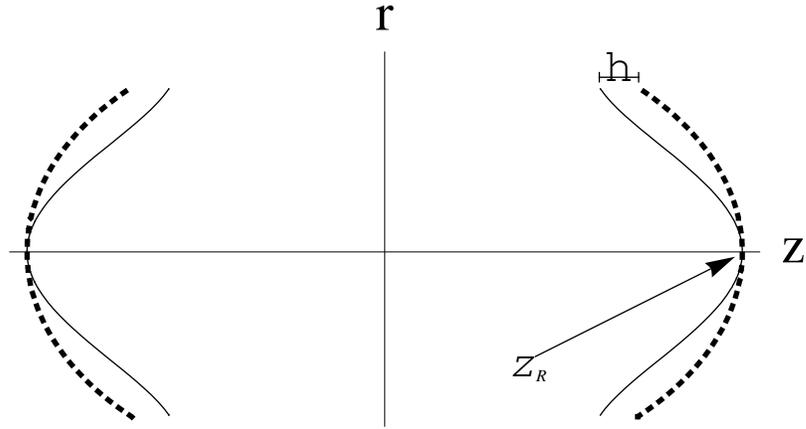


Figure 4.1: The cylindrical coordinate system that we use, with the ϕ coordinate suppressed. The dotted lines are the fiducial spheroids, while a mirror is shown schematically, with $h(r)$ indicated.

The cavity is fed with laser light with wavelength λ_0 , and the distance between the mirrors is fine-tuned so that the cavity resonates in its fundamental mode, with a field amplitude $U(r, z)$ and intensity $|U(r, z)|^2$. In this paper we focus on axisymmetric modes with no ϕ dependence in the beams. Non-axisymmetric modes are important for studies of the tilt and parametric instabilities [38, 53], but are not discussed in detail in this work. The narrowest possible Gaussian mode that can exist in a cavity of given length is called the minimal

Gaussian, which has the intensity distribution

$$|U(r, z)|^2 = \frac{2}{1 + (z^2/z_R^2)} \exp\left(-\frac{2r^2}{w_0^2(1 + (z^2/z_R^2))}\right), \quad (4.1)$$

where $w_0 = \sqrt{\lambda_0 z_R/\pi}$.

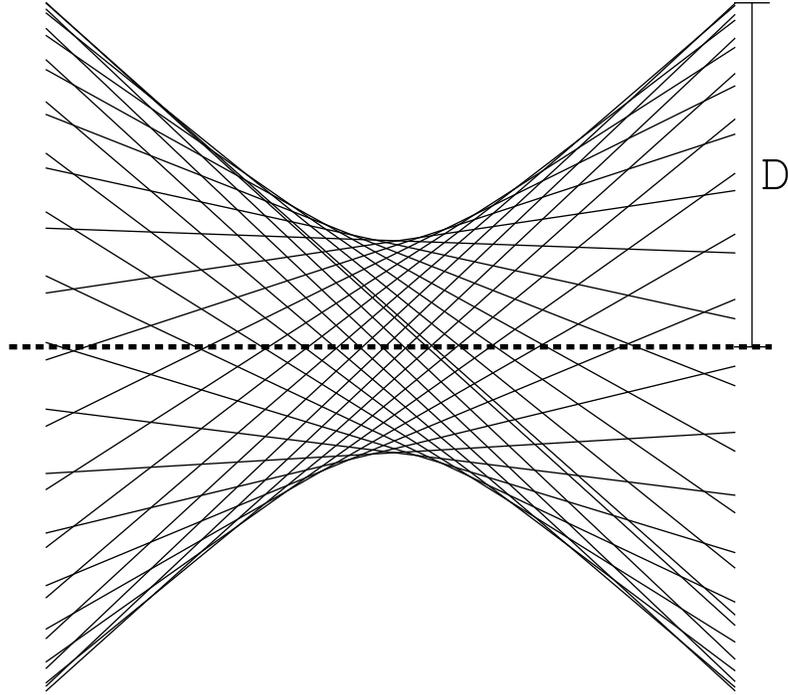


Figure 4.2: The generators of the minimal Gaussians that are superposed to produce the $\alpha = 0.8\pi$ hyperboloidal beam. For clarity, only the outermost set of generators are displayed.

A hyperboloidal beam is the superposition of minimal Gaussians chosen such that the symmetry axis of the individual minimal Gaussians are generators of a set of coaxial hyperboloids. An example set of generators is shown in Fig. 4.2. The beam family has two parameters: α , the twist angle one would have to rotate the two basis of a set of coaxial cylinders with respect to each other to obtain the hyperboloids and D , the radius of a section perpendicular to the optic axis of the outermost hyperboloid at the end of the cavity. In the case

$\alpha = 0$, the propagation axes are parallel and fill a cylinder of radius D . This is the Mesa beam supported by nearly-flat mirrors. For $\alpha = \pi$ the lines all cross at $z = 0$ forming two cones. This configuration also generates a Mesa beam, but one supported by nearly-concentric mirrors. Varying α smoothly deforms the beam and the mirror shape between the two configurations. Some examples of the beam shape

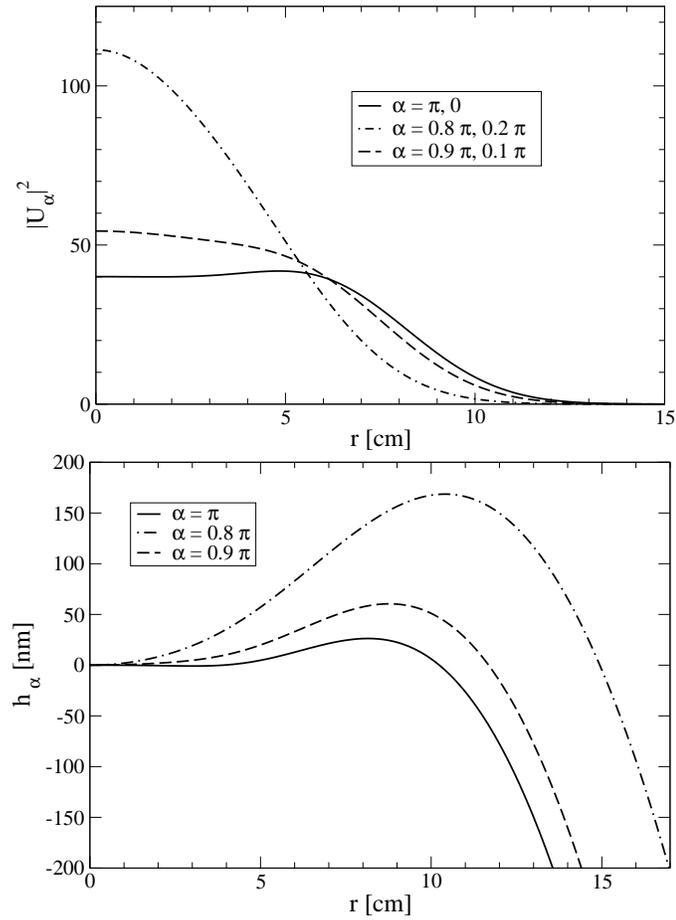


Figure 4.3: (a) The beam intensity profile $|U_\alpha|^2$ and (b) corrections h_α are shown at fixed $D = 10$ cm for different twist parameters α .

For the cavity to support the desired beams, the phase of the electric field of the beam should be constant on the mirror surface. We will focus our attention on the mirror on the positive z side of the cavity. The wavefront can be

approximated by the “fiducial spheroid”,

$$z = S_\alpha(r) = \sqrt{z_R^2 - r^2 \sin^2(\alpha/2)}. \quad (4.2)$$

For $\alpha = 0$ this is the mirror plane $z = z_R$; for $\alpha = \pi$, the fiducial spheroid is a sphere centered on $z = 0$, and clearly the lines which generate the hyperboloid are all radii of the sphere.

There are two equivalent expressions for the field amplitude evaluated on the fiducial spheroid. The first is the integral expression [42]

$$U_\alpha(r, S_\alpha) = \Lambda \int_0^{R_0} dr_0 \int_0^{2\pi} d\phi_0 r_0 \exp \left[i \frac{rr_0}{w_0^2} \sin \phi_0 \sin \alpha - \frac{(r^2 + r_0^2 - 2rr_0 \cos \phi_0)}{2w_0^2} (1 - i \cos \alpha) \right], \quad (4.3)$$

where Λ is a complex constant.

The second is the method that we use in this paper. The beam is constructed as shown in [54] by an expansion in Gause-Laguerre eigenbeams of spherical mirrors. They are closely related to the Gauss-Laguerre basis functions given by

$$\psi_m(\xi) = \sqrt{2} \exp(-\xi^2/2) L_m(\xi^2), \quad (4.4)$$

where L_m is the m^{th} Laguerre polynomial; these satisfy

$$\int_0^\infty \psi_m(\xi) \psi_n(\xi) \xi d\xi = \delta_{mn}. \quad (4.5)$$

Then the Gauss-Laguerre eigenbeams are

$$\Psi_m(r, z) = \frac{w_0}{w(z)} \psi_m \left[\frac{\sqrt{2}r}{w(z)} \right] \exp \left[\frac{ik_0 r^2}{2R(z)} \right] \times \exp[i(k_0 z - (2m + 1)\Phi(z))] \quad (4.6)$$

where

$$w(z) = w_0 \sqrt{1 + (z/z_R)^2}, \quad R(z) = z + z_R^2/z, \quad (4.7)$$

$$\Phi(z) = \arctan(z/z_R)$$

and $k_0 = 2\pi/\lambda_0$. The expansion is written as

$$U_\alpha(r, z) = \sum_{m=0}^{\infty} A_m^{(\alpha)} \Psi_m(r, z). \quad (4.8)$$

The expansion coefficients that result in hyperboloidal beams are

$$A_m^{(\alpha)} = (-\cos \alpha)^m \frac{\sqrt{2}w_0^2}{D^2} P\left(m+1, \frac{D^2}{2w_0^2}\right). \quad (4.9)$$

$P(a, x)$ is the incomplete gamma function

$$\begin{aligned} P(a, x) &= \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt \\ &= \frac{\int_0^x e^{-t} t^{a-1} dt}{\int_0^\infty e^{-t} t^{a-1} dt}. \end{aligned} \quad (4.10)$$

The mirror shape that supports a hyperboloidal beam is not exactly the fiducial spheroid. We make a correction $h(r)$ so that the surface of the mirror is given by $z_M(r) = S_\alpha(r) - h(r)$. The correction is chosen so that the mirror is located at a phasefront of the beam. We find the mirror surface from $U_\alpha(r, S_\alpha)$ by

$$h(r) = k_0^{-1}(U_\alpha(r, S_\alpha) - U_\alpha(0, S_\alpha)). \quad (4.11)$$

The shape of $h(r)$, or the mirror itself, is generally referred to as a ‘‘Mexican hat’’, and some examples are displayed in Fig. 4.3b.

As expected, the beams for α and $-\alpha$ are identical, as they correspond to hyperboloids that are simply twisted in opposite directions. There is a duality between α and $\pi - \alpha$. This was first mentioned in [42] for all α 's, then studied in more depth for $\alpha = 0$ and $\alpha = \pi$ in [38] and, finally, analytically understood in [37]. This duality extends to several quantities. The beam intensity profiles are identical. The corrections to the mirror shape are opposite, $h_\alpha(r) = -h_{\pi-\alpha}$. There are also dualities in the eigenvalues of the propagator [38, 37, 54, 42] that we will not discuss in this paper.

4.3 Asymptotic Behavior of Wide Hyperboloidal Beams

It was conjectured by Bondarescu and Thorne [42] that the beam becomes a Gaussian in the limit $D \rightarrow \infty$. We will prove this analytically for the intensity profile of the beam, evaluated on the plane $z = z_R$ which would be the surface of a perfectly flat mirror. The intensity varies slowly enough with z that this will also be the intensity profile on the mirror to a good approximation. Our proof uses the expression for the beam amplitude in terms of a summation of Gauss-Laguerre functions. The essential ingredient is the realization that the expansion coefficients take the form $A_m^{(\alpha)} = (\text{constant})^m$ as $D \rightarrow \infty$, where the constant depends only on α . This allows us to analytically perform the summation to obtain the beam profile.

