Effective Field Theories for Gravity

Matthew Dean Schwartz

A Dissertation

Presented to the Faculty

of Princeton University

in Candidacy for the Degree

of Doctor of Philosophy

Recommended for Acceptance

by the Department of

Physics

June, 2003

Abstract

Nobody knows how to incorporate the classical theory of gravity and the quantum field theory which describes the Standard Model into a single consistent framework. We can, however, study gravity and the Standard Model together with techniques of self-consistent effective field theory, as long as we are content to probe distances larger than the Planck length, 10^{-33} cm. In this dissertation, I demonstrate a number of predictions coming from the effective field theory approach in two broad areas. First, I show that in curved spaces, gravity can be regulated in a position-dependent way. This regulator is used to understand the thermodynamics of black holes and other gravitational systems, as well as the unification of the Standard Model couplings in certain phenomenologically interesting extra-dimensional models. The second part of this dissertation is concerned with consequences of breaking general coordinate invariance. I provide a simple technique for studying massive gravitons, and use it not only to explain a number of peculiar features of massive gravity in a transparent way but also to pinpoint the reason that models with discrete gravitational dimensions are difficult to construct. These effective field theory investigations leave leave a number of clues about what properties a fundamental theory of gravity might possess.

Acknowledgements

First and foremost I am indebted to my advisor, Lisa Randall, for her confidence in me and her constant support. After only working with me for a few months in Princeton, she graciously invited me to accompany her to MIT and to Harvard. I thereby had the privilege to work with three different departments, and profitted from an exposure and experience that graduate students are rarely offered. Working with Lisa has been a tumultuous pleasure, and I am very proud to be her student.

I have also had the good fortune to work with and learn from Nima Arkani-Hamed. Our collaboration, especially in the summer of 2002, provided some of the most productive and enjoyable experiences of my graduate career. I wish to thank Nima for his patience and encouragement and for profoundly influencing the way I think about physics.

I would also like to thank Howard Georgi for his enthusiasm and support. In addition, my education was immeasurably enhanced by discussions with the students, post-docs and faculty at Princeton, MIT and Harvard.

Finally, I am happy to acknowledge that I could never have gotten anywhere without the love and support of my friends and family.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Einstein's theory of gravity, general relativity, is a beautiful and tremendously successful classical theory. It makes detailed predictions about astrophysical phenomena beyond the Newtonian approximation, and well beyond what experimental technology can now, or is ever likely to explore. However, general relativity does not make predictions to infinite precision, the ultimate requirement of a complete fundamental theory. When integrated with quantum mechanics, GR is perturbatively non-renormalizable. This means, essentially, that we would need an infinite number of different experiments to predict the result of any one experiment to infinite accuracy. Moreover, it does not make predictions *at all* for experiments conducted at energies higher than the Planck scale $M_{\mathrm{Pl}} = 10^{33}$ cm^{-1}. While it is possible that the universe is described by such an elegant but marginally predictive theory, it is far more likely that our own understanding is too primitive. It is likely that entirely new ideas are needed to reconcile quantum field theory and gravity, for example, those provided by string theory. But even string theory is not predictive at high energy. Certainly, trying to guess the fundamental theory of gravity is a promising approach, but there is actually a lot that can be learned about quantum gravity from the approximation we already have. The key tool is effective field theory, which provides a way to sequester the problems associated with non-renormalizability, in a framework that is self-consistent at low energy.

1

Briefly, effective field theory is a way of parameterizing our ignorance of high energy physics. In any physical process, because of the uncertainty principle, virtual particles are created with arbitrarily high energy for an increasingly short period of time. If we modify the theory by cutting off the energy of these virtual particles at some fixed scale Λ, we should hope that the effects of the cutoff go away as Λ is taken to ∞. If this is true for all physical processes, the theory is said to be renormalizable. But even if a theory is not renormalizable, experiments conducted at low energy E should not depend on details of the high energy theory. It is a natural and apparently universal property of physical systems that the high energy/short distance details can be coarse-grained into a set of effective parameters relevant to the energy scales we probe. Corrections to this approximation are suppressed by factors of E/Λ simply by dimensional analysis: they grow as E approaches the coarse-graining scale and shrink as the coarse-graining becomes more fine. The point of effective field theory is that one does not actually need to know a fundamental theory to study its low energy approximation. As long as we include all the E/Λ effects which cannot be forbidden, the effective theory can accommodate all the complexity of the fundamental theory. The amazing thing is that symmetries can, and often do, prevent many bad effects from appearing. And the remaining effects may leave valuables clues as to what fundamental theory the effective theory is derived from. The purpose of this dissertation is to explore what we might learn about gravity from the effective field theory approach.

There are two basically distinct parts to this dissertation. The first is concerned with the consequences of applying effective field theory in a fixed gravitational background. Of course, in flat space, gravity is just a non-renormalizable theory. Because Einstein's theory is geometric, only geometric objects are generated in the effective theory, and these are suppressed by powers of E/M_{Pl}. In curved space, however, the effective theory is much more interesting – the cutoff is no longer strictly M_{Pl}, it must vary with position. In Chapter 2, I show how to implement this idea in a consistent way. It turns out that more of the space is excluded at high energy than at low energy. One of the consequences of this is that number of degrees of freedom will not scale with volume, and the entropy will not

be an extensive quantity. This is in agreement with semi-classical results about black holes and (anti) de Sitter space, which predict entropy to scale with the area of the boundary.

In Chapter 3 I apply this position dependent cutoff to a class of phenomenologically relevant constructions called Randall-Sundrum (RS) models. There are many felicitous features of RS, which I will discuss in Chapter 3, but one of the big drawbacks of RS had always been the apparent conflict with Grand Unification, which I will also explain. In general, RS makes good use of a strongly curved extra dimension. One of its inevitable predictions is that in the region where matter resides, gravity seems to become strong at a scale well below M_{Pl}. This allows RS to justify, in a natural way, the exponential factor in the scale of electroweak interactions: TeV $\sim 10^{-16} M_{Pl}$. That gravity is strong at a low scale is not actually in conflict with any experiment; but it seems to preclude one of the only hypotheses for extending the Standard Model which has any connection to experiment: that the various couplings in the Standard Model vary with energy and seem to unify at $\sim 10^{13}$ TeV. It was thought that because gravity is strong at $\sim$ TeV one cannot trust the unification calculations. It turns out that because the electroweak scale is well below the *five*-dimensional Planck scale, the unification calculation can be performed in a consistent effective field theory framework. Using the techniques of Chapter 2 I show that, somewhat remarkably, RS predicts Grand Unification at least as well as the Standard Model.

The second part of this dissertation uses effective field theory to characterize models where the general coordinate symmetry of classical gravity is explicitly broken. As I mentioned above, symmetries of a theory can be essential in controlling its low energy approximation. But because GR is non-renormalizable, it is only predictive in a restricted range $(E < M_{Pl})$ and so breaking its symmetries will just result in a lower cutoff. No serious damage is done to the theory, because it is not fundamental to begin with. In fact, from the phenomenological point of view, we should not assume GR is valid all the way up to $E \sim M_{Pl}$, when experiment has only tested gravity up to $E \sim$ mm$^{-1} \sim 10^{-30} M_{Pl}$. It is *a priori* doubtful that all theories with weakly broken general coordinate invariance will be incompatible with experiment. And so it is useful to have a system for classifying the

acceptable effects.

I begin in Chapter 4 by looking at what happens when we give the graviton a mass. The mass breaks general coordinate invariance, but scattering amplitudes are still well behaved at low energy. To understand the interactions of the massive graviton, an effective field theory formalism is introduced whose main ingredient is a set of Goldstone bosons for the broken symmetry. Massive gravitons had been studied before, but only classically, or at the quantum level without interactions. The effective field theory formalism proves to be remarkable, in that it provides a transparent explanation of many peculiar results that were around in the literature, but came from long and physically obscure derivations. For example, I will show how to understand the Fierz-Pauli form of the propagator, the macroscopic Vainshtein radius, and the van Dam-Veltman-Zakharov discontinuity with hardly any calculational effort. The most important result, which is completely new, is that interactions of the Goldstones, which represent the longitudinal modes of the massive graviton, get strong at a scale $\Lambda_5 = (m_g^4 M_{\rm Pl})^{1/5}$. This scale is well below $M_{\rm Pl}$, where the massless theory breaks down, but parametrically above the mass of the graviton, m_g.

The work on effective field theory for massive gravitons was inspired by the deluge of phenomenological applications coming from a new understanding of discrete extra dimensions in the gauge theory context. It has been shown that an $SU(k)$ gauge theory, which in five dimensions is non-renormalizable, can be written as $SU(k)^N$ gauge theory in four dimensions with controlled $SU(k)$ violating interactions. The large 4D theory can then be renormalized with the Higgs mechanism to produce an effective UV completion of the 5D theory. Moreover, parameters can be chosen so the 5D theory satisfies any phenomenological constraints. If the same trick could be used to write a 4D theory of gravity as a 3D theory with a large number of 3D gravitons, perhaps a similar UV completion could be derived. After all, gravity in 3D is technically renormalizable, as the graviton simply does not propagate. Thus, ultimately, we might hope to establish a renormalizable theory of gravity by defining it on a space-time lattice.

To pursue this idea, in Chapter 5 I apply the formalism from Chapter 4 to construct-

ing discrete gravitational dimensions. Immediately, we see that the problem is much more complicated than in the gauge theory case. Because the discretization breaks the general coordinate symmetry of the higher dimensional theory, the continuum limit may look nothing like the higher dimensional continuum. For example, if the lattice only has nearest neighbor interactions, the strong coupling at Λ_5 causes the theory to break down at energies well below the higher dimensional Planck scale M_{5D}. This constrains the way we approach the continuum. It turns out that any continuum limit of such a theory will appear highly non-local in the extra dimension. Of course, effective field theory guarantees that we really can mimic an extra dimension by simply ignoring its high energy degrees of freedom. I proceed to show that if we try to interpret such a truncated theory as a discrete dimension, it has highly non-local interactions on the lattice itself. Moreover, in any of these discretizations, the strong coupling scale depends on the size of the new space we are constructing. Thus we are brought back the UV/IR connection unearthed in Chapter 2.

As far as phenomenology is concerned, the Goldstone boson formalism of Chapter 4 is powerful enough to construct many realistic models. For example, practically all of the extensions of the Standard Model which come from extra dimensions, such as RS, can be completely mimicked by a model with a few sites. However, if we believe that the universe really is extra-dimensional, perhaps because of string theory, then we need to know which low-energy effective theories might really come from compactifications. This is the topic of Chapter 6. Chapter 6 actually ties up a number of loose ends from the two previous chapters. It includes a characterization of all non-linear Lagrangians for a massive spin two field and a calculation of tree-level scattering amplitudes in the truncated theory of Chapter 5. One cannot help but conclude that effective field theory is the key tool for determining what more fundamental physical system our low energy experiments may be describing.

Chapter 2

Effective Field Theory in Cuved Space

2.1 Introduction: semi-classical gravity and holography

Practically all the clues we have about quantum gravity come from semi-classical investigations. The general approach is to quantize fields on a fixed gravitational background and study the resulting quantum field theory at tree level. It is often possible to include the back reaction of the field energy on the classical background without animating gravity completely. While such a theory is not really self-consistent, as gravity and other fields are treated differently, it may provide clues of underlying principles which may guide our search for a fully consistent theory. For example, semi-classical investigations of Bekenstein [18, 19, 20. 21] and Hawking [58, 17, 59] in the 1970's led to the concept of black hole entropy and eventually the notion of space-time holography [103, 97]. It is now generally believed that any fundamental theory of gravity will have holography built in. First I will give a brief introduction to the subject, and then explain my own work which partially reconciles effective quantum field theory with the holographic principle.