In the limit $x \rightarrow \infty$, the incomplete gamma function $P(a, x) = 1$, giving $A_m = \sqrt{2}w_0^2(-\cos \alpha)^m/D^2$. The approximation $z = z_R$ yields

$$w(z_R) = \sqrt{2}w_0, \quad R(z_R) = 2z_R, \quad \Phi(z_R) = \pi/4, \quad (4.12)$$

and the Gauss-Laguerre propagators become

$$\Psi_m(\bar{r}, z_R) = \frac{1}{\sqrt{2}}\psi_m(\bar{r})e^{-i\frac{\pi}{2}m}e^{i\phi(\bar{r})}, \quad (4.13)$$

where $\bar{r} = r/w_0$ and the r -dependent part of the phase has been absorbed into $\phi(\bar{r})$. Since in the end we will only be interested in the intensity profile, the exact form of $\phi(\bar{r})$ is unimportant.

The expansion for U_α then becomes

$$\begin{aligned} U_\alpha(\bar{r}, z_R) &= \sum_{m=0}^{\infty} (-\cos \alpha)^m \frac{w_0^2}{D^2} \left(e^{-i\frac{\pi}{2}}\right)^m e^{-\frac{\bar{r}^2}{2}} L_m(\bar{r}^2) e^{i\phi(\bar{r})} \\ &= \left(\frac{w_0}{D}\right)^2 e^{-\frac{\bar{r}^2}{2}} e^{i\phi(\bar{r})} \sum_{m=0}^{\infty} (i \cos \alpha)^m L_m(\bar{r}^2). \end{aligned} \quad (4.14)$$

We now use the generating function for the Laguerre polynomials [55]

$$\sum_{m=0}^{\infty} L_m(x)t^m = \frac{1}{1-t} \exp\left(-\frac{t}{1-t}x\right) \quad (4.15)$$

to evaluate the sum in (4.14), with $t = i \cos \alpha$. The final result is an intensity profile

$$|U_\alpha(r)|^2 = \frac{1}{\pi w_0^2} \exp\left[-\frac{r^2}{\sigma^2}\right], \quad \sigma = \frac{w_0 \sqrt{1 + \cos^2 \alpha}}{\sin \alpha}. \quad (4.16)$$

The minimal Gaussian $\alpha = \pi/2$ is seen to have $\sigma = w_0$. The width is symmetric under $\alpha \rightarrow \pi - \alpha$, as expected from the duality relation [42, 37], and goes to infinity at $\alpha = 0$ or π . This includes every Gaussian beam capable of resonating in a cavity of the given length.

4.4 Thermal Noises

There are a number of noise sources limiting the sensitivity of ground-based gravitational-wave interferometers. Seismic noise causes an effective cutoff in the lowest frequencies that can be measured. Fundamental problems such as shot noise and radiation pressure noise, as well as technical issues, are important limitations on the sensitivity throughout the frequency band. However, the major contribution in the most sensitive frequencies of LIGO is the thermal noise in the mirrors. Reducing the thermal noise is the goal of this paper.

The mirror consists of a substrate with a coating, and we must consider noises due to fluctuations of both. The substrate and coating have both thermoelastic and Brownian contributions to the noise. Thermoelastic noise is caused by expansions in the material caused by random heat flow. Brownian noise is due to the coupling of normal modes of vibration by imperfections in the material.

As a practical matter, both types of noises in the coating have the same scaling law so they do not have to be considered separately. In fact, for the fused silica mirrors now under consideration for Advanced LIGO, the coating noises are the dominant contribution. However, use of a material like sapphire, with a higher coefficient of thermal expansion, would cause the substrate thermoelastic noise to dominate. We will calculate all three types of noise in this paper.

A set of simple scaling laws were derived by Lovelace [45] in parallel with O'Shaughnessy [46] that are applicable to beams of arbitrary shape. The noise is proportional to an integral depending on the intensity and an overall constant which is independent of the shape of the beam. The noises are given by

$$S_n = C_n \int_0^\infty \tilde{I}(k)^2 k^n dk , \quad (4.17)$$

$$\tilde{I}(k) = \int_0^R J_0(kr) |U(r)|^2 r dr , \quad (4.18)$$

where n specifies the type of noise under consideration, and $\tilde{I}(k)$ is the 2D axisymmetric Fourier transform of the beam intensity with k the radial wavenumber. The substrate Brownian noise has $n = 0$, coating Brownian and coating thermoelastic noises have $n = 1$, and substrate thermoelastic noise has $n = 2$. We are interested in comparing noises given by different beam shapes, so the overall constants C_n are not important. The resulting amplitude sensitivity is the square root of the noise, and has units of meters/ $\sqrt{\text{Hz}}$.

These scaling laws were derived for half-infinite mirrors, meaning that effects of the finite radius and thickness of the mirror are ignored. For the specific mirrors under consideration for Advanced LIGO, the width and thickness of the mirror are large enough compared to the beam width [45] that this should be a good approximation. The results of [49] suggest that we can expect corrections

of not much more than ten percent to the half-infinite scaling law expressions that we are using.

4.5 Eigenvalues of the Propagator

The idealized picture of a locked cavity is that the mirrors are perfectly aligned and a precise distance from each other. The beam should leave one mirror, reflect off the other, and when it returns it should be the same shape and exactly in phase. The beam will have lost some intensity due to diffraction and the finite extent of the mirror. In order to build up a very intense beam with a relatively weak laser, the beam must reflect very many times. The loss per half-trip (from one mirror to the other) must be below approximately 60 ppm. The majority of the loss will be due to absorption in the mirror and other factors, not diffraction. The commonly agreed upon budget for losses due to diffraction is 1 ppm.

In previous work, the clipping approximation is often used to estimate the diffraction loss by calculating the fraction of the intensity of the beam which falls outside the mirror. For an axisymmetric beam, it is given by

$$\mathcal{DL} \approx 2\pi \int_R^\infty |U_\alpha(r)|^2 r dr . \quad (4.19)$$

This is only an approximation because it does not take into account distortions in the beam shape caused by the diffraction of the beam from a finite mirror. Our numerical code is accurate and fast enough to compute the diffraction loss directly from the eigenvalues of the propagator, which is more accurate.

The propagator also allows us to estimate the difficulty of locking the interferometer. The finesse of Advanced LIGO is about 1200 [56]. This sets the width

of the resonance for the cavity [57] to be about $2\pi/1200 \approx 0.005$ radians. If any other modes with a small diffraction loss have an argument within this distance of the desired mode, there will be severe problems with locking the cavity. For axisymmetric modes, we find that this is not the case (see Table 4.5) and the arguments of the eigenvalues are well-separated enough that locking with Mesa or any of the other beams that we study should be no more difficult than locking the currently proposed Mesa.

Table 4.2: The diffraction loss and phase separation for eigenvalues with losses less than 10 percent, in three different hyperboloidal configurations with 1 ppm loss in the fundamental mode. Only axisymmetric modes are shown.

$\alpha = \pi$		$D = 9.62$ cm	
Diffraction Loss		Phase	
1.0 ppm		0.0	
121.7 ppm		-1.5104	
182.3 ppm		-0.5835	
334.8 ppm		-2.6677	
7941.6 ppm		2.2904	
45401.8 ppm		0.8325	
$\alpha = 0.95\pi$		$D = 9.71$ cm	
Diffraction Loss		Phase	
1.0 ppm		0.0	
136.2 ppm		-1.5079	
195.3 ppm		-0.5802	
302.3 ppm		-2.6630	
7225.3 ppm		2.2976	
44104.7 ppm		0.8422	
$\alpha = 0.90\pi$		$D = 11.01$ cm	
Diffraction Loss		Phase	
1.0 ppm		0.0	
18.6 ppm		-0.4797	
951.8 ppm		1.3048	
3400.8 ppm		2.7653	
5870.7 ppm		-2.3322	
61706.3 ppm		1.4493	

4.5.1 Integral Form of the Propagator

The propagator from a single point r, ϕ on one mirror to a point r', ϕ' on the other (see for instance [37]) is

$$\mathcal{K}(r, \phi, r', \phi') = \frac{ik_0}{4\pi\rho} (1 + \cos\theta) e^{-ik_0\rho}, \quad (4.20)$$

where ρ is the path length between the two points, θ the angle between the cavity axis and the path, and r and ϕ are the standard cylindrical radial coordinate and azimuthal angle. The cavity is very long compared to the radius of the mirrors so we can immediately make the paraxial approximation $\theta = 0$. The path length can be approximated as

$$\begin{aligned} \rho &= [(S_\alpha(r) + S_\alpha(r') - h(r) - h(r'))^2 \\ &+ (r^2 + r'^2 - 2rr' \cos(\phi - \phi'))]^{1/2} \\ &\approx L + \cos\alpha \left(\frac{r^2}{2L} + \frac{r'^2}{2L} \right) - h(r) - h(r') \\ &- \frac{rr'}{L} \cos(\phi - \phi') \end{aligned} \quad (4.21)$$

since $L \gg r$ and $L \gg h(r)$. This expression is the path length used to find the phase of the propagator. In the prefactor of the propagator we only need the much simpler $\rho \approx L$ in order to compute the amplitude to an accuracy of $R_{\text{mirror}}^2/L^2 \sim 10^{-9}$.

We assume that the cavity is axisymmetric, which means that the eigenmodes of the propagator can also be written as eigenstates of rotation. We write the complex beam amplitude of these eigenmodes as $V_m^{(i)}(r, \phi) = V_m^{(i)}(r) e^{-im\phi}$. Then any beam could be decomposed as

$$U(r, \phi) = \sum_{m=-\infty}^{\infty} \sum_{i=0}^{\infty} U_m^{(i)} V_m^{(i)}(r) e^{-im\phi}. \quad (4.22)$$

To determine the modes of the cavity, we have an eigenvalue problem given by the integral

$$\gamma_m^{(i)} V_m^{(i)}(r, \phi) = \int_0^{2\pi} \int_0^R r' dr' d\phi' \mathcal{K}(r, \phi, r', \phi') V_m^{(i)}(r', \phi'), \quad (4.23)$$

where $\gamma_m^{(i)}$ is the associated eigenvalue, and R is the mirror radius. Integrating over ϕ' produces

$$\gamma_m^{(i)} V_m^{(i)}(r) = \int_0^R r' dr' \mathcal{K}_m^r(r, r') V_m^{(i)}(r'), \quad (4.24)$$

$$\begin{aligned} \mathcal{K}_m^r(r, r') \equiv & \frac{i^{m+1} k_0}{L} J_m \left(\frac{k_0 r r'}{L} \right) \exp \left[i k_0 (-L + h(r) \right. \\ & \left. + h(r') - \cos \alpha \left(\frac{r^2}{2L} + \frac{r'^2}{2L} \right) \right], \end{aligned} \quad (4.25)$$

where for convenience we have defined a “radial kernel” \mathcal{K}_m^r . $J_m(x)$ is the m^{th} order Bessel function of the first kind. In what follows, we specialize to axisymmetric modes ($m = 0$), as we are not focusing on tilt or parametric instabilities which involve modes with $m > 0$.

4.5.2 Discrete Form of the Propagator

The above integral must be converted to a discrete sum to be suitable for numerical computation. We choose to do this using a Chebyshev quadrature [55, 58], which is appropriate for the finite range of r' . We use N collocation points r_k with associated integration weights w_k . The integral is converted to

$$\gamma_0^{(i)} V_0^{(i)}(r_j) = \sum_{k=1}^N \mathcal{K}_0^r(r_j, r_k) r_k w_k V_0^{(i)}(r_k). \quad (4.26)$$

This is a matrix eigenvalue problem which is easy to solve numerically. We will order the eigenstates by the number of radial nodes, the number of times that $V_0^{(i)}(r)$ goes to zero. The fundamental mode ($i = 0$) has zero nodes and as $R \rightarrow \infty$ it limits to the hyperboloidal beam that the cavity is designed to support. We will show that finite mirror effects cause the actual eigenstate to be slightly different.

We performed a convergence test where we varied the number of collocation points, N , and observed exponential convergence. The diffraction loss for the lowest eigenvalue for the $\alpha = \pi$ Mesa beam with $D = 10.67$ cm, $R = 17$ cm changes in the 5th digit (relative to its value of 1 PPM) when varying N between $N = 250$ and $N = 500$. We are using $N = 1000$ or $N = 500$ for all of the calculations in this paper. On a typical single processor 2 GHz laptop computer, our code takes ~ 10 seconds for $N = 500$ and ~ 1 minute for $N = 1000$, to calculate the eigenmodes of any given cavity.

4.5.3 Meaning of Eigenvalues

The eigenvectors of the propagator are the field amplitudes of the cavity's resonant modes. Fine-tuning the length L of the cavity to $L + \delta L$ changes the eigenvalue to $\exp(-ik_0\delta L)\gamma$. This tuning is used to select the desired mode; when the argument of the eigenvalue is zero it will resonate in the cavity. The magnitude of the eigenvalue is the fraction by which the amplitude changes during a half-trip. We must have $|\gamma| < 1$ for finite mirrors, because some light will always be lost to diffraction. We define the diffraction loss in parts per million (ppm) over

one half trip through the cavity as

$$\mathcal{DL} = 10^6(1 - |\gamma|^2). \quad (4.27)$$

Advanced LIGO requires a diffraction loss per half trip of about one ppm. In the next section, we will discuss the results of our analysis, where we study the noise characteristics of the hyperboloidal beams given the constraint on the diffraction loss.

4.6 Results

Our goal of reducing the noises in LIGO is constrained by the need to keep the diffraction loss at nearly 1 ppm. The hyperboloidal beams have two parameters, D and α . D is roughly the width of the beam, and α is the shape. The duality relation reduces the range of α that we need to consider. We focus our attention on $\pi/2 \leq \alpha \leq \pi$ since the intensity profiles are identical to those in the range from 0 to $\pi/2$ but the mirrors are nearly concentric as needed to decrease the tilt instability. The Mesa profile obtained for $\alpha = \pi$ has the flat-top shape required to decrease the thermal noises. As α goes toward $\pi/2$ the beam becomes more rounded, losing the flat top and sharp decay near the edge. We have shown that $D \rightarrow \infty$ the beam becomes a Gaussian whose width is $w_0\sqrt{1 + \cos^2 \alpha}/\sin \alpha$. For α near π , this Gaussian is nearly infinitely wide, and at $\alpha = \pi/2$ the beam becomes the minimal Gaussian of width w_0 . From the clipping approximation we can estimate that at $\alpha \approx 0.247\pi$ and $\alpha \approx 0.753\pi$, $D = \infty$ the diffraction loss is about 1 ppm. We do not have to consider any values of α between these since the widest beam consistent with the diffraction loss constraint would be the $D = \infty$ Gaussian.

4.6.1 Finite Mirror Effects

If the mirrors were infinite in extent, no light would propagate off of the mirror and there would be no diffraction loss, giving eigenvalues of unit magnitude. The clipping approximation assumes that the beam is the one supported by infinite mirrors. This is not the case as diffraction also causes the beam profile to change. The propagator is a more accurate calculation because it finds the precise beam profile supported by the mirrors. The clipping approximation is typically an underestimate of the diffraction loss [36, 35]. In Fig. 4.4a, we show that this is indeed usually the case. However, for some ranges of beam width D there is an anomalously low diffraction loss below the clipping approximation.

To study this effect, we varied the mirror radius and computed the diffraction loss as a function of D . The mirrors that we study in this paper have radius $R = 17$ cm, so we compared with $R = 16$ cm and $R = 18$ cm. The local minimum becomes narrower and shallower for increasing mirror radii. This is suggestive of a finite mirror effect that will disappear when the mirror radius is significantly larger than the radius of the beam.

Fig. 4.4b shows the deviation of the beam from the infinite mirror Mesa beam. A concentric Mesa ($\alpha = \pi$) beam with $D = 10.67$ cm is chosen. This is the location of a local minimum of the diffraction loss with respect to D . The beam intensity is computed directly from the eigenvector of the propagator. The intensity is normalized by integrating over the mirror rather than over the entire mirror plane as with the infinite mirror case. However, this only causes a fractional error in the normalization on the order of 10^{-6} . As shown in the figure, the finite mirror causes deviations from the infinite mirror beam; the plateau of the beam is less flat than expected. When the radius of the mirror is increased,

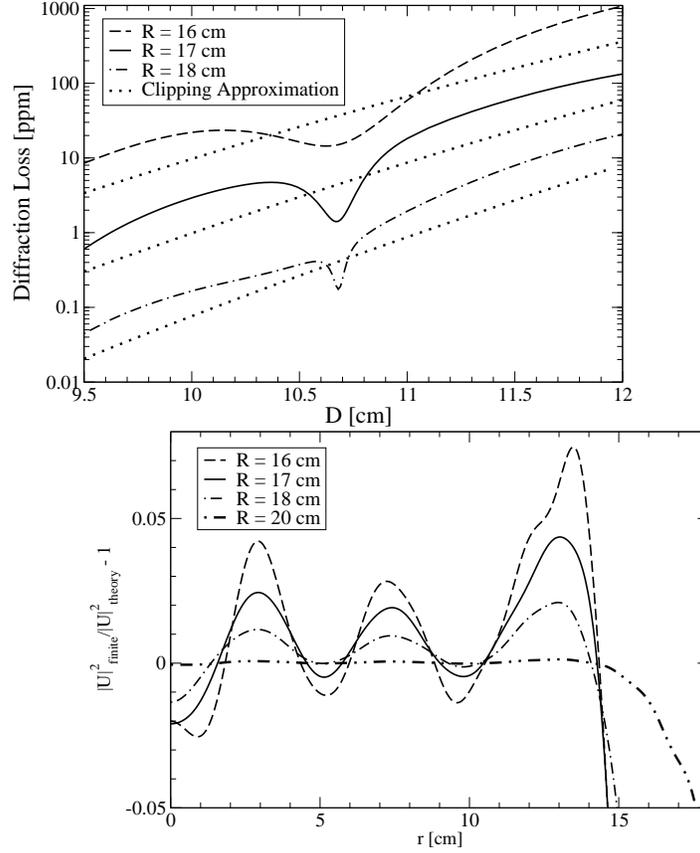


Figure 4.4: (a) The diffraction loss for an ($\alpha = \pi$) Mesa configuration ($\alpha = \pi$) is shown as a function of D on a logarithmic plot for several different mirror radii: $R = 16$ cm, $R = 17$ cm and $R = 18$ cm. The diffraction loss computed numerically using Eq. (4.27) (solid, dashed and dotted-dashed lines) exhibits local minima due to finite mirror effects. It can be seen that the minima get narrower as R increases and that they go below the values estimated using the clipping approximation (dotted lines). (b) The fractional difference $|U|_{\text{finite}}^2 / |U|_{\text{theory}}^2 - 1$ between the theoretical infinite-mirror beam intensity profile and the actual profile given by the first eigenvector of the propagator in Eq. (4.24) is plotted for $D = 10.67$ cm and $R = 16, 17, 18$ and 20 cm. The deviation decreases with R as expected.

the deviations retain their shape but decrease in size. For $R = 20$ cm, the beam is very close to the infinite mirror expectation. There is still a difference near the outer edge of the beam, causing the intensity to decay more quickly with radius than in the infinite mirror case.

Our numerical results suggest that the anomalous diffraction loss is related to the deviation of the beam from the ideal Mesa shape. As D increases, the clipping approximation predicts a smooth increase in diffraction loss due to the widening beam. Finite mirror effects increase with the ratio D/R , so they alone do not explain this unexpected behavior. For the values of D that yield an anomalous diffraction loss, the variations around the plateau (Fig. 4.4b) have an organized shape with an approximate wavelength of w_0 . The variations in these cases have a shape such that they alter the falloff of the beam at the edge, i.e. in the last two centimeters of the mirror. This has an obvious beneficial effect on the diffraction loss.

We expect that the fundamental mode of the cavity (the hyperboloidal shape, with no nodes), should have the lowest diffraction loss. However, this is not the case for all choices of α and D . Surprisingly, the diffraction loss of the first excited axisymmetric mode can decrease below that of the fundamental mode. This occurs, for $\alpha = \pi$, for a small range of D around 10.5 cm, with diffraction losses around 4 ppm. The arguments of two eigenvalues remain well separated. Fig. 4.5 shows the crossing of the two diffraction loss curves.

To investigate the cause of this crossing, we choose the specific value $D = 10.52$ cm and increase the mirror radius to 18 cm. The diffraction loss of the fundamental mode decreases from 3.7 ppm to 0.4 ppm, while the first excited mode only decreases from 3.6 ppm to 2.3 ppm so the diffraction loss of the

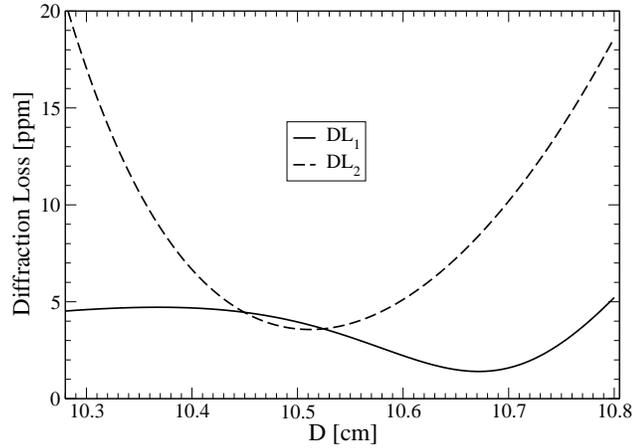


Figure 4.5: The diffraction losses of the fundamental mode and the first excited mode are plotted as a function of D for the $\alpha = \pi$ Mesa configuration. The two curves cross due to finite mirror effects causing an anomalous diffraction loss for the first excited mode.

fundamental mode is now lowest. Further increasing the mirror radius to 20 cm causes the losses of the second and third excited modes to cross. Despite changes of the mirror radius, the arguments of the eigenvalues change by less than a percent (we are only considering eigenvalues with losses less than 10,000 ppm because otherwise they would dissipate too quickly to be of interest). This dependence on mirror radius tends to confirm that this is a finite mirror effect. Having demonstrated that this effect can have substantial and beneficial effects on the diffraction loss, we now turn our attention to studying the parameter space of hyperboloidal beams in more detail. Further studies of the precise cause of the anomalous losses may want to focus on the deviation of the beam from its theoretical expectation, as well as the behavior with changing mirror radius.

4.6.2 Noises for fixed D

The width of the beam increases with increasing D , which averages the fluctuations over more of the mirror surface and therefore decreases the noise. It is less clear how the noise will behave when α is changed. We begin by fixing $D = 10$ cm and ignoring the diffraction loss constraint. Fig. 4.6 shows that all three types of noise increase as α moves away from 0 and π . Substrate thermoelastic noise is most affected by changing α , followed by the coating noises (recall that both types of coating noise follow the same scaling law).

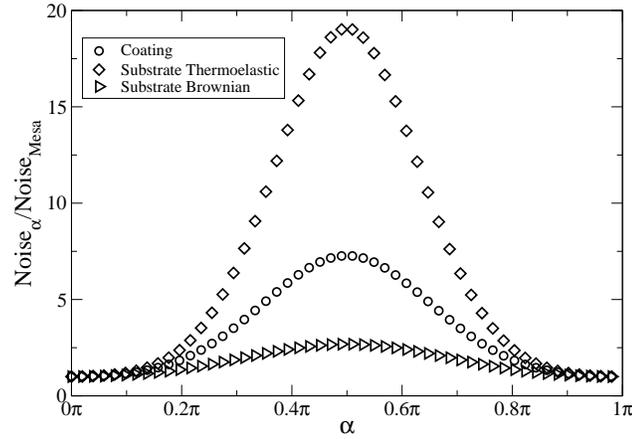


Figure 4.6: The noise ratios $\text{Noise}_\alpha / \text{Noise}_{\text{Mesa}} - 1$ for three types of noise are shown as a function of α for fixed $D = 10$ cm. The minimal Gaussian $\alpha = \pi/2$ has the highest noise and Mesa ($\alpha = 0, \pi$) has the lowest noise.