At the classical level, black holes satisfy laws which have a precise and direct analogy to the ordinary laws of thermodynamics [17]. The surface gravity of a black hole is like

temperature. the mass of a black hole is like internal energy and the area of the event horizon is like entropy. This analogy became much more than a formal correspondence in 1975 when Hawking showed [59], through a semi-classical calculation. that black holes really do emit thermal radiation. at precisely the temperature required. So. the temperature of a black hole is physical, and therefore the entropy. the Bekenstein-Hawking entropy, must be physical as well. But what possible statistical explanation could this entropy possibly have? First of all. there are theorems which prove that black holes can be described by only a handful of parameters. So their entropy, which according to standard statistical arguments. is the logarithm of the number of degrees of freedom. should be close to zero. In contrast, the Bekenstein-Hawking entropy is huge. Perhaps more troubling is the fact that the entropy of a black hole scales with area, not volume. so it is not extensive. This is also in contradiction with our expectations from statistical mechanics. Nevertheless, there is some structure to black hole thermodynamics, and this structure has inspired the "holographic principle". Like a hologram, black holes seem to encode information about a three dimensional space with the degrees of freedom of a two dimensional surface.

The results about black holes actually imply constraints on more general gravitational systems. In non-gravitational physics, one can have a box of fixed surface area $\mathcal{A}$ enclose an arbitrary amount of entropy. If this entropy were to exceed the entropy of a black hole of the same area, one could compress the matter into a black hole. and then the total entropy would decrease. Thus, when gravity is included, the entropy of *any* system should be bounded by its area in Planck units:

$$S \leq \frac{\mathcal{A}}{4\pi} M_{\text{Pl}}^2 \qquad (2.1)$$

To support for this conjecture, a number of semi-classical systems have been shown to satisfy (2.1). For example, de Sitter and anti de Sitter space also seem to have entropies which scale as their areas.

The general conclusion of semi-classical investigations of thermodynamics in curved space, as summarized by the holographic principle, is that the number of degrees of freedom of a space are determined by a space of one lower dimension. This lower dimensional space

is often taken to be the boundary [104, 22]. Indeed, in the one example where holography is known to hold, through the so called AdS/CFT correspondence [74], the degrees of freedom really are on the boundary. The correspondence, which has its support in some string theory calculations [74, 115, 56], states that a five-dimensional field theory on anti de Sitter space (AdS) is dual to a four-dimensional conformal field theory (CFT). The background for the CFT is the boundary of the AdS space. However, in general, we cannot expect a gravitational system to exhibit the holographic principle so cleanly. This chapter will explore how far one can get by just studying the higher dimensional gravitational theory, independent of whether a consistent boundary theory can be constructed.

The basic idea is that in a curved background, the size of the metric and hence the Planck scale varies with position. So the scale where new physics comes in, and where the energy cutoff must be imposed in an effective theory, is lower in some places and higher in others. In a very direct way, effective field theory dictates that we must use a position-dependent regularization procedure. When we use this regulator to calculate the entropy, we will find that the number of effective degrees of freedom are in accordance with the predictions of holography and are consistent with the Bekenstein formula (2.1). That is, the degrees of freedom fit on a space of one lower dimension. The main purpose of this chapter is to demonstrate that by eliminating the region of space-time where energies are above a local cutoff scale, we can understand intuitively why and how a theory becomes holographic, in the sense of having degrees of freedom reflecting a space of reduced dimensionality. It will be shown for global AdS, de Sitter space, and black holes.

Interpreting a cutoff on local energy as a cutoff on position is an example of a UV/IR correspondence. In many holographic situations, there seems to be a relationship between the high energy, or ultraviolet (UV), degrees of freedom and the low energy, or infrared (IR), degrees of freedom. This idea has yet to be turned into a precise theorem or principle, it is more like a common thread which seems to show up in various quantum gravity investigations. I will mention it a number of times in this dissertation. It appears in this chapter as a result of the self-consistent and logical assumptions of effective field theory.

It is not hard to see how the position dependent regulator gives a simple origin for the reduction of degrees of freedom. By forcing local energy, $\sqrt{g^{tt}}E$, to be less than a cutoff Λ, the high energy modes will in general probe less of the space than the low energy modes. The spatial cutoff implements the correct counting. It reflects the physical fact that at high energy, there are regions of space that are inaccessible, so effectively the boundary depends on energy. This reduction of volume at high energy can have the same effect on thermodynamic quantities as confining the theory to a subspace, for example. the boundary. There is not necessarily a boundary theory; however, in certain cases (when a horizon exists for example), the physics is dominated by the degrees of freedom concentrated near this region.

The method that will be described is quite distinct from one that postulates the existence of a boundary theory. In both approaches. we find in certain known examples that the number of degrees of freedom scales with area, not volume. In our description, the degrees of freedom reside throughout the space, but might be peaked on the boundary. In boundary description the degrees of freedom are fundamental degrees of freedom defined on a holographic screen. Our description only involves bulk degrees of freedom. We are not taking the holographic correspondence as an assumption. We simply note that the degrees of freedom reflect a theory of lower dimension, but we do not explicitly postulate such a theory.

At this point, I should point out that this procedure only works in a curved background. Clearly, this is not sufficient to derive area-law scaling for all systems. For example, it says nothing about flat space. Moreover, it does not incorporate any back-reaction. This is justified by the fact that we count only those states that can be properly treated with low-energy field theory. We must exclude states for which the back-reaction would alter the gravitational background considerably. such as those with cutoff sized black-holes. This is not quite consistent, as we cannot eliminate all possible configurations that can form a black hole or have some other strong gravitational back-reaction. But it should account for the vast majority of gravitational phenomena that can occur. Still, we sometimes find

certain parameter regimes that reflect the full dimensionality of the space.

Although it is not the whole story, I believe this simple procedure provides some guidance towards a more comprehensive understanding of holography, and of quantum gravity. It also suggests that when the metric is properly accounted for, there is not necessarily redundancy in the low-energy description which holography seems to imply. It should also be kept in mind that a cutoff really refers to a more fundamental theory which may differ significantly from the effective theory at high energy. The fact that our answers are cutoff dependent tells us how the entropy associated with the cutoff should scale with size of the system. Because the local cutoff changes with the curvature of the space, it suggests that degrees of freedom that are best thought of as free fields in some contexts are strongly bound states in others. This work tells only about the field theoretical contribution, which is in general less than the full counting.

One advantage of this method of counting, even when there exists a precise holographic description as in the AdS example, is that it indicates the wave function for states in the bulk theory associated with the holographic dual on the boundary. We will see this explicitly for global AdS where we can see why the Bekenstein bound applies. Another advantage of our method is that it regulates a theory to give finite entropy, even in the presence of a horizon, in a way well motivated by the physics.

The organization of this chapter is as follows. We will start by briefly presenting the regularization as it applies to the Randall Sundrum model. This will be discussed in detail in Chapter 3. In Section 2.3, we present the techniques we use for counting degrees of freedom or calculating entropy. We show that the answer will be the answer one should expect, the integrated local energy over the proper volume, where only states consistent with the local cutoff are permitted. We also demonstrate the relation between our cutoff procedure and Pauli-Villars, which it closely resembles. Section 2.4 applies our methodology to global anti-de Sitter space, where we again see the reduced dimensionality from our simple procedure. Sections 2.5 and 2.6 explore de Sitter space and black hole space-times. We will see that the local UV cutoff obviates the need for 't Hooft's brick wall cutoff. In the following section,

we speculate about the non-covariant nature of our result; in particular, why it might not entirely account for all states in time-dependent space-times. Finally, in section 2.8, we summarize our results and their implications. The rest of this chapter follows very closely to [86].

2.2 Poincare Patch AdS

Before looking at the spatially-varying cutoff in general space-times, we explain how it applies in the two brane Randall Sundrum model (RS) [89]. This model is fairly well understood, phenomenologically viable, and believed to be holographic [13, 90]. I will explain its merits in the Chapter 3. For now, I will just present its general features.

The background geometry of RS is the Poincare patch of 5D anti-de Sitter space with metric:

$$ds^2 = \frac{1}{(kz)^2}(-dt^2 + dx^2 + dz^2) \tag{2.2}$$

The bulk cosmological constant is $-3k^2$ which we assume to have a large magnitude, of order the Planck scale. In RS, the AdS horizon is cutoff by a Planck brane at $z_0 = \frac{1}{k}$ and a TeV brane at $z_1 = \frac{1}{T}$, where T is a mass scale of order 1 TeV. It is clear from (2.2) that the induced geometry at any fixed z is flat. It is also not hard to show that at low energies the bounded fifth dimension can be integrated out to get an effective theory which is flat as well.

Let's suppose we have a scalar field which is free to propagate in the bulk. A simple hypothesis is that the total number of degrees of freedom of such a field is given roughly by the volume of the space normalized with some UV cutoff Λ [22]:

$$g(\Lambda) \approx \Lambda^4 \int \sqrt{-G} d^3 x dz \approx \Lambda^4 L^3 \int_{1/k}^{1/T} \frac{1}{(kz)^4} dz \approx \Lambda^3 L^3 \Lambda(\frac{k^4 - T^4}{k^5}) \approx \Lambda^3 L^3 \frac{\Lambda}{k} \tag{2.3}$$

For large $k \approx \Lambda$, this looks like a 4D system.

Now, suppose we tried to count the degrees of freedom by adding up the Kaluza-Klein (KK) modes of the bulk field. The massless field in 5D can be decomposed into a set of 4D

modes with masses given by roughly $m_n = jT$ for integer j. Then each mode satisfies:

$$E^2 = p^2 + m_j^2 \tag{2.4}$$

where p is the 3-momentum. The number of states with energy less than E, $g(E)$, is given by:

$$g(E) = \int_0^{E/T} dj L^3 \int_0^{\sqrt{E^2 - (jT)^2}} p^2 dp \sim E^3 L^3 \frac{E}{T} \tag{2.5}$$

So $g(\Lambda) = \Lambda^3 L^3 \frac{\Lambda}{T}$. This superficially has the same form as (2.3), but we expect $\Lambda \approx k \gg T$. Since $\frac{1}{T}$ is the size of the fifth dimension, this wrongly implemented Kaluza Klein picture makes it seem like there are 5D degrees of freedom.

These two results are different because the volume calculation is scaling the cutoff with position, while the KK calculation, as presented above, is not. Indeed, the contribution to the volume at a position z is not $\Lambda^4 L^3 dz$ but $\frac{\Lambda^4 L^3}{k^5 z^5} dz$.

Historically, we first introduced the position-dependent regulator in [87, 88] to study unification in RS. This will be the subject of Chapter 3. Basically, by calculating Feynman diagrams in 5D with this regulator, we showed that gauge couplings in the 5D bulk run logarithmically, that is as in 4D, and that perturbative unification is feasible at a high scale in AdS$_5$ What we want to emphasize here is not the details of the calculation, but why the result is to be expected with a generally covariant cutoff in place.