As α is decreased from π , the noises increase if D is kept fixed. At the same time, the diffraction loss decreases. If we keep the diffraction loss fixed, the D can be increased as α decreases toward $\pi/2$. Widening the beam tends to decrease the noise, which partially offsets the increase from changing α . We fixed the loss at 1.4 ppm and found that the noise still increases for beams other than Mesa. We expect that for larger diffraction losses this result will still hold.

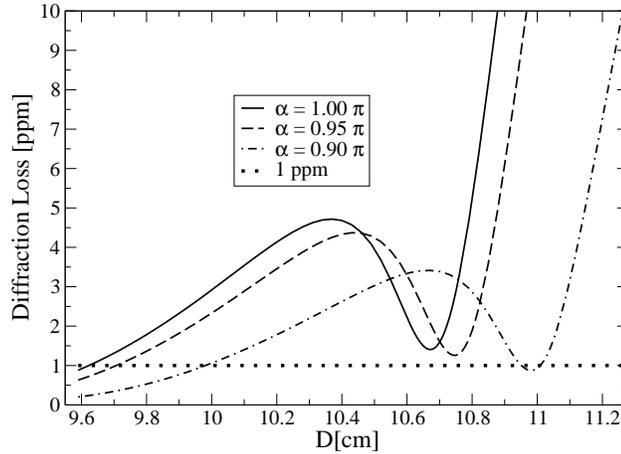


Figure 4.7: Diffraction loss as a function of D is displayed for $\alpha = \pi$, $\alpha = 0.95\pi$, $\alpha = 0.90\pi$. It can be seen that as α decreases the minimum diffraction loss is lower and occurs for larger D .

However, at 1 ppm the anomalous behavior of the diffraction loss due to finite mirror effects changes this conclusion.

4.6.3 Hyperboloidal beams with 1 ppm Diffraction Loss

As α decreases from π toward $\pi/2$, the beam loses its flat top and sharp falloff, and approaches the minimal Gaussian. Also, the D corresponding to the local minimum in diffraction loss increases, and the local minimum becomes deeper and wider. Fig. 4.7 shows the diffraction loss versus D for three values of α , while Fig. 4.8 gives D and diffraction loss at the local minimum for a range of α . The local minimum of the diffraction loss for a Mesa beam ($\alpha = \pi$) is at $D = 10.67$ cm, and has 1.4 ppm. As discussed above, the Mesa is has the lowest noise of the hyperboloidal family for this diffraction loss.

If a diffraction loss of strictly 1 ppm is required, the beam must be reduced

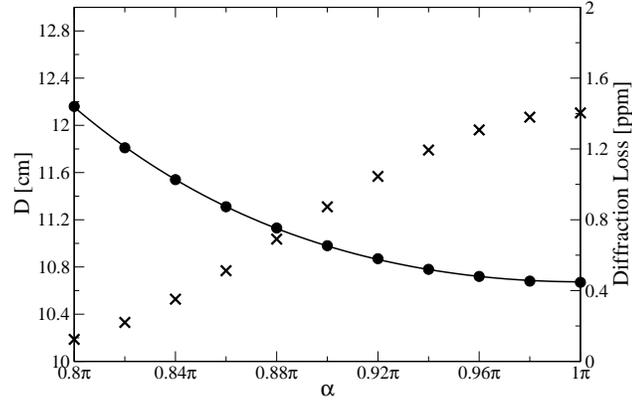


Figure 4.8: The minimum diffraction loss in ppm (represented by crosses) and the corresponding D (dots) are shown as a function of α . The solid line represents the best exponential fit of the form $a + \exp(b + c \sin^2 \alpha)$ with $a = 0.094$, $b = -4.34$ and $c = 2.20$.

to a width of $D = 9.62$ cm in the Mesa configuration. Alternatively, we can consider other values of α . The local minima are displayed in Fig. 4.8, which gives the values of D as well as the corresponding diffraction losses. Note that for $\alpha = 0.9\pi$, the diffraction loss at the local minimum is now below 1 ppm. Fig. 4.9a shows the maximum D that yields a 1 ppm loss, as α is varied. There is a discontinuity between $\alpha = 0.91\pi$ and 0.92π because below 0.92π the diffraction loss at the local minimum is below 1 ppm. The noises therefore drop substantially when $\alpha = 0.91\pi$ as in Fig. 4.9b. The coating noise decreases by 12% and the substrate thermoelastic by 19%, relative to the 1 ppm Mesa beam. A strict requirement of 1 ppm diffraction loss therefore combines with the finite mirror effects to make the $\alpha = 0.91\pi$, $D = 10.94$ cm configuration the best choice.

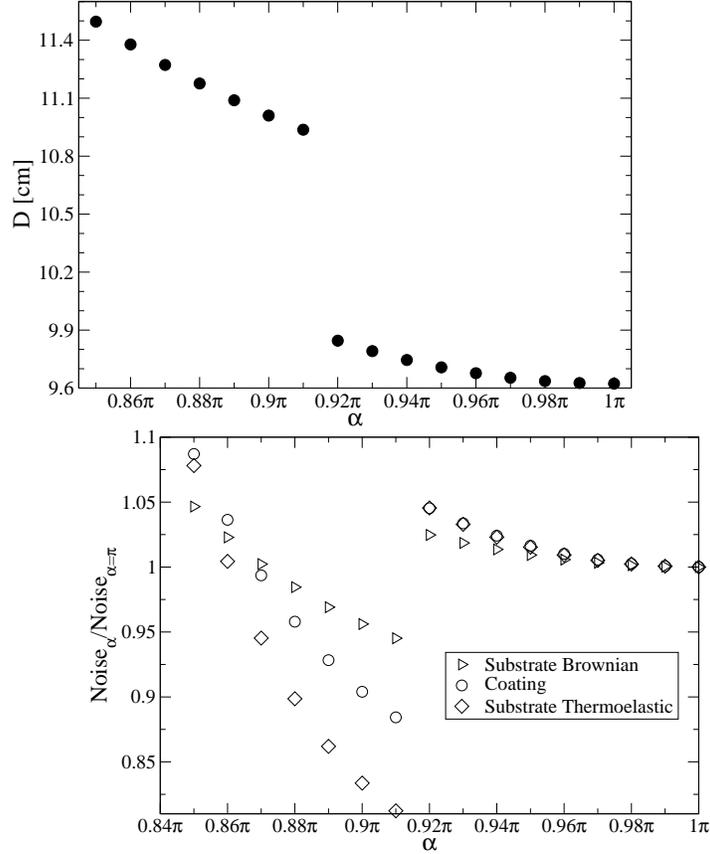


Figure 4.9: (a) The largest values of D giving a strict 1 ppm diffraction loss are shown as a function of α . The discontinuity is caused by finite mirror effects as explained in the text. (b) The corresponding values of the noise, normalized so that the noises for the 1 ppm Mesa ($\alpha = \pi$, $D = 9.62$ cm) are all equal to 1. The discontinuity in allowed maximum D leads to a sharp drop in noise at $\alpha = 0.91\pi$.

4.6.4 Correcting for Finite Mirror Effects

Restructuring the mirror to specifically account for finite mirror effects allows us to increase D in hyperboloidal beams (thereby reducing thermal noise), while keeping within given diffraction loss constraints. The restructured beams can reduce the diffraction loss by a factor of 30 to 100, allowing a wider beam. For the Mesa case this allows for a net noise reduction of 30% for the beam satisfy-

ing the 1ppm diffraction loss constraint.

As noted above, the original Mesa beam used to construct the mirror is infinite in extent. The mirror is designed to be a phasefront of the theoretical beam. Since the mirror is actually finite, for $D \sim R$ there can be substantial effects due to missing light that was incident on the mirror plane outside the mirror radius. To account for these finite mirror effects we reconstruct the mirror, with the goal of making the phase of the first eigenbeam constant at the mirror surface, rather than the phase of the idealized infinite beam.

The propagator formulation allows us to explicitly calculate the phase of the eigenbeams, as a function of r . As the mirror deviation from the fiducial spheroid $h(r)$ enters into the calculation of phase through the propagator, we use an iteration scheme to adjust the mirror to match the eigenbeam phasefront motivated by the argument of the propagator:

$$\begin{aligned} \arg[\mathcal{K}(r, r')] &\simeq \pi/2 + k_0 (h(r) + h(r') - L) \\ &\quad - k_0 \cos \alpha \left(\frac{r^2}{2L} + \frac{r'^2}{2L} \right). \end{aligned} \quad (4.28)$$

We see in Eq. (4.28) that if the phase of the eigenbeam is too large at some radius r , reducing the value of $h(r)$ should act to roughly reduce the phase of the new eigenbeam.

With this motivation, we apply the simple iteration scheme:

$$h_{i+1}(r) = h_i(r) - c \times \arg \left[V_i^{(1)}(r) / V_i^{(1)}(0) \right], \quad (4.29)$$

where the i denotes the i^{th} iteration, $V_i^{(1)}(r)$ is the first eigenbeam for the mirror with deviation $h_i(r)$, and $c > 0$ is an arbitrary constant less than unity, chosen to prevent overshoot.

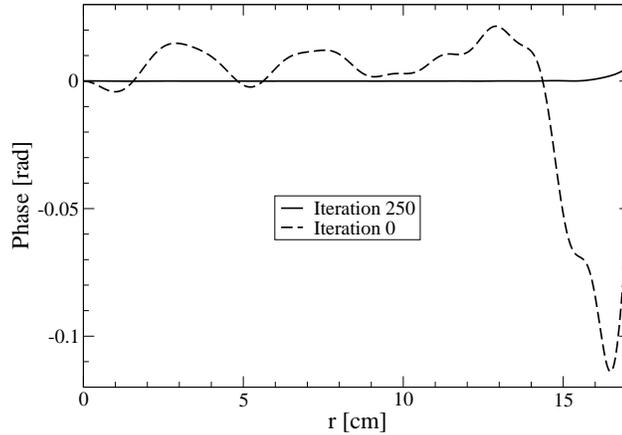


Figure 4.10: The phase of the fundamental eigenbeam as a function of radius is shown for iteration zero and 250. It can be seen that the iteration scheme drives it closer to zero as expected.

This iteration scheme successfully reduces the relative phase of the eigenbeam, as shown in Fig 4.10. Optimizing the mirror surface to match the phase-front of the primary eigenbeam also acts to reduce the diffraction loss for that mode in general, with the iteration scheme providing convergence towards an apparent lower bound for the diffraction loss, while increasing the diffraction loss for other higher-order eigenbeams. This lower bound increases with D (Fig 4.11).

This diffraction loss is plotted against the iteration number for the Mesa ($\alpha = \pi$) case with $D = 11.35$ cm in Fig 4.12. The iteration scheme is shown to lower the diffraction loss for this D from 46.5 ppm to a ~ 1 ppm lower bound, satisfying the required design constraint. The diffraction losses of higher order modes are more than doubled in the iterated case as illustrated in Table 4.6.4.

The beam for the iterated mirror with $D = 11.35$ cm is close to the original Mesa, but with variations in the central plateau of relative amplitude $\sim 1/30$ and variations of radial wavelength $\sim w_0$ (Fig 4.13). This seems to be an unavoidable

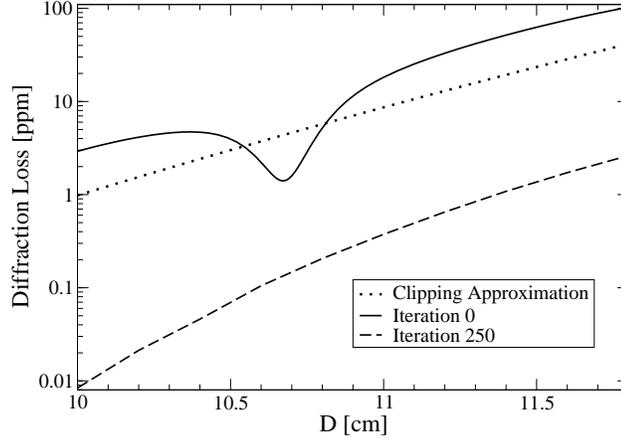


Figure 4.11: The diffraction loss in ppm calculated using the clipping approximation is compared to that using the propagator eigenvalues for the iterated and original mirrors as a function of D . As before, the configuration studied is $\alpha = \pi$ Mesa with $R = 17$ cm. The iteration process lowers the diffraction loss by a factor of 30 to 100.

consequence of a finite R , as even the original eigenbeam has roughly similar features. Despite changing the variations in the plateau of the beam intensity the process of iteration does not significantly affect the noises computed using Eq. (4.17).