When we did a field theory calculation in AdS$_5$, we did not use explicit KK modes, but did our calculation in the five-dimensional space with a mixed position space/momentum space formulation, integrating over p_μ and z. We applied a cutoff on momentum p that varied with position in accordance with the warp factor, that is $p(z) < \Lambda$ or equivalently, $p < \Lambda(z)$ where

$$\Lambda(z) = \frac{1}{kz}\Lambda \tag{2.6}$$

This is the spatially-varying cutoff on energy; the UV cutoff on p depends on position (see Figure 2.1). The highest cutoff is on the Planck brane, where $\Lambda(\frac{1}{k}) = \Lambda$. On the TeV brane, p can only go up to $\Lambda \frac{T}{k} \ll \Lambda$. Because $p < \frac{1}{kz}\Lambda$ is the same constraint as $z < \frac{\Lambda}{kp}$ and $p < \Lambda$, the position-dependent cutoff is equivalent to a usual cutoff on 4D momenta
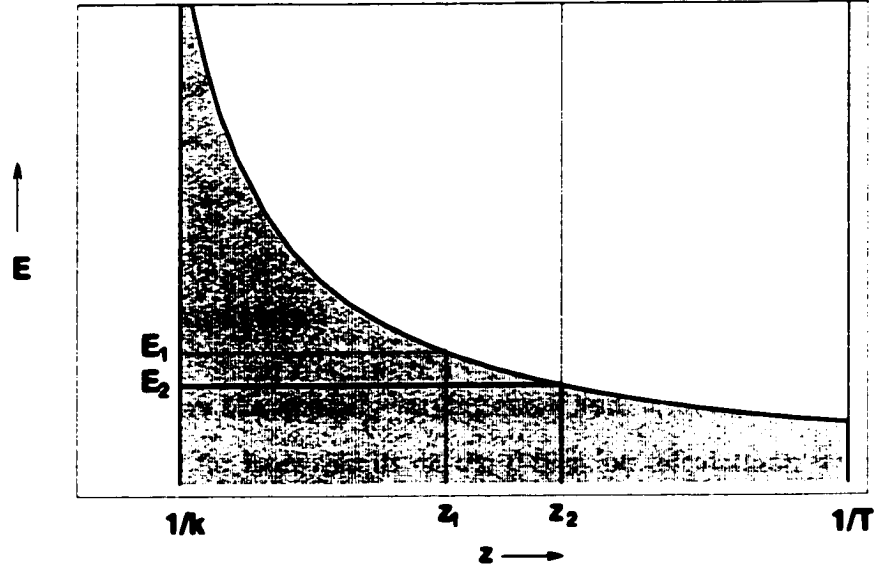
Figure 2.1: In RS model, an IR boundary at $z = 1/T$ is effectively replaced for energies $E > \frac{\Lambda}{k}T$. The new boundary, shown for two modes with energies E_1 and E_2, is at $z_{1(2)} = \frac{\Lambda}{k}\frac{1}{E_{1(2)}}$. The area of the (E, z) plane satisfying $E < \Lambda(z)$ is shaded.

and an energy-dependent limit on z. By taking this constraint on z as the new boundary of the system, one can readily see that the number of states is drastically reduced and agrees with the holographic expectation.

We emphasize that there are two aspects to this procedure. First, there is a spatial-cutoff for a given energy. Second there is a re-quantization reflecting this spatial cutoff for each energy. For this RS Poincare patch example, we get the right counting even without moving the boundary. However, in general, we have to move the boundary explicitly for each energy to reproduce what we expect from the more detailed mode analysis. This was the actual procedure used in [87, 88], which be discussed at length in Chapter 3.

Not only does the above calculation demonstrate 4D behavior, it shows that there is an intermediate regime that appears to be 5D [13]. The cutoff is position dependent only for $p > \frac{\Lambda}{k}T$. At lower energies, the cutoff is fixed at $z = \frac{1}{T}$. In this regime, the constraint on z would have implied a brane deeper in the IR than the TeV brane, so it is irrelevant. At very low energies $p < T$, no KK modes are excited and the theory is 4-dimensional:

$g(E) = (EL)^3$. If $T < p < \frac{\Lambda}{k}T$ then $E^2 = \mathbf{p}^2 + j^2T^2$. Therefore in this regime, one obtains the usual 5D dispersion relation, and $g(E) = \frac{E}{T}(EL)^3$. When $E/T > \Lambda/k$ we are in the holographic regime. In summary, the theory appears to be 4D for $E < T$, 5D for $\frac{\Lambda}{k}T > E > T$ and 4D for $E > \frac{\Lambda}{k}T$.

2.3 Other Geometries

Now that we have seen how the spatially-varying cutoff leads to holographic thermodynamics in RS (see Chapter 3 for more details on the β-function calculation), we generalize to other geometries. For time-independent metrics, the killing vector ∂_t allows us to assign an energy E to any state. As we justified above, this energy should never be greater than the locally measured cutoff $\Lambda(r) = \sqrt{g_{tt}(r)}\Lambda$. The state with energy E should only probe the region where $\Lambda(r) \geqslant E$. So the boundary r_E at an energy E is determined by:

$$E = \sqrt{g_{tt}(r_E)}\Lambda \tag{2.7}$$

The E-dependent cutoff on r will determine the quantization of momentum in the r direction, and hence the quantization of E itself. In some cases, one can derive details about the spectrum; however, even without the precise spectrum, we can evaluate approximately the density of states.

In the examples below, we focus on metrics which take the form

$$\mathrm{ds}^2 = -V(r)\mathrm{d}t^2 + \frac{1}{V(r)}\mathrm{d}r^2 + r^2 d\Omega^2 \tag{2.8}$$

Some examples are global anti-de Sitter space $V(r) = 1 + k^2 r^2$, static de Sitter $V(r) = 1 - k^2 r^2$, and the exterior of Reissner-Nordstrom black holes $V(r) = 1 - \frac{2m}{r} + \frac{q^2}{r^2}$. This is an interesting class of metrics in that they are examples in which area laws are known to apply. Furthermore, as we will discuss, they all manifest a UV/IR correspondence.

Consider the covariant Klein-Gordon equation:

$$\frac{1}{\sqrt{g}}\partial_\mu(\sqrt{g}g^{\mu\nu}\partial_\nu\phi) = M^2\phi \tag{2.9}$$

We first consider states in the 4D system. In the background (2.8). and with the ansatz $\phi = \varphi(r)Y_{lm}(\theta.\phi)e^{iEt}$. (2.9) becomes:

$$\frac{E^2}{V(r)}\varphi + \frac{1}{r^2}\partial_r(r^2V(r)\partial_r\varphi) - \frac{l(l+1)}{r^2}\varphi - M^2\varphi = 0 \qquad (2.10)$$

This is just a 1D quantum mechanics problem. The spatially-varying cutoff would enter through the boundary conditions which will depend on E.

With the exact spectrum, one can evaluate the density of states. Even without the exact spectrum, one can often use the semi-classical (WKB) approximation [102]. Alternatively, one can evaluate the density of states based on the local energy and proper distance in accordance with the metric. We will see that both these last two methods give yield the same formula for calculating the density of states.

WKB applies when the phase of the wavefunction changes much faster than the amplitude. We can then write $\varphi = \sqrt{\rho}e^{iS(r)}$ and assume that $k(r) = S'(r)$ is large. This allows us to solve (2.10) implicitly (for $M = 0$):

$$k(r.E.l) = \frac{1}{V(r)}\sqrt{E^2 - \frac{V(r)}{r^2}l(l+1)} \qquad (2.11)$$

The number of oscillations of the phase of φ over the whole space is the number of nodes of the approximate wavefunction. This is roughly the number of modes with energy less than E. Indeed, as E is lowered to zero, each of these nodes should disappear: whenever a node hits the endpoint, there is another state. Thus the number of states with energy less than E is given by

$$g(E) = \int dr \int k(r,E,l)(2l+1)dl = \int \frac{dr}{V(r)} \int (2l+1)dl\sqrt{E^2 - \frac{V(r)}{r^2}l(l+1)} \qquad (2.12)$$

The integral over l is for values of l which keep k positive, that is $l(l+1) < \frac{r^2E^2}{V(r)}$. This gives:

$$g(E) = \frac{2}{3}E^3 \int \frac{r^2}{V(r)^2}dr \qquad (2.13)$$

Note that $g(E)$ already seems to have 4D E dependence. So, unless the r integral depends on E, we will have 4D thermodynamics. In our regularization, we limit r by $E < \sqrt{V(r_E)}\Lambda$.

This will add E-dependence to the r integral. For example, if $V(r_0) = 0$ at a horizon, then the integrand will have a pole. The only modes that can probe up to the horizon have $E = 0$. so the pole is an $E = 0$ pole and must change the E dependence of the density of states for low energy. We will show this explicitly in later sections.

In fact. (2.13) is precisely the answer we expect in four dimensions once the metric is properly accounted for. That is. for an n-dimensional space, we expect the number of degrees of freedom to be

$$
\begin{aligned}
g(E) &\approx \int^{r_E} \left(\sqrt{g^{tt}}E\right)^{n-1} r^{n-2}\sqrt{g_{rr}}dr \\
&\approx \int \frac{E^{n-1}}{V(r)^{\frac{n}{2}}}r^2 dr
\end{aligned}
\tag{2.14}
$$

One can partially understand the relation between the position-dependent cutoff and a Pauli-Villars regulator by examining the form of the equation of motion. Since we are concerned with time-independent metrics, we write $\phi = \varphi(x)e^{iEt}$ and the Klein-Gordon equation (2.9) becomes

$$
g^{tt}(x)E^2\varphi + \frac{1}{\sqrt{g}}\partial_i(\sqrt{g}g^{ij}\partial_j\varphi) = M^2\varphi
\tag{2.15}
$$

The Green's function for a quantum field and for its PV regulator at an energy E will satisfy (2.15). The PV field has negative norm so its propagator will cancel the regulated field wherever the PV mass has a negligible effect on the propagator. In particular, there will be a thorough cancellation for all values of x which satisfy $g^{tt}(x)E^2 \gg M_{PV}^2$. This is equivalent to equation (2.7) for $\Lambda = M_{PV}$.

As an example of the equivalence between our regularization procedure and Pauli-Villars, we look again at the RS model. PV was used in this scenario in [82]. The propagators for massless and massive fields were derived in [87] and involve Bessel functions such as $\mathcal{J}_\nu(pz)$ where $\nu = \sqrt{1 + M^2/k^2}$ and p is the momentum. These functions have the nice property that $\mathcal{J}_\nu(x)$ is independent of ν (up to a phase) for $x \gg \nu$. Thus, the propagator for a field and its PV regulator will cancel if $z > \frac{\nu}{p} \approx \frac{M}{kp}$. This is exactly the condition we use when applying the position-dependent cutoff.

Once we have $g(E)$, we can generate all the important thermodynamic quantities. The free energy is:

$$F = -\frac{1}{\pi} \int \frac{g(E)}{e^{\beta E} - 1} dE \qquad (2.16)$$

Which gives $U = \frac{\partial}{\partial \beta}(\beta F)$ and $S = \beta(U - F)$. For our examples, the scaling of $g(E)$ with E will be the same as the scaling of $S(T)$ with T.

2.4 Global AdS

In [115, 60], the thermodynamics of global AdS was considered. Witten found that at a temperature of order k (the AdS curvature scale), there is a transition from pure AdS to AdS-Schwarzschild, above which both the bulk and boundary theory reflect the number of degrees of freedom of a four-dimensional theory in that the entropy scales as T^4. We now show that our counting of states agrees with the above transition between that of a low-energy theory and one for which it reflects the full dimensionality (our estimate of states will not reflect the gap). More precise agreement would require choosing Λ and the number of fields in accordance with the holographic correspondence (see below).

The AdS potential is $V(r) = 1 + k^2 r^2$. To apply our procedure, we introduce a boundary regulator brane at a position $r = R$. Also, for simplicity, we consider AdS_5 rather than $AdS_5 \times S^5$. As explained in the previous section, we expect the number of states is approximately (note that now we are in 5D):

$$g(E) \sim E^4 \int_{r_E}^{R} \frac{r^3 dr}{V(r)^{5/2}} \sim E^4 \int_{r_E}^{R} \frac{r^3 dr}{k^5 r^5} \qquad (2.17)$$

where the last approximation is valid for $r \gg 1/k$. In this equation, r_E is chosen in accordance with a spatially varying cutoff as described in Section 2.3, $\Lambda(r) = \sqrt{1 + k^2 r^2}\Lambda$. Setting $\Lambda(r_E) = E$ leads to

$$r_E = \frac{1}{k}\sqrt{\frac{E^2}{\Lambda^2} - 1} \sim \frac{E}{k\Lambda} \qquad (2.18)$$

The number of states is now readily evaluated to find

$$g(E) \sim \left(\frac{E}{k}\right)^4 \left(\frac{\Lambda}{E}\right) \sim \left(\frac{E}{k}\right)^3 \frac{\Lambda}{k} \qquad (2.19)$$

This shows the dependence on energy of a four-dimensional theory when E is sufficiently large, as anticipated. Notice the confinement of the high energy modes to the region near the boundary yields the factor of $1/E$ that converts the behavior from five to four-dimensional.

For low energies, $E < \Lambda$, the answer will not be that of a lower-dimensional theory. That is to be expected, since for these energies, one is probing distance scales less than $1/k$, for which the theory should resemble flat space. This follows when applying our regulator, since for energies less than Λ, the entire space, down to arbitrarily small r, can support the state. Actually, a more careful analysis with the precise modes would reflect the band gap that would limit this 5D scaling behavior to energies between k and Λ. In summary, the theory appears to have 5-dimensional counting of states for $k < E < \Lambda$ and appears 4-dimensional for $E > \Lambda$. Notice that this behavior closely resembles what we found for the Poincare patch calculation, which had the fewest modes at low energies, then had a five-dimensional regime, then evolved to the holographic four-dimensional regime.

We can ask what we learn from this method that we did not already know by using the precise holographic dual [101]. The answer is that we have a guide to the precise nature of the correspondence, since we can see what states exist in the bulk dual theory. For example, in the state counting done by Susskind and Witten, the Bekenstein bound was assumed for the bulk. But if we use the additional information of precisely what the cutoff should be as determined by the duality, we can see this counting of bulk states explicitly. To see this, we use the coordinate parameterization of AdS space assumed in that paper, where $V(r) = 1/(1 - (kr)^2)^2$ and k sets the AdS curvature scale. Then with a regulator brane a distance δ from the boundary, the 4D theory would have a cutoff on distance δ so that the cutoff on proper distance is $1/k$. Then the volume of the AdS$_5$ is $(1/k) \times \mathcal{A} \times k^4$, where $\mathcal{A}$ is the 3D area and we have used the cutoff k provided by the dual holographic theory, which we know for this example. From this, we almost have a result well within the Bekenstein bound. But this was for a single field in the bulk. We expect that the description of the holographic dual would require N^2 such fields so that the total maximum entropy would be

$N^2 A k^3$, where N is a parameter from the dual theory. [1]

We can also obtain the answer for the density of states through studying directly the eigenvalue spectrum. For simplicity, we show a 4D AdS example, rather than the 5D example we just looked at. When $E > \Lambda$ we restrict r by $r_E < r < R$. Let us take the limit $Rk \gg 1$ and $E \gg \Lambda$. If we change variables to $\rho = \frac{r_E}{r}$ and define $\varphi = \rho \chi(\rho)$ then in this limit, (2.10) simplifies to an analog quantum mechanics problem:

$$-\partial_\rho^2 \chi + \frac{2}{\rho^2}\chi - \frac{\Lambda^2}{k^2}(1 - \frac{l(l+1)k^2}{E^2})\chi = 0 \tag{2.20}$$

Note that we have absorbed the spatially-varying cutoff into the normalization of ρ so this equation has boundary conditions at fixed ρ. The solutions are Bessel functions (in fact, they are the same Bessel functions as in the 4D Poincare patch case). The eigenvalues for each l are integers j_l related to the energy and other parameters as:

$$j_l^2 = \frac{\Lambda^2}{k^2}(1 - \frac{l(l+1)k^2}{E^2}) \tag{2.21}$$

In other words:

$$E^2 = \frac{l(l+1)k^2}{1 - (\frac{nk}{\Lambda})^2} \tag{2.22}$$

Therefore the number of modes at each l is bounded by $n < \frac{\Lambda}{k}$. The total number of modes less than E is $g(E) \sim \frac{\Lambda}{k}(\frac{E}{k})^2$. This density of states is 3D which is holographic to the 4D background.

Notice that the energy scale here is different from the Poincare Patch example. With the regulator brane there is the potential for a phenomenologically viable 4D theory of gravity if we choose the energy cutoff to be of order M_{Pl}, which is related to the 5D Planck scale M by $M_{Pl} = (\frac{M}{k})^{3/2} k^2 R$. For large R this is a much higher cutoff than the natural expectation $\Lambda \approx M$ which we have used.

2.5 Black Hole Thermodynamics

In this section, we consider the contribution to the entropy from a scalar field in the exterior of a black hole [102, 39]. Although this is not the fundamental contribution to a black

---

[1] We thank Massimo Porrati for discussions of this calculation.

hole's entropy, it can give us information about the scaling with cutoff and dimension of this fundamental contribution which should have the same dimension-dependence.

Black holes and de Sitter space (see next section) differ from the AdS examples we just studied in that states are concentrated not on a boundary where g_{tt} is largest but on a horizon, where it is smallest. It might seem surprising that our method of counting states would yield a concentration of states on the horizon, since, as we have emphasized, the number of states is reduced by restricting states to a region where they are consistent with the local cutoff. By this reasoning, we would expect high energy states to be concentrated away from the horizon so that the entropy would also be concentrated there.

However, state counting proceeds differently in the black hole and de Sitter space examples, because we assume that there is a fixed temperature, T, as measured at infinity (or the origin for de Sitter space). Therefore, in general, the energy stays well below the local cutoff. In fact, energy at a given position r will be chiefly of order $T/\sqrt{V(r)}$, where $V(r)$, given in Section 2.3, goes to zero near the horizon. This means that for a fixed T, the highest energy region is where $V(r)$ is smallest, the opposite of what happens if we allow all energies up to the cutoff at all r. In fact, energies only achieve the local cutoff near the horizon. With no cutoff on energy, they would diverge as one approaches the horizon. We assume a cutoff on energy Λ which we expect to be of order M_{Pl}. This in turn implies a cutoff on position: states cannot get too close to the horizon unless they have arbitrarily low energy (as measured at infinity).

For a black hole, $V(r) = 1 - \frac{r_S}{r}$, where $r_S = \frac{2m}{M_{Pl}^2}$ is the location of the horizon and m is the mass of the hole. Generalizations to rotating or charged black holes are straightforward. The unregulated density of states diverges at $r = r_S$ because $V(r_S) = 0$. Following the prescription of Section 2.3, we define the cutoff to be Λ at $r = \infty$. Then the closest a state of energy E can get to the horizon is determined by $\sqrt{V(r_E)}\Lambda = E$. So,

$$r_E = r_S \frac{\Lambda^2}{\Lambda^2 - E^2} \tag{2.23}$$

The solution to the analog quantum mechanics problem has been extensively studied without energy-dependent boundary conditions (see, for example, [23]). For our purposes, it

will be sufficient to estimate the density of states using Section 2.3.

In order to regulate the IR divergence from the asymptotic Minkowski space, we restrict space to a box of size L. Then the number of states calculated with (2.23) and (2.13) is

$$g(E) \approx \frac{2}{3}r_S^3\Lambda^2 E + \frac{2}{9}L^3 E^3 + \mathcal{O}(\frac{r_S}{L}) + \cdots \qquad (2.24)$$

The $L^3 E^3$ term is just what we expect from flat space. The other term is the leading contribution associated with the black hole. The cutoff dependence comes from a factor of $r_E - r_S$, which scales as $(E/\Lambda)^2$. The entropy which follows from the L-independent part of this $g(E)$ scales with temperature as:

$$S(T) \propto r_S^3 T \Lambda^2 + \cdots \qquad (2.25)$$

Substituting in the Hawking temperature, we see the black hole contribution scales with the horizon area.

If we used the Reissner Nordstrom black hole potential, $V(r) = (1 - \frac{r_+}{r})(1 - \frac{r_-}{r})$ (2.25) would have been replaced by

$$S(T) \propto \frac{r_+^4}{r_+ - r_-}T\Lambda^2 + \cdots \qquad (2.26)$$

Note that the RN temperature is $T_{RN} = \frac{r_+ - r_-}{4\pi r_+^2}$ which makes $S(T_{RN}) \propto \mathcal{A}\Lambda^2$ where $\mathcal{A} = 4\pi r_+^2$ is the area of the horizon.

This computation follows closely that performed by 't Hooft in [102], in which he assumed a sharp cutoff at a coordinate h corresponding to a proper distance of order $1/M_{Pl}$. In our approach, states are kept away from the horizon due to a cutoff on energy, so that a state with finite energy cannot reach the horizon. So the minimum distance of a state from the horizon depends on energy. Because the energy at a position r is of order $T/\sqrt{V(r)}$, the contribution is heavily concentrated in the high proper energy states that are closest to the horizon. In fact, it is easy to see that the proper distance of these states from the horizon is of order $1/M_{Pl}$, as in 't Hooft's calculation.

Our calculation is even closer in spirit to that of Demers, Lafrance, and Myers [39]. They also calculated the entropy contributed from a scalar field external to a black hole.

Their calculation employed a Pauli-Villars regulator, which we have already shown closely matches our regulator. They furthermore verified the suggestion of Susskind and Uglum [100] by demonstrating the renormalization of the entropy was consistent. In flat space. Newton's constant gets renormalized as:

$$\frac{1}{G_N} \to \frac{1}{G_N} + \frac{B}{12\pi} \qquad (2.27)$$

where B is some quadratically divergent function of the five PV masses. Then they use the same PV fields to regulate the divergence of the entropy outside a black hole. The entropy they get, using the same WKB technique, is $S(T) \propto \frac{r_+^4}{r_+ - r_-} TB$. Note the similarity to (2.26). They then interpret this as a renormalization of the black hole entropy:

$$S_{\mathrm{BH}} = \frac{\mathcal{A}}{4G_N} \to \frac{\mathcal{A}}{4}(\frac{1}{G_N} + \frac{B}{12\pi}) \qquad (2.28)$$

The important point is that this is the same quadratically divergent function B as in (2.27).

It is also straightforward to work out the logarithmic corrections. We obtain a term in (2.26) proportional to $\frac{r_+^4}{r_+ - r_-} T^3 \log(\Lambda)$. This piece corresponds to the divergent contribution $T^3 \log(M)$ in [39]. The scale for the logarithm is naturally set by the black hole temperature T_{RN}. We can interpret this logarithm as a constant contribution related to higher curvature terms on the gravitational action [39, 63], being absorbed in the renormalization of higher dimension operators in the gravity action.

This result would be especially appealing in a theory of induced gravity. There the fact that both G_N and the entropy receive the same contribution which grows with the number of degrees of freedom would guarantee there is no problem with a large number of species. We also observe that in this case, one might be able to understand the small value of G_N as a result of a large number of species.

Clearly, there is a contribution to black hole entropy not associated with an external field. Very likely, this contribution is associated with whatever regulates the quantum field theory at energies of order the cutoff Λ. From this perspective, it is not surprising that there exists a string theory example where the full entropy is reproduced [95]. Furthermore, this shows on quite general grounds that whatever the cutoff physics contribution is, it

should scale quadratically with associated mass scale and be proportional to area, in order to absorb the cutoff dependence of our field theory result, rendering it scheme-independent.