Similarly the iterated mirror has variations of the central mirror shape of similar radial scale, with amplitude on the order of 2 nm, shown in Fig 4.14. The most significant feature of the iterated mirror is the inward tilting of the outer edge of the mirror, though preliminary studies show that the primary contribution to reducing the diffraction loss is due to the variations near the center of the mirror.

Reformulating the mirror to account for finite mirror effects allows us to increase the D parameter of the beam from 9.62 cm to 11.35 cm for a concentric Mesa beam while still maintaining a 1ppm diffraction loss. Though this design

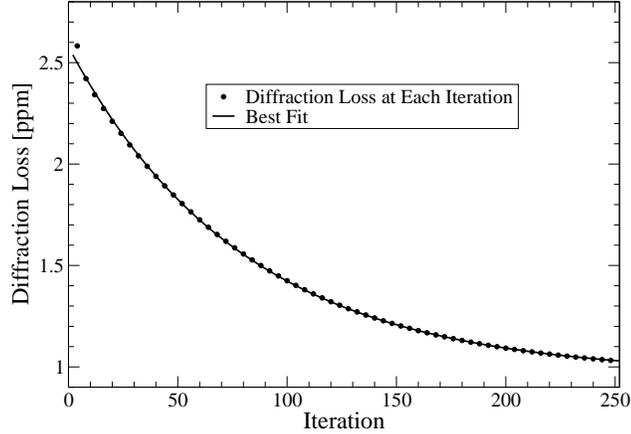


Figure 4.12: The diffraction loss is shown as function of iteration number for a $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm. The original beam (not shown) has diffraction loss at 46.5 ppm, and it can be seen that after a few iterations the diffraction loss begins to converge to an exponential with lower bound ~ 1 ppm. The best fit exponential is given by $0.96 + 1.616 \exp(-0.013i)$ ppm, where i is the iteration number.

may introduce more complications in the construction of the mirror itself, it allows a significant reduction in noise by broadening the beam. This iteration scheme can also be used for other values of α , as shown in Table 4.6.4, where the iterated mirror for $\alpha = 0.9\pi$ has a diffraction loss lower bound of 1 ppm for $D = 11.87$ cm. However, we find that $\alpha = \pi$ is optimal for noise reduction.

4.7 Conclusions

In this paper, we studied thermal noise and diffraction loss for the hyperboloidal family of light beams and mirror shapes in detail for the first time. This family had been initially proposed to unify the concentric ($\alpha = \pi$) and nearly-flat ($\alpha = 0$) Mesa configurations through variations of the twist angle α . In this

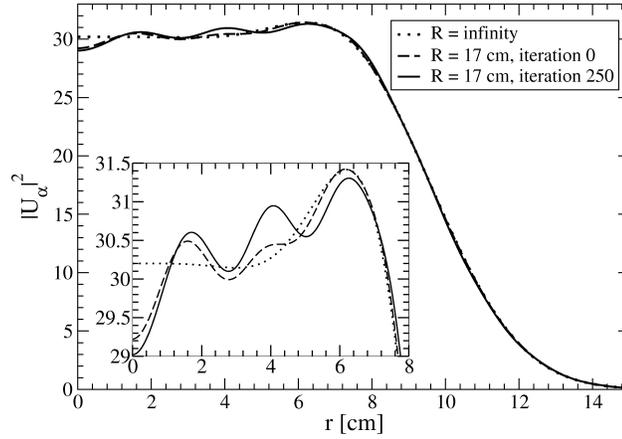


Figure 4.13: The intensity profile $|U_\pi|^2$ for the mirror with $R = \infty$, $R = 17$ cm uniterated and $R = 17$ cm at iteration 250 are compared for the $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm. The finite mirror effects induce oscillations in the intensity profile that do not disappear when the mirror is corrected. An inset shows the central 8 cm ‘plateau’ of the beam in detail.

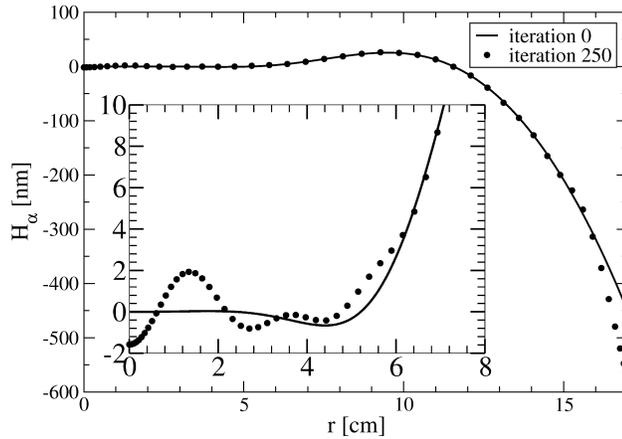


Figure 4.14: The correction h_α to the mirror at iteration 0 and at iteration 250 are compared for the $\alpha = \pi$ Mesa configuration with $D = 11.35$ cm, $R = 17$ cm. The iteration scheme introduces some bumps on the mirror of the size ~ 2 nm. The inset shows the central 8 cm of the mirror in more detail.

paper we also presented an analytic proof that Gaussian beams are a limiting case of the hyperboloidal beam as $D \rightarrow \infty$. This was previously conjectured in Ref. [42]. We developed a pseudo-spectral code both fast and accurate enough to calculate the diffraction loss directly from the beam propagator. We find that the finite radius of the mirror causes beam shapes to deviate significantly from the infinite-mirror theoretical expectations. This causes a previously unnoticed local minima in the diffraction loss that can be exploited to find a natural beam width D for the current diffraction loss constraints of about 1 ppm. For an $\alpha = \pi$ Mesa beam a local minimum occurs at $D = 10.67$ cm giving 1.4 ppm diffraction loss for a mirror of radius $R = 17$ cm. If one requires a strict enforcement of the 1 ppm diffraction loss we show that a hyperboloidal beam with $\alpha = 0.91\pi$ and $D = 10.94$ cm has lower noise than that of the $\alpha = \pi$ Mesa with 1 ppm diffraction loss. The coating noise changes by about 12% and the substrate thermoelastic noise and substrate Brownian noise change by 5% and 19%, respectively.

We also propose new mirror and beam shape configurations that explicitly account for finite mirror effects by reformulating the mirror surface to coincide with the phasefront of the primary eigenbeam. These beams reduce the diffraction loss by more than an order of magnitude for the range of D considered here (between 10 cm and 11.8 cm). This allows the use of wider beams for the same diffraction loss constraints on the primary eigenmode, while the diffraction losses of higher order modes (both axisymmetric and non-axisymmetric) are increased. We are able to widen the $\alpha = \pi$ Mesa beam for a $R = 17$ cm mirror while keeping the diffraction loss fixed at 1 ppm from a width of $D = 9.62$ cm to $D = 11.35$ cm. This lowers the coating thermal noise by about 30% (compared to the smaller D Mesa) and the other noises (substrate Brownian and substrate thermoelastic noise) by comparable factors. However, feasibility of the

construction of the mirrors must also be taken into account.

The non-iterated beams taking advantage of the local minimum in diffraction loss discussed above are supported by strict hyperboloidal or Mesa mirrors, which would be no harder to make than the current Mesa designs, and would still lower the coating thermal noise by 12%. If one is to consider the more ambitious goal of lowering the coating thermal noise by 28% while using a beam that is very similar to Mesa through the iteration scheme described, the limitations on mirror manufacturing errors are likely to be more stringent, but still less than the currently considered conical beams [51, 40]. In addition, the methods developed here for reducing the diffraction loss of the Mesa beam may be applied in the case of the conical beams previously considered [40, 51]. The phase fronts of conical beams considered there have not been optimized to match the finite mirror surface.

Recently, parametric instability [53, 59, 60, 61] was found to be a serious problem in Advanced LIGO. Choosing D at the minimum of the diffraction loss curve of the hyperboloidal beams increases the diffraction loss of the higher eigenmodes, thus in principle somewhat improving the parametric instability. The effect is most pronounced in the case of the iterated mirrors which also have the most drastic thermal noise reduction.

Table 4.3: The phase separation for the axisymmetric ($m = 0$) modes with diffraction losses less than 10%, both before and after the iteration scheme is applied. The phases do not change significantly as the mirror is iterated. The absolute value of the eigenvalues with nonzero phase (and hence the diffraction loss) increase upon iterating. Preliminary calculations show that non-axisymmetric modes have diffraction losses increased by the iteration process while the phases change by no more than 5%.

$\alpha = \pi$		$D = 11.35$ cm (Uniterated)	
Diffraction Loss	Phase		
46.5 ppm	0.0		
128.5 ppm	-0.4313		
341.7 ppm	-1.1895		
10530.8 ppm	-2.1470		
38445.0 ppm	3.0277		
$\alpha = \pi$		$D = 11.35$ cm (Iterated)	
Diffraction Loss	Phase		
1.0 ppm	0.0		
320.6 ppm	-0.4319		
1100.1 ppm	-1.1920		
26167.4 ppm	-2.1593		
66808.7 ppm	2.9873		
$\alpha = 0.9\pi$		$D = 11.87$ cm (Uniterated)	
Diffraction Loss	Phase		
43.5 ppm	0.0		
205.0 ppm	-0.4173		
371.2 ppm	-1.1665		
10626.6 ppm	-2.1082		
50723.6 ppm	3.0878		
$\alpha = 0.9\pi$		$D = 11.87$ cm (Iterated)	
Diffraction Loss	Phase		
1.0 ppm	0.0		
366.6 ppm	-0.4185		
2013.4 ppm	-1.1701		
32628.3 ppm	-2.1267		
87359.8 ppm	3.0335		

Table 4.4: The coating, substrate Brownian and substrate thermoelastic noise are displayed after the iteration process. The diffraction loss is kept constant at 1 ppm. The noises are normalized to noises of the original ($\alpha = \pi$) Mesa (with $D = 9.62$ cm) which gives the 1 ppm diffraction loss. The iteration scheme lowers the noise by about 30% for $\alpha = \pi$ by allowing larger D for the same 1 ppm diffraction loss.

α	D [cm]	Coating Noise	Substrate Brownian Noise	Substrate Thermoelas- tic Noise
π	11.35	0.72	0.84	0.63
0.9π	11.87	0.80	0.90	0.69

CHAPTER 5
F(R) MODIFICATIONS OF GRAVITY

5.1 Introduction

5.1.1 The dark energy problem

Cosmological measurements indicate that the expansion of the Universe is accelerating. Cosmology treats the Universe as homogeneous and isotropic, which is a good approximation on scales larger than 100 megaparsecs. Observations of the cosmic microwave background and galaxy clustering determine the spatial geometry to be very nearly flat, which is the result expected if the Universe underwent a period of inflation early in its evolution. The stress-energy tensor is modelled as a cosmological fluid made up of several components. Radiation and other relativistic particles were an important contribution only in the early Universe. For most of the history of the Universe, matter has been the dominant component, and as a result the expansion has been decelerating. However, relatively recently the expansion has begun to accelerate. The most economical explanation in standard General Relativity is that the remaining component of the cosmological fluid is a cosmological constant. As the matter density has been diluted by the expansion, the cosmological constant has begun to dominate.

While in principle the cosmological constant could take any value, there is a quantum mechanical contribution whose scale can be estimated. Loop effects from quantum field theory contribute to the cosmological constant. The con-

tribution from these loop effects is divergent, but the cosmological constant is obviously not infinite. We can impose a cutoff on the zero-point energy by assuming that there is some new physics which first appears at the lowest energy that has not yet been probed with accelerators, which is approximately 1 TeV. The mass density resulting from a cutoff at energy scale E is on the order of $E^4/\hbar^3 c^5$, which gives a mass density $\rho \sim 10^{32}$ kg/m³.

The actual value of the cosmological constant, which makes up approximately 70 percent of the current energy density of the cosmological fluid, is $\rho_\Lambda = 7 \times 10^{-27}$ kg/m³. The natural value of the cosmological constant is about 60 orders of magnitude above the value that is actually measured. Obtaining a value this low for the zero-point energy would require a tremendous amount of fine-tuning. This is the so-called cosmological constant problem. A similar question is the coincidence problem of why it is only relatively recently in cosmological history that dark energy has begun to dominate.

5.1.2 The F(R) Idea

Another way to explain the cosmological observations is by modifying gravity. In this approach the dark energy is not a component of the cosmological fluid; instead, we modify Einstein's equation so that a Universe with only matter and radiation can produce acceleration. The modification considered in this paper is F(R) gravity, which is derived from the action [65]

$$S = \int d^4x \sqrt{-g} F(R) . \quad (5.1)$$

Here R is the Ricci scalar and $F(R)$ is a function of the Ricci scalar. The usual Einstein-Hilbert action with a cosmological constant Λ corresponds to $F(R) =$

$R - 2\Lambda$. The equations of motion resulting from the $F(R)$ action imply covariant conservation of the stress-energy tensor, just as in GR. However, the equations of motion now contain fourth-order derivatives of the metric, rather than two derivatives as in GR.

We will often use the notation $F(R) = R + f(R)$, so that $f(R)$ is the correction to the usual Hilbert action. We will use primes to denote derivatives with respect to whatever variable is indicated, while derivatives of f with respect to R will be denoted with subscripts, i.e. $f_R = df/dR$. We set $\hbar = c = 1$ throughout, and make use of the reduced Planck mass $M_{Pl} = (8\pi G)^{-1/2}$. The Hubble parameter in these units corresponds to a mass $H_0 = 1.5 \times 10^{-33}$ eV.

5.1.3 Original CDTT Model

The original model introduced by Carroll et. al. to explain the cosmological acceleration [65] had the form

$$f(R) = -\mu^2 \left(\frac{\mu^2}{R} \right)^n \quad (5.2)$$

where n is a positive integer. This model is often referred to as the ‘‘CDTT’’ model. The mass scale μ is of order the Hubble scale H_0 . One would expect that in local tests of gravity, the density of matter is very high and the curvature should also be high, making the correction Eq. (5.2) small. By contrast, at the low density at the cosmological scale today, this correction Eq. (5.2) contributes significantly to the action and causes a modification of the dynamics.

Chiba [66] showed that the model Eq. (5.2) does not work as intended. In GR, we have $R = \rho/M_{Pl}^2$ for pressureless matter with density ρ . In $F(R)$ gravity, the identification between high density and high curvature breaks down.

In models like CDTT, R is generally of order μ^2 even in high density regions. Chiba showed this by showing that $F(R)$ theories are equivalent to scalar-tensor theories in which the scalar field mass is of order μ^2 . The resulting very long-ranged scalar field causes easily detectable deviations of the metric of the Sun from the predictions of GR.

5.1.4 The Chameleon Effect

Khoury and Weltman [67, 68] and others [69] have shown that scalar-tensor theories can be constructed where the perturbation to the cosmological background value of the scalar field due to the Sun cannot be treated as a linear perturbation, as is often assumed. In these theories, the scalar field perturbation generated by the Sun is much smaller than the linear prediction, suppressing the modifications to Solar System observables. Because they hide the scalar field, these are named “Chameleon” models. Other papers [70, 72, 73] have constructed $F(R)$ theories that utilize the chameleon effect. The scalar-tensor description of an $F(R)$ theory has a nonlinear potential for the scalar field. The effective mass of the scalar field is of order H_0^2 outside a planet, as in the CDTT model (Eq. (5.2)). However, the potential in a chameleon theory is such that a small perturbation of the scalar field pushes it into a nonlinear regime of the potential where the scalar field becomes much more massive. The result is that only a thin shell near the surface of an object contributes to the scalar field; at points interior to the thin shell, the scalar field is effectively short-ranged and does not contribute to the field outside. Because of the Chameleon effect, some $f(R)$ models can evade the problem pointed out by Chiba and are not ruled out by Solar System observations.

5.1.5 Observational Constraints on F(R) Models

The first requirement is that the F(R) theory reproduce the observed cosmic acceleration. We will simplify our discussion by considering only spatially flat FLRW metrics. The Friedmann equation in GR relates the Hubble parameter to the density. There are two more derivatives of the metric in F(R) theories, so the modified Friedmann equation involves the next two derivatives of the scale factor a in the form of the deceleration parameter and the jerk, which are dimensionless second and third derivatives of the scale factor. Our cosmological measurements effectively tell us the Hubble parameter, deceleration, jerk, and matter density at the current time, and with the Friedmann equation these give us one constraint on the F(R) parameters.

Solar System tests of gravity measure the metric of the Sun to high precision, and the results can be expressed as constraints on the post-Newtonian parameters. These constraints set a limit on how thick the thin shell of the Sun can be, in a chameleon model. This gives a constraint

$$|\Phi| \lesssim 5 \times 10^{-11} \quad (5.3)$$

for the asymptotic value of Φ outside the Solar System. (See Sec. 5.4 for details).

Considerations of cosmological structure formation by Hu and Sawicki [70] place constraints on the value of the scalar field on cosmological scales, where the curvature is on the scale of H_0^2 . It is not certain whether these constraints are valid because the evolution of the scalar field in chameleon models during structure formation is not yet known. We will assume that the galaxy has a thin shell, which implies the constraint

$$|\Phi| \lesssim 10^{-6} \quad (5.4)$$

for the asymptotic value of Φ outside the Galaxy.

We will show that for $F(R)$ models that utilize the chameleon effect, satisfying the above constraints requires an unnatural fine-tuning. Our conclusion is that $F(R)$ models are not appealing as a solution to the dark energy problem. We will also show that the fine-tuned models which successfully evade the Galaxy constraint tend to easily satisfy the Solar System constraint as well.

5.2 Equations of Motion

The equations of motion can most easily be written in terms of a new scalar field

$$\Phi = f_R(R) \quad (5.5)$$

(recall that $F = R + f$ and $f_R = df/dR$). They are

$$(1 + \Phi) \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) + g_{\mu\nu} \square \Phi - \nabla_\mu \nabla_\nu \Phi - \frac{1}{2} g_{\mu\nu} (f - R f_R) = M_{Pl}^{-2} T_{\mu\nu} . \quad (5.6)$$

A more useful way of writing Eq. (5.6) involves splitting into the trace and traceless parts. Taking the trace gives an equation for the scalar field:

$$\square \Phi - \frac{1}{3} (R + 2f(R) - R f_R(R)) = \frac{1}{3M_{Pl}^2} T . \quad (5.7)$$

This is a wave equation for the field Φ , with a potential term and the trace T of the stress-energy tensor as a source term. The potential term $(R + 2f - R f')/3$ can be written in terms of Φ by inverting the relation $\Phi = f_R(R)$. An alternative form of this equation is

$$f_{RR} \square R + f_{RRR} (\nabla_\mu R) (\nabla^\mu R) - \frac{1}{3} (R + 2f - R f_R) = \frac{1}{3M_{Pl}^2} T , \quad (5.8)$$

which does not require solving for R in terms of Φ . Because the function $R(\Phi)$ typically includes branch cuts and other problems, this second form may be

more amenable to computation; however, in our analysis we will use the first form.

Accompanying the trace equation is the traceless part of Eq. (5.6), which is

$$(1 + \Phi) \left(R_{\mu\nu} - \frac{1}{4} g_{\mu\nu} R \right) - \left(\nabla_\mu \nabla_\nu - \frac{1}{4} g_{\mu\nu} \square \right) \Phi = M_{Pl}^{-2} \left(T_{\mu\nu} - \frac{1}{4} g_{\mu\nu} T \right). \quad (5.9)$$

The equations of motion therefore consist of a nonlinear wave equation for the scalar field Φ , equations for the traceless part of the Ricci tensor, and the relation Eq. (5.5) between Φ and the Ricci scalar R .

5.2.1 Exact Solutions

We will first consider some exact solutions to the theory with no stress tensor, $T_{\mu\nu} = 0$. In standard GR, the vacuum solution is simply the flat space Minkowski metric which has $R_{\mu\nu} = 0$. The vacuum solution of GR with a cosmological constant is either deSitter space or anti-deSitter space, depending on the sign of the cosmological constant.

We can easily show that the vacuum solutions of F(R) theories are either deSitter or anti-deSitter. The deSitter metric and Ricci tensor are

$$ds^2 = - \left(1 - \frac{\Lambda}{3} r^2 \right) dt^2 + \left(1 - \frac{\Lambda}{3} r^2 \right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2), \quad (5.10)$$

$$R_{\mu\nu} = \Lambda g_{\mu\nu}, \quad R = 4\Lambda. \quad (5.11)$$

Here Λ is the cosmological constant; the metric is deSitter for $\Lambda > 0$, anti-deSitter for $\Lambda < 0$, and flat Minkowski spacetime for $\Lambda = 0$. From now on, we will simply refer to all three possibilities as deSitter metrics. The constant Ricci scalar implies constant Φ , and the terms in the equation of motion involving derivatives are now zero. We see that Eq. (5.9) is satisfied because the Ricci tensor has

no traceless part. The allowed values of Λ are given by Eq. (5.7), which yields the condition $R + 2f - Rf' = 0$. A given F(R) theory may have several allowed values of Λ , possibly including $\Lambda = 0$.

Some other exact solutions of F(R) gravity are the Schwarzschild-deSitter and Kerr-deSitter solutions for non-spinning and spinning black holes with a cosmological constant. The Ricci tensor of these metrics is still $R_{\mu\nu} = \Lambda g_{\mu\nu}$. Birkhoff's theorem in GR says that the metric exterior to any spherical star is the Schwarzschild metric, but this result does not apply to F(R) gravity. Black holes are the same in F(R) and GR, but the exterior solutions of even spherical stars are different.

5.2.2 Cosmological Evolution

The Universe has been evolving cosmologically from high curvature to low curvature, so it is only at late times that the cosmological evolution of F(R) models will diverge from Λ CDM. Cosmological measurements effectively give us knowledge of time derivatives of the scale factors at the present epoch. We will need three derivatives: the Hubble parameter H , the deceleration parameter q , and the jerk j . These are defined by

$$H = \frac{1}{a} \frac{da}{dt}, \quad q = -\frac{1}{aH^2} \frac{d^2a}{dt^2}, \quad j = \frac{1}{aH^3} \frac{d^3a}{dt^3}. \quad (5.12)$$

Assuming a spatially flat Universe, the Ricci scalar is

$$R = 6 \left[\left(\frac{1}{a} \frac{da}{dt} \right)^2 + \frac{1}{a^2} \frac{d^2a}{dt^2} \right] = 6H^2(1 - q). \quad (5.13)$$

The modified Friedmann equation [70] for F(R) gravity is

$$H^2 + \frac{1}{6}f - f_R(H^2 + HH') + H^2 R' f_{RR} = \frac{1}{3M_{Pl}^2} \rho, \quad (5.14)$$

where a prime denotes the derivative with respect to $\ln a$. Substituting in the definitions above, the Friedmann equation becomes

$$H^2 + \frac{1}{6}f + qH^2 f_R + 6(j - q - 2)H^4 f_{RR} = \frac{1}{M_{Pl}^2} \Omega_M, \quad (5.15)$$

with $R = 6H^2(1 - q)$, and Ω_M the observed fraction of matter. This equation will provide a single constraint on the parameters of the F(R) model. If the next derivative, called the ‘‘snap’’, could be measured accurately another constraint would be possible by taking a derivative of the Friedmann equation and using energy conservation on the right hand side.

Values suggested by the WMAP 5-year data [75] for the Λ CDM model are $\Omega_M = 0.26$, $q = -0.61$, and $j = 1$. In fact, a GR model with only matter and cosmological constant (radiation is negligible) always has $j = 1$, and q is determined by Ω_M and the flatness assumption. The increased range of freedom of an F(R) model allows us to fit cosmological measurements that would not be consistent in Λ CDM. Rapetti et.al. [71] have used various cosmological observations to find H , q , and j without assuming Λ CDM. The values that they give are $q = -0.81 \pm 0.14$, $j = 2.16^{+0.81}_{-0.75}$, and $\Omega_M = 0.306^{+0.042}_{-0.040}$. We will also use $H = 72 \text{ km/sec / Mpc}$.