Finally, it is of interest to reflect on the form of the result. We found that for a black hole at the Hawking temperature, the entropy had a term that scaled with the volume of the space and an additional term that is attributed to the black hole that scales as $S \times \Lambda^2 R^2$, where Λ is the cutoff and R is the Schwarzschild radius. In fact, the answer really scales as $E^3 R^3 (\Lambda/E)^2$, and when averaging over the thermal ensemble, E gets replaced by T, which is taken to be the Hawking temperature. Clearly, the first factor is what one would expect for the volume inside the black hole horizon. From this perspective, we see that the number of degrees of freedom is in fact far *greater* than the expected number for a system at such low temperature. How are we to understand this result?

Let us consider the value of $r_H - r_S$, where r_H is the minimum r permitted by our regularization procedure for a mode of energy $T_H = 1/R$. One finds $(r_H/r_S - 1) = (T_H/\Lambda)^2$. This corresponds to the *proper* distance from the horizon equal to $1/\Lambda$, what one would expect to be the minimum distance permitted in a system with cutoff Λ. Notice that in this sense, our regularization is naturally leading to the notion of a stretched horizon [98]. The degrees of freedom are concentrated at and around the coordinate r_H. Of course, we are not making any further claims; we do not consider a membrane as a physical object. We are simply making the observation that with our cutoff procedure, the degrees of freedom naturally tend to congregate at this position.

Furthermore, we understand the enhancement of the number of degrees of freedom over that which is expected. We see that the proper energy at r_H is actually Λ, not the much smaller T_H. This is not surprising as this is how r_H was chosen. But we see that the large number of degrees of freedom is due to the much higher proper energy of the modes at r_H. In fact, we can eyeball the result by taking the quantity $E_p^3 \delta r_p^3$, where $E_p = E/(\sqrt{r_h/r_S - 1})$ is the proper energy and $\delta r_p = (r_h - r_S)/(\sqrt{r_h/r_S - 1})$. Of course, this is restating the interpretation given in general in Section 2.3.

2.6 Static Patch of de Sitter Space

The final example we consider is de Sitter space. for which the entropy is bounded and is expected to take the form $S_{BH} \propto M_{Pl}^2 R^2$. We can compute the quantum corrections to this formula, as above for the exterior of the black hole.

The de Sitter function is $V(r) = 1 - \frac{r^2}{R^2}$ with $0 < r < R$. Since $V(r)$ decreases as we approach the boundary, the cutoff $\Lambda(r) = \sqrt{V(r)}\Lambda$ has its maximum at $r = 0$. Setting $\Lambda(r_E) = E$ gives $r_E = R\sqrt{1 - \frac{E^2}{\Lambda^2}}$. The state counting of Section 2.3 gives:

$$g(E) = \frac{R^3}{3}\Lambda^2 E \left(\sqrt{1 - \frac{E^2}{\Lambda^2}} - \frac{E^2}{\Lambda^2}\text{Tanh}^{-1} (\sqrt{1 - \frac{E^2}{\Lambda^2}}) \right) \tag{2.29}$$

This function looks like a hump peaked around $E \approx 0.46\Lambda$.

However, we are only interested in counting the states at low energy. since we assume the space is at the de Sitter temperature. Then.

$$g(E) = \frac{R^3}{3}\Lambda^2 E + \cdots \tag{2.30}$$

The linear energy-dependence arose because the distance to the horizon of a state of energy E scales as $(E/\Lambda)^2$. as in the black hole example. which has the same near horizon structure.

The entropy is

$$S \propto R^3\Lambda^2 T \tag{2.31}$$

Evaluated at the de Sitter temperature $T = T_{dS} \approx \frac{1}{R}$ this says that $S_{dS} = R^2\Lambda^2$. We expect that as with the black hole entropy, there is an associated renormalization of Newton's constant. There are also logarithmically divergent pieces which correspond to the renormalization of higher dimension operators, as with black holes.

Notice that as with the black hole case, this is a much *larger* number of states than would be expected in a box of radius R in flat space at temperature T. which scales like $R^3 T^3$. This enhancement arises in a similar manner to the enhancement of states for a black hole at low temperature that we already discussed. Notice also that the entropy we calculated is not obviously a restriction on the number of states, but just corresponds to the

number of states at the de Sitter temperature. The fundamental description of the system in principle involves many more states. most of which do not participate.

2.7 Coordinate Dependence

In the previous sections. we illustrated the utility of the spatially varying cutoff in reproducing qualitatively the known results for the entropy in several different geometries. The AdS case is well understood. although there are still puzzling features such as the explicit implementation of the UV/IR correspondence [80]. The black hole case is less well understood and exhibits some interesting features that we discuss in more detail in this section.

Although we understood the form of our answer in terms of local energy and proper volume, there is an obvious issue that we have so far skirted which is the coordinate dependence of our result. Some of the coordinate dependence of course comes from the non-covariant nature of the way we implemented our cutoff. A more careful analysis would use Pauli-Villars, which we have shown our method closely approximates. We use our procedure because the counting is very simple.

However, more worrisome for the black hole example is that one can choose coordinates that are completely smooth at the horizon. Where would the cutoff dependence of black-hole entropy then arise? We will necessarily be speculative in our considerations in this section. However. we find it suggestive that our results overlap with proposals that have already been made.

The key criterion that distinguishes the metrics in which we do our calculations from those we have avoided is that the ones we use are *time-independent*. For metrics that are strongly time-dependent, our cutoff procedure simply does not apply and we do not know how to do a calculation of the sort we have outlined since we cannot count energy eigenstates as we have been doing. We will next explain what goes wrong in several alternative parameterizations of the black hole metric. We will also argue by following the details of the coordinate transformation that there might exist long wavelength modes or nonlocal correlations that are not accounted for in low-energy field theory and which are sensitive

to the cutoff physics. This would be a correction to low-energy field theory when counting degrees of freedom. However, these modes should be very low energy and long wavelength, and therefore irrelevant at a practical level, in which field theory should still apply.

2.7.1 Kruskal Coordinates

To get coordinates which are smooth at the horizon, we first transform from Schwarzschild (r,t) to advanced and retarded Eddington-Finkelstein coordinates:

$$U = t + r - r_s \log|r/r_s - 1| \tag{2.32}$$

$$V = t - r - r_s \log|r/r_s - 1| \tag{2.33}$$

The metric in these coordinates is still not well-behaved at the horizon, but it remains Minkowski for large r. In Kruskal coordinates: $x = e^{U}$, $y = e^{V}$, the metric is:

$$ds^2 = \frac{32M^3}{r} e^{-r/r_S} dx\, dy + r^2 d\Omega^2 \tag{2.34}$$

There is no longer a singularity at the horizon. However for large r, the metric looks like $ds^2 = \frac{dx\, dy}{xy} + r^2 d\Omega^2$. This is not Minkowski and one has to worry about physics far away from the horizon.

As an example of how we must be careful using our physical intuition in these coordinates, consider a mode near the horizon. It has very small energy measured at infinity, and we define $\epsilon = \frac{E}{\Lambda}$. The distance to the horizon this mode can probe is $r_{min} = r_S + \epsilon^2 r_S$. To measure this mode, one would need an amount of Schwarzschild time approximately equal to $\Delta t = \frac{1}{E} = \frac{1}{\epsilon \Lambda}$. In Eddington-Finkelstein coordinates this means that $\Delta U \approx \Delta V \approx \frac{1}{\epsilon}$. While this is large, the space-like combination $U - V$ is still small. However, in Kruskal coordinates, $\Delta x \approx \Delta y \approx \epsilon e^{\frac{1}{\epsilon \Lambda}}$. Now the distance measured with respect to the space-like combination $x - y$ is exponentially large. Since the space is flat near the horizon, this corresponds to an exponentially large proper distance as well.

Tidal effects make any given region grow with t in terms of x and y coordinates. We see that taking the minimum time that would be necessary to probe low energy states near the horizon, the near-horizon region is transformed into a huge region. We would expect a

large number of low-energy states associated with this patch. These are not anticipated in the field theory associated with the patch that corresponded to the near-horizon region.

It is of interest to consider the possibility of such low-energy states. Because they are very large and low-energy, they would not affect any local physics so they are not precluded by the success of low-energy field theory. In fact, they are not necessarily states but could be correlations in existing states that store information. In this sense, they would be similar to the "precursors" suggested in [80, 99]. Also, because these states are sensitive to the cutoff of the theory, as has already been emphasized, we see this would implement a UV/IR connection [101, 77]. Finally, it is perhaps not unexpected that such large nonlocal states should exist in this new coordinate system, since the existence of an event horizon summarizes physics that is extended in time. In the new coordinates, where time and space are mixed, there might be new extended states.

Although we did the exercise for Kruskal coordinates, we expect similar results for any time-dependent coordinate system. One might expect things to simplify in coordinates which are well-behaved at the horizon but smoothly approach Minkowski space far away. We briefly consider one family of such coordinates, nice slice, in the next section and demonstrate that the situation is just as confusing.

2.7.2 Nice Slice

By Birkhoff's theorem, we know that the only set of coordinates in which the black hole metric is time-independent and asymptotically Minkowski is Schwarzschild coordinates. Nevertheless, we can go to a set of coordinates which are nonsingular near the horizon and have a sort of minimal time dependence [79]. We define a one parameter (R) family of slicings implicitly by:

$$-e^{T/2r_S}y + e^{T/2r_S}x = 2R \qquad (2.35)$$

x and y are Kruskal coordinates defined above. R can take values from 0 to $2r_S$. For $R = 0$, T is just Schwarzschild time t. So, by varying R we can introduce time-dependence in a controlled way. To get the metric, we can solve (2.35) for T and then solve for a coordinate

Z orthogonal with respect to the black hole metric. The solution is:

$$T = -2r_S \log\left(\frac{R + \sqrt{R^2 + 4r_S^2 xy}}{2r_S x}\right) \tag{2.36}$$

$$Z = -2\sqrt{R^2 + 4r_S^2 xy} + 2R\log\left(\frac{R + \sqrt{R^2 + 4r_S^2 xy}}{2r_S x}\right) \tag{2.37}$$

Note that $\sqrt{R^2 + 4r_S^2 xy} = \frac{1}{2r_S}(RT + r_S Z)$ which lets us express xy in terms of $RT + r_S Z$. both of which are independent of Schwarzschild time t. The metric in nice-slice coordinates is:

$$ds^2 = \frac{r_S}{4r}e^{-\frac{r}{r_S}}\left(-\frac{(RT + r_S Z)^2}{4r_S^4}dT^2 + dZ^2\right) \tag{2.38}$$

Although ∂_t is still a killing vector. this metric does depend on the new time parameter T. For any R. $T \to t$ as $r \to \infty$. For $R = 0$. T becomes t everywhere. but the metric becomes singular at the event horizon $r = r_S$ ($RT + r_S Z = 2r_S R$). For nonzero R space is flat (by construction) at the event horizon. but the coordinate singularity has moved to $RT + r_S Z = 0$. This is a surface of constant r solving $\frac{R^2}{4r_S^2} = e^{r/r_S}(1 - \frac{r}{r_S})$. For $R = 2r_S$. the horizon moves to the essential singularity at $r = 0$. In some sense, all we are doing is moving the coordinate singularity between the event horizon and the physical singularity by varying R.

Now suppose we tried to do quantum field theory in such a background. Again. we would have to worry about states within ϵ of the horizon. In nice slice coordinates. the position of the horizon depends on time as depends on time as $Z \approx \frac{2Rt}{r_S}$. Since these modes have time uncertainty $\Delta t = \frac{1}{\epsilon\Lambda}$ we get that $\Delta Z \approx \frac{R}{\epsilon\Lambda}$. So while in Schwarzschild coordinates. the mode is localized within ϵ of the horizon. in nice slice the mode is not localized at all. That is. while the horizon is just a line in nice slice. the near horizon region is very very large. If we tried to regulate the theory, we would have to include finite time in our regulator to somehow take this into account. Clearly this is a peculiar situation and not something we really understand. It might also be connected with the existence of new. low-energy states or nonlocal correlations.

2.8 Conclusions

Let us summarize our method and its justification. We define a space-dependent condition on energy, or equivalently a cutoff on local energy. By interpreting this as a cutoff on position, we find an energy-dependent boundary for our space which is used to quantize the system and evaluate the density of states. One can also use the space-dependent cutoff to do field theory, as I will demonstrate in the next chapter. Another application is to testing qualitative features of gravitational theories. For example, it is readily seen by a field theory calculation with our regulator that the proposal of [105] to address the cosmological constant solely through the holographic nature of the theory does not work. For example, in the AdS case, all the degrees of freedom are concentrated where the energy is largest, with no additional red-shift factor.

Our counting relies on the fact that any state above the local cutoff cannot be treated simply in field theory. Above the cutoff, one expects to find either black holes or states intrinsic to the fundamental theory providing the cutoff, e.g. string theory. At short distance, these are not weakly interacting field theory states. For this reason, the standard argument, relying on a UV fixed point, that a boundary theory must reflect an asymptotically AdS space does not apply.

In addition, the bulk theory description that we have given suggests additional "stored" degrees of freedom that have become strong bound states at the cutoff. For example, a rolling scalar field in AdS space might change the vacuum energy so that many more or fewer states appear in the region that appears non-holographic from a field theory vantage point. These states must be present once the curvature changes; do they emerge out of nowhere or have they dropped from above to below the cutoff? So we interpret the entropy bounds as bounds on the degrees of freedom that can be simultaneously excited in practice.

We do not assume the existence of a boundary theory. However, in all cases we have studied with a monotonic g_{tt} that varies sufficiently strongly, we find the degrees of freedom concentrated on the boundary of the space. It is not clear that there exists a more useful boundary description in general.

We also have seen that this approach is coordinate dependent. However, for most of our examples, there is a unique choice of metric for which there is not time-dependence. If there is a gauge-invariant formulation, our approach should be the result of a particular gauge choice. It seems clear that time-dependent metrics are more subtle to understand. Despite the existence of a covariant formulation of the holographic principle [25], it is not clear how to apply any of the defining quantities we have used, namely energy eigenstates, temperature, and entropy in a strongly time-dependent vacuum.

That low-energy states or non-local correlations can exist and be consistent with all known field theory successes is important. Local measurements would not know about big low-energy states since they carry low energy and their effects would be suppressed by small wave function overlap, or equivalently, a large normalization factor. However, if they are present, they can provide correlations that can in principle be measured over large distances. Such non-local effects should be relevant to the information problem of black holes [76].

It is interesting to contrast our approach with the "gauge theory" approach that has been suggested by which one would have some principle through which one could eliminate redundant degrees of freedom. As we have said, our procedure depends on a time-independent coordinate choice. It does not however preclude existence of a "gauge" theory where this is more readily embodied in a covariant way; in such a formulation our answer should correspond to a choice of gauge.

Of course, there is much more to fully understanding the Bekenstein bound and holography than the simple procedure we have outlined. However, it does well approximate the counting of states in spaces that are highly curved even without exciting additional fields. Also it might serve as a limit or special case of a more general more covariant formulation.

The reason our procedure does not supply the final answer is that we have not incorporated any back-reaction and we have not supplied the physics of the cutoff in our counting. However, it is a remarkably simple procedure to use in order to deduce qualitative features of a gravitational system. If we could also exclude extended (large) black holes, we might always find area-law behavior. The problem is that this involves a constraint on both en-

ergy and size: there is not way to do this in a conventional field theory. However. better understanding this constraint might yield some insight into UV/IR.

2.9 Acknowledgements

Chapter 3

Unification in RS

3.1 Introduction: The Randall-Sundrum Model

The ideas of the previous chapter find precise and practical application in the Randall-Sundrum (RS) model [89]. RS proposes that the universe is really five dimensional anti-de Sitter space, with a huge cosmological constant Λ. This is not in obvious contradiction with observation if one of two criteria is satisfied: 1) the extra dimension is compact, so that at low energy the theory looks four-dimensional or 2) matter is confined to a four-dimensional subspace, or 3-brane, within the five-dimensional bulk. Both of these possibilities are quite natural in string theory, and historically were partially inspired by string theory results. The model which seems to be the most viable phenomenologically is commonly called RS1, and satisfies both of these criteria, *i.e.* it has matter on a 3-brane in a compact extra dimension. RS1 has the added bonus that it naturally solves the hierarchy problem, which I will now explain.

The Randall-Sundrum model takes place in a part of five dimensional anti de-Sitter space (AdS$_5$) known as the Poincare patch. This part is covered by the 4D coordinates x^μ and the extra dimensional coordinate r. The background metric is:

$$ds^2 = e^{-2kr}\eta_{\mu\nu}dx^\mu dx^\nu + dr^2 \tag{3.1}$$

The 5D cosmological constant is expressed in terms of k and the 5D Planck scale M_{5D} as

32

$\Lambda_{5D} = -k^2 M_{5D}^3$. These coordinates are chosen so that the induced metric at fixed r is flat. In addition, we assume the space is compact, with 3-branes at $r = 0$ and $r = r_c$. In order for (3.1) to satisfy Einstein's equations, we must impose what are known as orbifold boundary conditions on the branes. We can ignore the orbifold structure for now, as it is not essential for the introductory discussion.

The important feature of (3.1) is the presence of the "warp factor" e^{-2kr}. This factor implies that the natural scale on a slice at fixed r is exponentially related to the natural scale on other slices. For example, the induced Planck scale is $M_{Pl}(r) \sim e^{-2kr} M_{Pl}$ where M_{Pl} is the Planck scale on the "Planck brane" at $r = 0$. The warp factor lets us configure the natural scale on the second brane to be $\sim$ TeV $\sim 10^{-16} M_{Pl}$ which corresponds approximately to the strength of forces in the Standard Model, with an r of around 30. If matter is confined to the "TeV brane" but gravity is free to propagate throughout the 5D bulk, then the strength of the gravitational and electromagnetic forces will be exponentially related in the low energy effective theory. Thus the hierarchy problem, which classically is the fact that the gravitational force is 10^{40} times weaker than electric force, can be naturally solved by the RS setup. Moreover, the problem is solved at the quantum level as well, at least within the validity of effective field theory, because quantum corrections to the weak scale are controlled by the warp factor.

The reason RS1 solves the hierarchy problem is because the scale on the TeV brane where gravity becomes strong is $\sim$ TeV but the low energy Planck scale is still $M_{Pl} \sim 10^{16}$TeV. This would be perfect if we had no experimental clues about physics above a few TeV. However, one of most inspirational features of the Standard Model is that the strengths of the strong, weak, and electromagnetic forces vary with energy, and seem to converge at around $E \sim 10^{16}$GeV [49]. This result is known from perturbative calculations, and so if non-perturbative physics comes in at low energy, such as at $\sim$ TeV in RS1, the calculations cannot be trusted. This is the reason that many models with extra dimensions are considered incompatible with grand unification. What I will show in this chapter is that because of the holographic properties of the AdS background of RS, perturbative unification

is still viable.

We saw in Chapter 2 that in strongly curved backgrounds, in particular in AdS_d, the effective number of field degrees of freedom is the same as in a flat $d-1$ dimensional space. In the same way, the coupling for a d dimensional gauge field on an AdS background will run as if it were $d-1$ dimensional. In particular, in AdS_5, a gauge coupling will run logarithmically, like it does in 4D, instead of as a power law, as it does in flat 5D space [41, 11]. Although it is difficult to do this calculation in an effective low energy four-dimensional field theory, one can do it in the full five-dimensional context. We will see that not only do the couplings run logarithmically, but generically, as in the Standard Model, the couplings almost unify. For specific choices of cutoff and number of scalar multiplets, there is good agreement with the measured couplings and the assumption of high scale unification.

The calculation we do is similar in some respects to that in [82], in which Pomarol considered bulk gauge boson running in RS but with the assumption that the quarks and leptons live on the Planck brane. He made the nice observation that unification can occur in this scenario. The problem is that if matter is on the Planck brane, the model no longer naturally solves the hierarchy problem. Here, we will allow for matter to be on the TeV brane by considering a general scenario with gauge bosons in the bulk. Another essential difference between our calculation and Pomarol's is in the regularization scheme. Pomarol uses Pauli-Villars, which provides light ghosts to cancel the contribution to power-law running from the light KK modes. This is problematic from the effective field theory point of view, and will be discussed further below. Our position dependent regulator is more natural from the holographic point of view (Chapter 2) and leads to a consistent effective field theory.

The organization of this chapter is as follows: first, we review some aspects of RS. In particular, we explore some difficulties with the KK picture for the high scale calculation required for running the couplings. In section 3.4, we derive the position/momentum space propagators in the R_ξ gauges. We then present some of the Feynman rules and in section 3.6 we explore the position/momentum space Green's functions in various limits. We introduce

and motivate our regularization scheme in section 3.7. Through some toy calculations we show that it is necessary to modify the boundary conditions and renormalize the Green's functions for energies greater than T in addition to introducing a position dependent cutoff. We then set up an illustrative calculation, the contribution of a scalar field to the vacuum energy. In section 3.9, we calculate the 1-loop β-functions. We use the background field method, where the Ward identities are manifest. Finally, section 3.10 explores unification in two specific examples. We conclude that coupling constant unification is entirely feasible in a warped 2-brane model. The remainder of this chapter follows very closely to [87].

3.2 Setup

As in [89], we postulate the presence of a fifth dimension, and an anti-deSitter space metric:

$$ds_5^2 = \frac{1}{k^2 z^2}(dt^2 - dx^2 - dz^2).$$
(3.2)

We will generally keep the z-dependence explicit and contract 4D fields with $\eta^{\mu\nu}$. We include two branes: the Planck brane at $z = 1/k$ and the TeV brane at $z = 1/T$. T is related to the size of the extra dimension R by $T = k\exp(-kR)$, and defines the energy scale on the TeV brane. If we take $T \approx$ TeV, we can naturally explain the weak scale in the Standard Model. The fifth dimension can be integrated out to get an effective four-dimensional theory, valid at energies below T. The effective 4D Planck scale M_{Pl} is given by

$$M_{\mathrm{Pl}}^2 = \frac{M_{\mathrm{5D}}^3}{k}\left(1 - \frac{T^2}{k^2}\right).$$
(3.3)

It is generally assumed that $T \ll k$ and $k \approx M_{\mathrm{5D}} \approx M_{\mathrm{Pl}}$.

Now we put gauge bosons in the bulk. The first question is what is the quantity we wish to calculate. We know we cannot run the coupling on the TeV brane above the strong gravity scale, which is roughly TeV. Furthermore, we are ultimately looking at the zero mode, the only light mode in the theory at energies of order TeV or below. This is important; it means the logarithmic running of the coupling considered in [13, 42] does not apply; it is a result of the sum of all the excited gauge modes, all but one of which are heavy from a low-energy

perspective. Therefore. the holographic computation of large threshold corrections does not apply. We are ultimately interested only in the relation of the zero-mode coupling of the four-dimensional theory to the high energy five-dimensional coupling.

3.3 Effective four-dimensional theory

It is of interest to consider the four-dimensional effective theory. As mentioned above. we will abandon this in favor of the full five-dimensional calculation. but here we discuss the theory and why it is problematic.

The action for a 5D gauge boson is:

$$S_{5D} = \int d^4x\, dz \sqrt{G} \left[-\frac{1}{4} F_{MN} F^{MN} + \frac{1}{2} \left(a^2 k \delta \left(z - \frac{1}{T} \right) + \tilde{a}^2 k \delta \left(z - \frac{1}{k} \right) + m^2 k^2 \right) A^M A_M \right]. \quad (3.4)$$

Here a is a mass term on the TeV-brane. $\tilde{a}$ is a mass term on the Planck brane. and m is the bulk mass. The signs are consistent with our metric convention (3.2).