5.2.3 Perturbative Solution

We now focus our attention on the scalar field equation of motion:

$$\square\Phi - V'(\Phi) = \frac{\kappa}{3}T. \quad (5.16)$$

The potential term $V'(\Phi)$ is equal to

$$\frac{1}{3}(R + 2f - Rf_R) , \quad (5.17)$$

except rewritten as a function of Φ by using the relation Eq. (5.5) between R and Φ . We find the effective mass of the scalar field from

$$m_\Phi^2 = V''(\Phi) = \frac{(1 + f_R - Rf_{RR})}{3f_{RR}} . \quad (5.18)$$

The mass of the field depends on the background value of Φ . Consider a background with low curvature $R \sim H_0^2$; for an F(R) theory that explains the cosmic acceleration, the mass on this background is also order H_0^2 .

We now attempt to model an isolated spherical star in a cosmological setting. There is a time-dependent background cosmological value of Φ that we will write as $\Phi_c(t)$, and we will assume that $m_\Phi(\Phi_c(t)) \sim H_0$ on this background. The background cosmological stress-energy tensor has trace $T_c(t)$. We write the field and the source as $\Phi = \Phi_c(t) + \delta\Phi$ and $T = T_c(t) + \delta T$. Eq. (5.16) becomes

$$\square\Phi_c(t) + \square\delta\Phi - V'(\Phi_c(t) + \delta\Phi) = -\frac{1}{3M_{Pl}^2}T_c(t) - \frac{1}{3M_{Pl}^2}\delta T . \quad (5.19)$$

This can be split into an equation for the background evolution and for $\delta\Phi$ as

$$\square\Phi_c(t) - V'(\Phi_c(t)) = -\frac{1}{3M_{Pl}^2}T_c(t) , \quad (5.20)$$

$$\square\delta\Phi - V'(\Phi_c(t) + \delta\Phi) + V'(\Phi_c(t)) = -\frac{1}{3M_{Pl}^2}\delta T . \quad (5.21)$$

The background equation, evaluated with an FLRW metric (neglecting backreaction from the perturbation), is the companion cosmological equation of motion to the modified Friedmann equation, Eq. (5.15). It can be derived from Eq. (5.15) by taking a time derivative and using conservation of the stress-energy tensor.

The key assumption will be that

$$V'(\Phi_c + \delta\Phi) - V'(\delta\Phi) = m_\Phi^2(\Phi_c) \delta\Phi + \mathcal{O}(\delta\Phi^2). \quad (5.22)$$

When the length scale of the source is much shorter than m_Φ^{-1} and Eq. (5.22) is a valid approximation, we will say that the field is in the perturbative regime.

We will only be concerned with the field at short distances and for short periods of time compared with cosmological distance and time scales. We consider only weakly gravitating stars, so the d'Alembertian in Eq. (5.16) becomes a flat-space Laplacian; we also can neglect the pressure of the star so that $T = -\rho$. We also assume that the length scale of the source is much shorter than m_Φ^{-1} , so we can neglect the mass of the scalar field. We have

$$\nabla^2 \delta\Phi = -\frac{1}{3M_{Pl}^2} \rho. \quad (5.23)$$

The variables are now simply functions of r . If M is the mass of the star, it is easy to show that the solution is

$$\delta\Phi(r) = \frac{1}{12\pi M_{Pl}^2} \frac{M}{r} = \frac{2GM}{3r}, \quad (5.24)$$

where we have restored Newton's constant in the second equality. In the perturbative regime Φ couples to matter with about the same strength as gravity. Chiba's transformation to a scalar-tensor theory [66] shows that for a spherical star in the perturbative regime, there will be large, measurable deviations from GR. The same result can be obtained considering the parameterized post-Newtonian metric [74] outside a spherical static star with a scalar field perturbation given by Eq. (5.24).

5.3 The Chameleon Effect

The nonlinearity of the potential $V(\Phi)$ of the scalar fields means that the effective mass (5.18) is dependent on the value of the Φ field. The cosmological background value of the field generically has a mass of order the Hubble scale. The field is in the perturbative regime when it is near this value of Φ . The perturbation of the field produced by a massive source, assuming it stays in the perturbative regime, is

$$\delta\Phi = \frac{2}{3}\phi_N \quad (5.25)$$

where $\phi_N = GM/r$ is the Newtonian potential of the source. If this perturbation is large enough to push the field to a value where it becomes sufficiently massive, the perturbative approximation breaks down.

The chameleon effect requires the field to become massive and hence short-ranged at large curvatures. If the perturbation of the field reaches the value of Φ where the scalar field becomes massive compared to $1/R_s$ where R_s is a characteristic size of the source, then the perturbative approximation breaks down. When the scalar field becomes very massive, the derivatives in the scalar field equation (5.16) become insignificant. The field “locks in” to a value

$$\Phi_{ad} = \Phi_{ad}(\rho) \quad (5.26)$$

of the field given by the algebraic equation

$$V'(\Phi_{ad}) = \frac{1}{3M_{Pl}^2}\rho. \quad (5.27)$$

This is called the adiabatic regime; the derivatives of the field are no longer important and the field tracks the minimum of an effective potential given by

$$V_{eff}(\Phi) = V(\phi) - \frac{1}{3M_{Pl}^2}\rho \Phi. \quad (5.28)$$

Substituting the definition of the potential, Eq. (5.17), into Eq. (5.28), yields the equation

$$R_{ad} + 2f(R_{ad}) - R_{ad}f_R(R_{ad}) = \frac{\rho}{M_{Pl}^2} \quad (5.29)$$

for the curvature in the adiabatic regime, R_{ad} . We typically consider models where $f(R) \ll R$ and $f_R \ll 1$ when $R \gg H_0^2$. In this case, the field tracks the GR value of $R \approx \rho/M_{Pl}^2$.

If the field becomes short-ranged at some radius inside the star, the mass interior to that radius does not contribute to the scalar field outside. A massive scalar field has a Yukawa potential

$$\Phi = \frac{1}{r} \exp^{-m_\Phi r} \quad (5.30)$$

which is exponentially suppressed at distances larger than $1/m_\Phi$, so matter in the regime where the scalar field has an effective mass $m_\Phi \gg 1/R_s$ does not contribute to the field outside. The effective coupling of the scalar field to the matter in the star is decreased, and so deviations from GR are suppressed. The suppression of the effective coupling to matter is called the chameleon effect. It is this behavior that allows an F(R) theory to pass Solar System tests.

To evaluate whether the chameleon effect works for a specific source, we use the formula for the thickness of the thin shell from [67]. This formula assumes a spherical, static source of constant density, with a different constant density outside. Translated into our notation we have

$$\frac{\Delta R_s}{R_s} = \frac{\Phi_{ad}(\rho_s) - \Phi_\infty}{2GM_s/R_s} \quad (5.31)$$

where ΔR_s is the thickness of the thin shell as compared to the radius of the source R_s ; M_S and ρ_s are the mass and density of the source. $\Phi_{ad}(\rho_s)$ is the

value of the scalar field inside the source, assuming that the field is in the adiabatic regime. Φ_∞ is the value of the scalar field far (relative to the size) from the source. Φ_∞ may be the adiabatic value of the scalar field corresponding to the exterior density. It may instead be the cosmological value as discussed in Section 5.2.3.

The thin shell condition is that

$$\frac{\Delta R_s}{R_s} \ll 1. \quad (5.32)$$

The chameleon effect requires that the thin shell condition be satisfied. If this is the case, and if the scalar field is short ranged at the adiabatic value inside the source ($m(\Phi_{ad}(\rho_s)) \gg 1/R_s$), then the scalar field outside the source is suppressed to

$$\Phi(r) = \frac{2}{3} \frac{\Delta R_s}{R_s} \frac{GM_s}{r} \quad (5.33)$$

rather than the perturbative result

$$\Phi(r) = \frac{2}{3} \frac{GM_s}{r}. \quad (5.34)$$

5.4 Applying the Constraints

We will consider two cases when applying constraints. The first case treating the Sun as if it were embedded directly into the cosmological background, so that we have a single spherical source. The Solar System constraints require two facts. First, the Sun's Newtonian potential is $GM_s/R_s = 2.1 \times 10^{-6}$. Second, the metric of the Sun has been measured quite precisely [74], and the effective coupling of the sun to a long-range scalar field must be less than 2.3×10^{-5} . The thickness of the thin shell then must obey the constraint $\frac{\Delta R_s}{R_s} \leq 2.3 \times 10^{-5}$, or

$|\Phi_{ad}(\rho_s) - \Phi_{infly}| \leq 4.8 \times 10^{-11}$. The example models that we will consider in the next section have values for $f(R)$ that are negative at $R \sim H_0$ and approach zero as R goes to infinity. The density of the sun is very high compared to H_0 , so $\Phi_{ad}(\rho_s)$ is effectively zero. Therefore, we have constrained $|\Phi_\infty| \leq 4.8 \times 10^{-11}$. In this case, $\Phi_{infly} = \Phi_c$ and we have constrained the cosmological value of the scalar field.

The other case that we consider is for the Galaxy to have a thin shell, so that it is in the adiabatic regime. We treat the galaxy as an idealized, uniform density sphere. With an average density of $\rho_g = 10^{-21} \text{kg/m}^3$ and radius of about $R_g = 25 \text{ kpc}$, for the Galaxy $GM_g/R_g \sim 10^{-6}$. The galactic density is about 5 orders of magnitude higher than the cosmological density, and so the value of $\Phi_{ad}(\rho_g)$ is once again negligible compared to Φ_{infly} . We can therefore constrain the cosmological value to be $|\Phi_\infty| \leq 10^{-6}$.

When the scalar field is in the adiabatic regime in the galaxy, if we embed the Sun within the Galaxy, we can treat it as a perturbation to $\Phi_{ad}(\rho_g)$ rather than to Φ_c . In this scenario, $\Phi_{\infty, \text{Sun}} = \Phi_{ad}(\rho_g)$, yielding a constraint on the adiabatic field at galactic densities. Our model must now satisfy both $|\Phi_c| \leq 10^{-6}$ and $|\Phi_{ad}(\rho_g)| \leq 4.8 \times 10^{-11}$. We no longer treat the value of $\Phi_{ad}(\rho_g)$ as negligible. However, once the chameleon is tuned to satisfy the first constraint at a cosmological density, the second constraint is very easily satisfied. The next section will explain why this is, and why these models require fine-tuning.

5.5 Fine-Tuning

To argue that chameleon models require fine-tuning, we will write them in terms of the cosmological curvature today, R_0 . We write the model in the form

$$f(R) = R_0 \bar{f} \left(\frac{R}{R_0} \right), \quad f_R(R) = \bar{f} \left(\frac{R}{R_0} \right). \quad (5.35)$$

We can now write the constraints in dimensionless form. In terms of cosmological parameters, $R_0 = 6(1 - q_0)H_0^2$, where H_0 and q_0 are the Hubble and deceleration parameters respectively. Taking the most probable value from [71], we will use $q_0 = -0.81$.

Our first constraint is $\Phi_c \sim -10^{-6}$, which translates directly into $f'(1) \sim -10^{-6}$. More difficult to deal with is the Friedmann equation constraint (Eq. (5.15)). The Hubble and deceleration parameters are reasonably well-measured, but the jerk is more uncertain. We will simply neglect the term containing f_{RR} . The Friedmann equation becomes

$$\frac{R_0}{6(1 - q_0)} + \frac{1}{6}R_0\bar{f} + q\frac{R_0}{6(1 - q_0)}\bar{f}' = \Omega_M\frac{R_0}{6(1 - q_0)}. \quad (5.36)$$

This produces the constraint $\bar{f}(1) \sim -.4$. We see that introducing the curvature scale R_0 is a natural choice because $|\bar{f}|$ is of order unity. However, this function is not natural because the dimensionless derivative is 6 orders of magnitude smaller. Any function that satisfies both the condition on $\bar{f}(1)$ and $\bar{f}'(1)$ must be fine-tuned.