To begin with. we take the bulk boson to be massless, and set $a = m = \tilde{a} = 0$. We expand the 5D bosons in terms of an orthonormal set of KK modes [29]-[37]:

$$A_\mu(x, z) = \sqrt{k} \sum_{j=0} A_\mu^{(j)}(x) \chi_j(z) \quad (3.5)$$

$$A_5(x, z) = \sqrt{k} \sum_{j=1} A_5^{(j)}(x) \frac{1}{m_j} \partial_z \chi_j(z). \quad (3.6)$$

The expansion of A_5 is chosen to diagonalize the couplings between A_5 and A_μ (see below). Keep in mind that the mass dimensions are $[z] = -1$, $[g_{5D}] = -1/2$, $[A_M] = 3/2$, $[A_5^{(n)}] = [A_\mu^{(n)}] = 1$, and $[\chi_n] = 0$. The eigenfunctions satisfy:

$$\partial_z (\frac{1}{z} \partial_z \chi_j) = -\frac{m_j^2}{z} \chi_j, \quad \int \chi_i(z) \chi_j(z) \frac{dz}{z} = \delta_{ij} \quad (3.7)$$

and therefore have the form:

$$\chi_j(z) = z(\mathcal{J}_1(m_j z) + \beta_j \mathcal{Y}_1(m_j z)). \quad (3.8)$$

We assume that the 5D boson, and all the KK modes, have even parity under the Z_2 orbifold transformation. Consequently their derivatives must vanish on both boundaries.

That is, even parity leads to Neumann boundary conditions. Similarly, odd parity leads to Dirichlet boundary conditions. which we will discuss in more detail later on. This leads to the quantization condition:

$$\beta_j = -\frac{\mathcal{J}_0(m_j/k)}{\mathcal{Y}_0(m_j/k)} = -\frac{\mathcal{J}_0(m_j/T)}{\mathcal{Y}_0(m_j/T)}.$$

(3.9)

For $m_j \ll k$, $\mathcal{Y}_0(m_j/k)$ blows up. and so the masses are basically determined by the zeros of $\mathcal{J}_0(m_j/T)$. Therefore, they are spaced in energy by approximately πT. This spacing is quite general: it is independent of bulk or boundary mass terms, and of the spin of the bulk field; because $\mathcal{J}_\nu$ oscillates with the same period for any ν. bulk fields will always have excitations of order T.

We chose the conditions (3.5), (3.6) and (3.7) to normalize the kinetic terms and diagonalize the couplings in the effective 4D action:

$$S_{4D} = \int d^4x \left[-\frac{1}{4}(F_{\mu\nu}^J)^2 + \frac{1}{2}(\partial_\mu A_5^J)^2 - \frac{1}{2}m_j^2(A_\mu^J)^2 + m_J(\partial_\mu A_5^J)A^{\mu J} \right].$$

(3.10)

This action is explicitly gauge invariant. Indeed. the 5D gauge invariance

$$A_M \rightarrow A_M - \frac{1}{g_{5D}}\partial_M \alpha(x, z)$$

(3.11)

has its own KK decomposition (we expand $\alpha = \alpha_i \chi_i$):

$$
\begin{aligned}
A_\mu(x, z) \quad &\rightarrow \quad \sqrt{k}A_\mu^i(x)\chi^i(z) - \frac{1}{g_{5D}}\partial_\mu \alpha_i(x)\chi^i(z) \\
&= \quad \sqrt{k}\left[A_\mu^i(x) - \frac{1}{\sqrt{k}g_{5D}}\partial_\mu \alpha_i(x) \right]\chi^i(z) \\
A_5(x, z) \quad &\rightarrow \quad \sqrt{k}A_5^i(x)\frac{1}{m_i}\partial_z \chi^i(z) - \frac{1}{g_{5D}}\alpha_i(x)\partial_z \chi^i(z) \\
&= \quad \sqrt{k}\left[A_5^i(x) - \frac{m_i}{\sqrt{k}g_{5D}}\alpha_i(x) \right]\frac{1}{m_i}\partial_z \chi^i(z).
\end{aligned}
$$

We can plug this back into the action to see that each mode of the 5D gauge field has an independent gauge freedom.

At this point, it is standard to set $A_5 = 0$. We see immediately that this breaks all but the zero mode of the gauge invariance (there is no $A_5^{(0)}$ since $\partial_z \chi_0(z) = 0$). All modes of A_5 are eaten by the corresponding excited modes of the gauge boson. This is the unitary

gauge. The Goldstone boson (A_5) is eliminated and the massive gauge boson propagators must take the form:

$$\langle A_5^j(p) A_5^j(-p)\rangle = \frac{-i}{p^2 - m_j^2}\left(\eta^{\mu\nu} - \frac{p^\mu p^\nu}{m_j^2}\right). \tag{3.12}$$

Although this makes the 4D action look very simple. it is problematic for evaluating loop diagrams. When we do the full 5D calculation later on. we will use the Feynman-'t Hooft gauge. in which A_5 is included as a physical particle.

Because it will be useful for interpreting our results. we present the mass of the lightest KK modes for states of various spin and mass. For the general Lagrangian (3.4). with arbitrary mass parameters, the spectrum of a bulk gauge boson is determined by:

$$\frac{\left(-\bar{a}^2/2 + 1 - \nu\right)\mathcal{J}_\nu\left(m_j/k\right) + m_j/k\mathcal{J}_{\nu-1}\left(m_j/k\right)}{\left(-\bar{a}^2/2 + 1 - \nu\right)\mathcal{Y}_\nu\left(m_j/k\right) + m_j/k\mathcal{Y}_{\nu-1}\left(m_j/k\right)} =$$
$$= \frac{\left(a^2/2 + 1 - \nu\right)\mathcal{J}_\nu\left(m_j/T\right) + m_j/T\mathcal{J}_{\nu-1}\left(m_j/T\right)}{\left(a^2/2 + 1 - \nu\right)\mathcal{Y}_\nu\left(m_j/T\right) + m_j/T\mathcal{Y}_{\nu-1}\left(m_j/T\right)}. \tag{3.13}$$

where $\nu = \sqrt{1 + m^2}$. If m or $\bar{a}$ is nonzero, the lowest mass is of order T. It cannot be lowered below T unless m or $\bar{a}$ are of the order $T/k \approx 10^{-13}$. For the WZ bosons to pick up a weak scale mass without fine tuning. it must come from the TeV brane. It is not hard to show that for $m = \bar{a} = 0$ but $a \neq 0$, the lowest mass is given by [62]:

$$m_0 \approx \frac{a}{\sqrt{2kR}}T. \tag{3.14}$$

If $T = \text{TeV}$, then we can get the weak scale with $a \approx 0.1$.

To get a feel for how the spectrum depends on various parameters, we present in table 3.1 the exact numerical KK spectrum with $e^{kR} = k/T = 10^{13}$.

Next, we look the couplings between the KK modes in a non-Abelian theory.

$$S \supset \int d^4x \left[-g_{ijk}f^{abc}(\partial_\mu A_\nu^{ia})A^{j\mu b}A^{k\nu c} - \frac{g_{ijkl}}{4}(f^{eab}A_\mu^{ia}A_\nu^{jb})(f^{ecd}A^{k\mu c}A^{l\nu d})\right], \tag{3.15}$$

where the coupling constants are given by overlap integrals.

$$g_{ijk} = g_{5D}\sqrt{k}\int \frac{dz}{z}\chi_i\chi_j\chi_k \quad \text{and} \quad g_{ijkl} = g_{5D}^2 k\int \frac{dz}{z}\chi_i\chi_j\chi_k\chi_l. \tag{3.16}$$

k/T $= 10^{13}$	massless vector	massive vector $m = 1$	massive vector $m = 5$	massive vector $m^2 = -1$	massless Dirichlet vector	massless scalar	massive scalar $m_s = 1$
m_0/T	0	2.869	6.873	1.297	3.832	0	4.088
m_1/T	2.458	6.086	10.762	4.134	7.016	3.832	7.321
m_2/T	5.575	9.249	14.196	7.213	10.174	7.016	10.498

Table 3.1: The exact numerical KK spectrum with $e^{kR} = k/T = 10^{13}$.

In flat space, momentum conservation in the fifth dimension implies that KK number is conserved. For example, g_{234} would have to vanish, but g_{224} would not. In curved space this is not true: $g_{ijk} \neq 0$ in general. We can say something for the zero mode, however. Since we have included no masses, its profile is constant and equal to:

$$\chi_0(z) = \frac{1}{\sqrt{kR}} \, . \tag{3.17}$$

Because the χ's are orthonormal and χ_0 is constant:

$$g_{0ij} = \frac{g_{5D}}{\sqrt{R}} \delta_{ij} \quad \text{and} \quad g_{00ij} = \frac{g_{5D}^2}{R} \delta_{ij} \, . \tag{3.18}$$

If we set $g_{5D}^2 = g_{4D}^2 R$, then the effective theory with just the zero mode looks identical to a 4D system. Moreover, the zero mode couples with equal strength to all the KK modes.

It's easy to get a rough idea of how the coupling would run if we took the effective theory at face value. That is, we assume all the KK modes are separate particles, and we use a 4D regularization scheme. At low energy, below $m_1 \approx T$, only the zero mode can run around the loops. As the energy is increased to q, q/T modes are visible. The result is power law running, similar to what has been observed in [41]. This is not the correct result.

One possible improvement, suggested by Pomarol in [82], is to regulate with a Pauli-Villars field with a five-dimensional mass. This field will have a KK spectrum roughly matching the KK spectrum of the gauge boson, except that it will have a heavy mode near its 5D mass instead of a massless zero mode. Thus all the propagators for the KK modes will roughly cancel and only the zero mode will contribute to running. It will turn out that this is superficially similar to the result we will end up getting from the 5D calculation.

However, Pauli-Villars requires that we take the mass of the regulator to infinity. in order to decouple the negative norm states. But an infinitely massive field no longer has a TeV scale KK masses. so it no longer is effective as a regulator. Moreover. it has no hope of telling us threshold corrections. as unitary is violated in the regime where the regulator works.

The root of the problem is that the effective theory breaks down at about a TeV. and so the KK picture is not trustworthy at the high energy scales necessary to probe unification. One can calculate on the Planck brane, but one still has to deal with bulk gauge bosons. The holographic calculation would be at strong coupling. A rigorous perturbative approach is to explore the 5D theory directly.

3.4   5D Position/momentum space propagators

To study the 5D theory. we will work in position space for the fifth dimension. but momentum space for the other four.