As an example, there is a class of chameleon models with a transitional behavior from high to low curvatures. At high curvature, f is effectively constant and so it mimics a cosmological constant. At low curvatures, f goes to zero. Hu

and Sawicky [70] consider a model of the form

$$f(R) = -\frac{c_1 H_0^2}{1 + \left(\frac{R}{c_2 H_0^2}\right)^{-n}}, \quad (5.37)$$

where c_1 and c_2 are dimensionless parameters and n is an integer. Starobinsky [76] considers

$$f(R) = c_1 H_0^2 \left(\left(1 + \frac{R^2}{c_2^2 H_0^4} \right)^{-n} - 1 \right). \quad (5.38)$$

We have taken the liberty of altering the notation. Also, we have used the curvature scale H_0^2 to clarify the physical meaning. Both models have tunable parameters n , c_1 , and c_2 . The degree of the power law is controlled by n , the value of the cosmological constant by c_1 , and c_2 controls the value of the curvature at which the transition begins.

We use the same constraints as in the more general case, except that now we use the full Friedmann equation without neglecting f_{RR} . The parameter n is freely specifiable, so we will consider $n = 1$ and $n = 2$. In the Hu and Sawicki model with $n = 1$, $c_2 = 2.8 \times 10^{-5}$. With $n = 2$, we have $c_2 = 1.2 \times 10^{-2}$. The $n = 1$ Starobinsky model coincides with the Hu and Sawicki model with $n = 2$, so we will consider $n = 2$, which gives $c_2 = 3.1 \times 10^{-1}$. In all cases, $c_1 \approx 4.16$.

There is a fine-tuning required for the model to work; the fine-tuning is that the mass scale of the cosmological constant is quite different from the scale at which the transition to zero effective cosmological constant occurs. In our notation, c_1 controls the value of the cosmological constant while c_2 controls the value of R at which a transition occurs to behavior that is distinguishable from a cosmological constant.

As n increases, the tuning becomes seemingly less extreme. However, for none of the models considered is c_2 more than one. This represents a transition

that becomes significant when $R \approx H_0^2$, while today and in the past $R > 9H_0^2$.

5.6 Conclusion

We have reviewed F(R) gravity, and its behavior in both the perturbative and chameleon regimes. The chameleon mechanism is required to hide deviations from General Relativity in Solar System tests. If the Galaxy is assumed to be in the adiabatic regime, these Solar System tests are easily satisfied. However, the requirement that the Galaxy be in the adiabatic regime puts strong constraints of the parameters in the models that we consider. We have made a general argument that any model which satisfies the constraints must be fine-tuned. We gave an example of two models where we found the values of the parameters that had to be fine-tuned for the F(R) theory to be viable. We have also shown that these models are therefore nearly equivalent to a cosmological constant, with corrections that must be tuned to be very small. These corrections define a curvature scale which is at least an order of magnitude smaller (for $n = 1$ and $n = 2$) than the effective cosmological constant. The curvature scale at which the corrections to the effective cosmological constant become important is much lower than the curvature scale of the Universe today. While we have not ruled out these models, this fine-tuning makes them less attractive as an alternative to the cosmological constant.

BIBLIOGRAPHY

- [1] J.D. Brown and J.W. York, Jr., Phys. Rev. D **47**, 1407-1419 (1993)
- [2] C-C.M. Liu and S-T. Yau, Phys. Rev. Lett. **90**, 231102 (2003)
- [3] R.A. Hulse and J.H. Taylor, ApJ Lett. **195** L51 (1975)
- [4] S.M. Carroll, *Spacetime and Geometry* (Addison Wesley, CA) (2004)
- [5] F. Zwicky, Physica Acta **6**, 124 (1933)
- [6] V.C. Rubin, W.K. Ford, and N. Thonnard, ApJ **238**, 471 (1980)
- [7] E. Corbelli and P. Salucci, MNRAS **311**, 441 (2000)
- [8] D. Clowe et.al., ApJ Lett. **648**, L109 (2006)
- [9] M. Kowalski et.al. arXiv:0804.4142v1
- [10] H. Bondi, M.G.J. van der Burg, A.W.K. Metzner, Proc. Roy. Soc. Lond. **A269**, 21-52 (1962)
- [11] R. Arnowitt, S. Deser, C.W. Misner (1962), "The Dynamics of General Relativity", in *Gravitation: An Introduction to Current Research*, edited by L. Witten (New York: Wiley)
- [12] C.W. Misner, Phys. Rev. **130**, 1590-1594 (1963)
- [13] L.D. Landau and E.M. Lifshitz, *The Classical Theory of Fields* (Addison-Wesley, Reading, MA) (1962)
- [14] J.D. Brown, S.R. Lau and J.W. York, Jr., Phys. Rev. D **59**, 064028 (1999)
- [15] H.W. Braden, J.D. Brown, B.F. Whiting, and J.W. York, Phys. Rev. D **42**, 3376-3385 (1990)
- [16] J.W. York, Phys. Rev. D **33**, 2092-2099 (1986)
- [17] L.B. Szabados, *Living Rev. Relativity* **7**, (2004), 4. URL (cited on 05/27/2006): <http://www.livingreviews.org/lrr-2004-4>

- [18] C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (W.H. Freeman and Co., NY) (1973)
- [19] S.R. Lau, *Class. Quant. Grav.* **13**, 1509-1540 (1996)
- [20] I.S. Booth and R.B. Mann, *Phys. Rev. D* **59**, 064021 (1999)
- [21] R.J. Epp, *Phys. Rev. D* **62**, 124018 (2000)
- [22] A. Armenti, Jr., in *Proceedings of the 8th International Conference on General Relativity and Gravitation - GR8, Waterloo, Ontario, Canada*, p. 72 (1977)
- [23] S. W. Hawking and D. N. Page, *Comm. Math. Phys.* **87** no. 4, 577–588 (1983)
- [24] J. Maldacena, *Adv. Theor. Math. Phys.* **2** 231 (1998); hep-th/9711200
- [25] E. Witten, *Adv. Theor. Math. Phys.* **2**, 505–532 (1998) hep-th/9803131
- [26] G. 't Hooft (1993), gr-qc/9310026
- [27] A. Chamblin, R. Emparan, C. V. Johnson, and R. C. Myers, *Phys. Rev. D* **60** 064018 (1999); hep-th/9902170
- [28] A. Chamblin, R. Emparan, C. V. Johnson, and R. C. Myers, *Phys. Rev. D* **60**, 104026 (1999); hep-th/9904197
- [29] S. Carlip and S. Vaidya, *Class. Quant. Grav.* **20**, 3827–3838 (2003)
- [30] G. Gibbons and S. W. Hawking, *Phys. Rev. D* **15**, 2752–2756 (1977)
- [31] J. W. York, *Phys. Rev. D* **33**, 2092–2099 (1986)
- [32] B.F. Whiting and J.W. York, *Phys. Rev. Lett.* **61**, 1336-1339 (1988)
- [33] A. Abramovici et. al., *Science*, **256**, 325-333 (1992).
- [34] Advanced LIGO.
<http://www.ligo.caltech.edu/advLIGO>
- [35] E. D'Ambrosio, R. O'Shaughnessy, S. Strigin, K. S. Thorne and S. Vyatchanin, [arXiv:gr-qc/0409075].

- [36] R. O’Shaughnessy, S. Strigin and S. Vyatchanin, [arXiv:gr-qc/0409050].
- [37] J. Agresti, Y. Chen, E. D’Ambrosio and P. Savov, [arXiv:gr-qc/0511062].
- [38] P. Savov and S. Vyatchanin, Phys. Rev. D **74**, 082002 (2006). [arXiv:gr-qc/0409084]
- [39] E. D’Ambrosio, Phys. Rev. D **67**, 102004 (2003).
- [40] M. Bondarescu, Ph.D. Thesis, California Institute of Technology.
<http://resolver.caltech.edu/CaltechETD:etd-05282007-231321/>
- [41] J. Sidles and D. Sigg, LIGO Report Number P030055-B (2003).
www.ligo.caltech.edu/docs/P/P030055-B.pdf
- [42] M. Bondarescu and K. S. Thorne, Phys. Rev. D **74**, 082003 (2006). [arXiv:gr-qc/0409083].
- [43] M. Tarallo, J. Miller, J. Agresti et.al., Appl. Opt. **46**, 6648-6654 (2007).
- [44] M. Tarallo, Thesis (Tesi di Laurea Specialistica), Universita di Pisa.
<http://etd.adm.unipi.it/theses/available/etd-09092005-145828/>
- [45] G. Lovelace, Class. Quantum Grav. **24**, 4491 (2007). gr-qc/0610041.
- [46] R. O’Shaughnessy, Class. Quantum Grav. **23** 7627 (2006).
- [47] Y. Levin, Phys. Rev. D **57**, 659 (1998).
- [48] Y. Liu and K. Thorne, Phys. Rev. D **62**, 122002 (2000).
- [49] J. Agresti and R. DeSalvo, LIGO report LIGO-T050269-00-R
<http://www.ligo.caltech.edu/docs/T/T050269-00/T050269-00.pdf>
- [50] V. Pierro, V. Galdi, V. Castaldi, I. Pinto, J. Agresti, and R. DeSalvo,
[arXiv:0707.0579]
- [51] M. Bondarescu, Y. Chen and O. Kogan, in preparation.
- [52] P. Barriga, B. Bhawal, L. Ju, D. Blair, J. Opt. Soc. Am. A **24**, 1731 (2007).

- [53] V. Braginsky, S. Strigin, S. Vyatchanin Phys. Lett. A **287** 331 (2001). [arXiv:gr-qc/0107079]
- [54] V. Galdi, G. Castaldi, V. Pierro, I. M. Pinto, J. Agresti, E. D'Ambrosio and R. DeSalvo, Phys. Rev. D **73**, 127101 (2006). [arXiv:gr-qc/0602074].
- [55] M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions", Handbook of Mathematical Functions, New York: Dover, 1972.
- [56] O. Miyakawa et.al., Phys. Rev. D **74**, 022001 (2006).
- [57] B. Saleh and M. Teich, "Fundamentals of Phototonics," John Wiley and Sons, USA, 1991.
- [58] J. Boyd, "Chebyshev and Fourier Spectral Methods," Dover, NY, 2001.
- [59] V. Braginsky, S. Strigin, S. Vyatchanin, Phys. Lett. A **305** 111 (2002). [arXiv:gr-qc/0209064]
- [60] C. Zhao, L. Ju, J. Degallaix, S. Gras, D. Blair Phys. Rev. Lett. **94** 121102 (2005). [arXiv:gr-qc/0502079]
- [61] V. Braginsky, A. Gurkovsky, S. Strigin, S. Vyatchanin, Phys. Lett. A **362** 91 (2007). [arXiv:gr-qc/0608007]
- [62] GNU Scientific Library:
<http://www.gnu.org/software/gsl/>
- [63] IT++: <http://itpp.sourceforge.net/>
- [64] AMD Core Math Library: <http://www.amd.com/acml>
- [65] S.M. Carroll et.al., Phys. Rev. D **70**, 043528 (2004)
- [66] T. Chiba, Phys.Lett. **B 575**, 1 (2003)
- [67] J. Khoury and A. Weltman, Phys. Rev. Lett **93**, 171104 (2004)
- [68] J. Khoury and A. Weltman, Physical Review D **69**, 044026 (2004)
- [69] P. Brax et.al., Physical Review D **70** 123518 (2004)

- [70] W. Hu and I. Sawicki, Phys. Rev. D **76**, 064004 (2007)
- [71] D. Rapetti et.al., MNRAS **375**, 1510 (2007)
- [72] T. Faulkner et.al., Phys. Rev. D **76**, 063505 (2007)
- [73] I. Navarro and K. Van Acoleyen, JCAP **0702**, 022 (2007)
- [74] C.M. Will, Living Reviews in Relativity **9**, 3, (2006)
URL (cited 07/2008) <http://www.livingreviews.org/lrr-2006-3>
- [75] G. Hinshaw et.al., arXiv: astro-ph/0803.0732
- [76] A.A. Starobinsky, JETP Lett. **86**, 157 (2007)