3.4.1 R_ξ Gauges

Before gauge fixing. the quadratic terms in the 5D Lagrangian are:

$$L = \frac{1}{2kz}\left[A_\mu(\partial^2\eta^{\mu\nu} - z\partial_z(\frac{1}{z}\partial_z)\eta^{\mu\nu} - \partial^\mu\partial^\nu)A_\nu + 2A_5\partial_z\partial^\mu A_\mu - A_5\partial^2 A_5\right].\tag{3.19}$$

We would like to set $A_5 = 0$ and then choose the Lorentz gauge $\partial_\mu A^\mu = 0$. But these conditions are incompatible. Instead. we will use the following gauge-fixing functional:

$$\Delta L = -\frac{1}{2\xi kz}\left[\partial_\mu A^\mu - \xi z\partial_z\left(\frac{1}{z}A^5\right)\right]^2.\tag{3.20}$$

This produces:

$$L + \Delta L = \frac{1}{2kz}\left[A_\mu\left(\partial^2\eta^{\mu\nu} - z\partial_z\left(\frac{1}{z}\partial_z\right)\eta^{\mu\nu} - \left(1 - \frac{1}{\xi}\right)\partial^\mu\partial^\nu\right)A_\nu + \right.$$
$$\left. + A_5(-\partial^2)A_5 + \xi A_5\partial_z\left(z\partial_z\left(\frac{1}{z}A_5\right)\right)\right].\tag{3.21}$$

We can then read off the equation that the A_μ propagator must satisfy:

$$\langle A^\mu A^\nu \rangle = -iG_p(z,z') \left(\eta^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right) - iG_{\frac{p}{\sqrt{\xi}}}(z,z') \left(\frac{p^\mu p^\nu}{p^2} \right). \tag{3.22}$$

where $p_\mu = i\partial_\mu$ is the four-momentum, and

$$\left[\partial_z^2 - \frac{1}{z}\partial_z + p^2 \right] G_p(z,z') = zk\delta(z-z'). \tag{3.23}$$

We define A_5's propagator as $\langle A_5 A_5 \rangle \equiv i\frac{1}{\xi} G^{1,1}_{\frac{p}{\sqrt{\xi}}}(z,z')$. where

$$\left[\partial_z^2 - \frac{1}{z}\partial_z + \frac{1}{\xi}p^2 + \frac{1}{z^2} \right] G^{1,1}_{\frac{p}{\sqrt{\xi}}}(z,z') = zk\delta(z-z'). \tag{3.24}$$

We choose this notation for the following reason. If we had included a bulk mass $\frac{1}{2}m^2 k^2 A_M A^M \sqrt{G}$ in L, (3.23) would have been

$$\left[\partial_z^2 - \frac{1}{z}\partial_z + p^2 - \frac{m^2}{z^2} \right] G^{1,m}_p(z,z') = zk\delta(z-z'). \tag{3.25}$$

Then we can interpret (3.24) as saying that A_5 has the Green's function of a vector boson with bulk mass $m^2 = -1$. For simplicity we will continue to write $G_p = G^{1,0}_p$ for the gauge boson.

We can also work out the ghost Lagrangian by varying the gauge fixing functional.

$$L_{\text{ghost}} = \frac{1}{kz}c \left(-\partial_\mu D^\mu - \xi z\partial_z \left(\frac{1}{z}\partial_z \right) \right) c. \tag{3.26}$$

Which makes the ghost propagator:

$$\langle cc \rangle = i\frac{1}{\xi} G_{\frac{p}{\sqrt{\xi}}}(z,z'). \tag{3.27}$$

Note that the ghosts do not couple to A_5 directly, as they would not couple directly to Goldstone bosons in a conventional spontaneously broken gauge theory.

If we take $\xi = \infty$, we get the unitary gauge. The gauge boson propagator looks like:

$$\langle A^\mu A^\nu \rangle = -iG_p(z,z') \left(\eta^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right) - iG_0(z,z') \left(\frac{p^\mu p^\nu}{p^2} \right). \tag{3.28}$$

The first term is the transverse polarization states of all the KK modes. We can think of the second term as subtracting off the longitudinal form of the zero mode. Then the

zero mode's contribution has just the regular $\eta^{\mu\nu}$ tensor structure. In this gauge, the A_5 and ghost propagators are zero. It is easy to imagine how this gauge would make loop calculations very problematic.

$\xi = 0$ is the Lorentz gauge. Here the A_μ propagator is purely transverse. and A_5 and the ghosts have 4-dimensional propagators:

$$\langle A^\mu A^\nu \rangle = -iG_p(z,z') \left(\eta^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right) \qquad \langle A_5 A_5 \rangle = \frac{i}{Rp^2} \qquad \langle cc \rangle = \frac{i}{Rp^2}. \qquad (3.29)$$

Finally, $\xi = 1$ is the Feynman-'t Hooft gauge. The propagators are:

$$\langle A^\mu A^\nu \rangle = -iG_p(z,z')\eta^{\mu\nu} \qquad \langle A_5 A_5 \rangle = iG_p^{1,1}(z,z) \qquad \langle cc \rangle = iG_p(z,z). \qquad (3.30)$$

This is the most intuitive gauge. The A_5 supplies the longitudinal polarizations to the excited modes of A_μ.

3.4.2 Solving the Green's functions

To solve (3.23), we first find the homogeneous solution. Defining $u \equiv \min(z,z')$ and $v \equiv \max(z,z')$, this is:

$$G_p(u,v) = u(A\mathcal{J}_1(pu) + B\mathcal{Y}_1(pu)) = v(C\mathcal{J}_1(pv) + D\mathcal{Y}_1(pv)) \qquad (3.31)$$

$\mathcal{J}$ and $\mathcal{Y}$ are Bessel functions. For positive parity under the orbifold Z_2, we must impose Neumann boundary conditions at both branes:

$$\partial_u G_p \left(\frac{1}{k}, v \right) = \partial_v G_p \left(u, \frac{1}{T} \right) = 0. \qquad (3.32)$$

Finally, matching the two solutions over the delta function leads to the fully normalized Green's function:

$$G_p(u,v) = \frac{\pi}{2} \frac{kuv}{AD - BC}(A\mathcal{J}_1(pu)) + B\mathcal{Y}_1(pu))(C\mathcal{J}_1(pv) + D\mathcal{Y}_1(pv)), \qquad (3.33)$$

where

$$\begin{aligned} A &= -\mathcal{Y}_0(p/k) \quad &\text{and} \quad C &= -\mathcal{Y}_0(p/T) \\ B &= \mathcal{J}_0(p/k) \quad &\text{and} \quad D &= \mathcal{J}_0(p/T). \end{aligned} \qquad (3.34)$$

If the gauge boson has negative parity under the orbifold Z_2, then it must satisfy Dirichlet boundary conditions at both branes (hence we will call it a Dirichlet boson). Its Green's function will have the same form as (3.33) but with:

$$A = -\mathcal{Y}_1(p/k) \quad \text{and} \quad C = -\mathcal{Y}_1(p/T)$$
$$B = \mathcal{J}_1(p/k) \quad \text{and} \quad D = \mathcal{J}_1(p/T). \tag{3.35}$$

Fields of other spin can be derived analogously. For example, the Green's function for a massless scalar is:

$$S_p(u,v) = \frac{\pi}{2} \frac{k^3 u^2 v^2}{A_s D_s - B_s C_s}(A_s \mathcal{J}_2(pu)) + B_s \mathcal{Y}_2(pu))(C_s \mathcal{J}_2(pv) + D_s \mathcal{Y}_2(pv)) \tag{3.36}$$

with

$$A_s = -\mathcal{Y}_1(p/k) \quad \text{and} \quad C_s = -\mathcal{Y}_1(p/T)$$
$$B_s = \mathcal{J}_1(p/k) \quad \text{and} \quad D_s = \mathcal{J}_1(p/T) \tag{3.37}$$

In general, for scalars ($\sigma = 2$), fermions ($\sigma = 1/2$) or vectors ($\sigma = 1$) and with bulk mass m, as in (3.4), the Green's functions are:

$$G_p^{\sigma m}(u,v) = \frac{\pi}{2} \frac{k^{2\sigma-1} u^\sigma v^\sigma}{A_{\sigma m} D_{\sigma m} - B_{\sigma m} C_{\sigma b}} \times$$
$$\times (A_{\sigma m} \mathcal{J}_\nu(pu)) + B_{\sigma m} \mathcal{Y}_\nu(pu))(C_{\sigma m} \mathcal{J}_\nu(pv) + D_{\sigma m} \mathcal{Y}_\nu(pv). \tag{3.38}$$

where $\nu = \sqrt{\sigma^2 + m^2}$ and

$$A_{\sigma m} = -\mathcal{Y}_{\nu-1}(p/k) + (\nu - \sigma)\frac{k}{p}\mathcal{Y}_\nu(p/k)$$

$$B_{\sigma m} = \mathcal{J}_{\nu-1}(p/k) - (\nu - \sigma)\frac{k}{p}\mathcal{J}_\nu(p/k)$$

$$C_{\sigma m} = -\mathcal{Y}_{\nu-1}(p/T) + (\nu - \sigma)\frac{T}{p}\mathcal{Y}_\nu(p/T)$$

$$D_{\sigma m} = \mathcal{J}_{\nu-1}(p/T) - (\nu - \sigma)\frac{T}{p}\mathcal{J}_\nu(p/T). \tag{3.39}$$

Similar results for KK decompositions can be found in [50]. Keep in mind that although A_5 and the ghosts are scalars, their propagators involve the spin-1 Green's function. Intuitively, this is expected because they are necessary for gauge invariance.

We will eventually have to perform a Wick rotation, so that a Euclidean momentum cutoff can be imposed on all the components of p_μ. To this end, we will need the Green's

functions with $p \to iq$. These functions are still real. It is easiest to get them by re-solving equations like (3.23) with $p^2 \to -q^2$. The result is:

$$G_q^{\sigma m}(u, v) = \frac{k^{2\sigma - 1} u^\sigma v^\sigma}{A_{\sigma m} D_{\sigma m} - B_{\sigma m} C_{\sigma m}} \times$$
$$\times (A_{\sigma m} \mathcal{K}_\nu(qu)) + B_{\sigma m} \mathcal{I}_\nu(qu))(C_{\sigma m} \mathcal{K}_\nu(qv) + D_{\sigma m} \mathcal{I}_\nu(qv)). \quad (3.40)$$

where $\nu = \sqrt{\sigma^2 + m^2}$ as before and

$$A_{\sigma m} = \mathcal{I}_{\nu-1}(q/k) - (\nu - \sigma)\frac{k}{q}\mathcal{I}_\nu(q/k)$$

$$B_{\sigma m} = \mathcal{K}_{\nu-1}(q/k) + (\nu - \sigma)\frac{k}{q}\mathcal{K}_\nu(q/k)$$

$$C_{\sigma m} = \mathcal{I}_{\nu-1}(q/T) - (\nu - \sigma)\frac{T}{q}\mathcal{I}_\nu(q/T)$$

$$D_{\sigma m} = \mathcal{K}_{\nu-1}(q/T) + (\nu - \sigma)\frac{T}{q}\mathcal{K}_\nu(q/T). \quad (3.41)$$

3.5 Feynman rules

It is fairly straightforward to derive the Feynman rules in these coordinates. External particles are specified by their 4-momentum and their position in the fifth dimension. The vertices have additional factors of the metric which can be read off the Lagrangian. Both loop 4-momenta and internal positions must be integrated over. The Feynman rules for a non-Abelian gauge theory are (in the Feynman-'t Hooft gauge):

$$z \, \text{~~~~} \, z' \;=\; -iG_p(z, z')\eta^{\mu\nu} \quad (3.42)$$

$$= g_{5d}\frac{1}{kz}f^{abc}[\eta^{\mu\nu}(k - p)^\rho + \eta^{\nu\rho}(p - q)^\mu + \eta^{\rho\mu}(q - k)^\nu] \quad (3.43)$$

$$= -ig_{5d}^2 \frac{1}{kz} N_{abcd}^{\mu\nu,\rho\sigma} \quad (3.44)$$

where $N^{\mu\nu,\rho\sigma}_{abcd}$ is the standard 4-boson vertex tensor and group structure. Then there are the A_5 contributions:

$$z \bullet\!\cdots\!\bullet z' \;=\; iG_p^{1,1} \tag{3.45}$$

$$= \; g_{5d}\frac{1}{kz}f^{abc}(\partial^1_z - \partial^2_z)\eta^{\mu\nu} \tag{3.46}$$

$$= \; g_{5d}\frac{1}{kz}f^{abc}(p^\mu - q^\mu) \tag{3.47}$$

$$= \; g^2_{5d}\frac{1}{kz}\eta^{\mu\nu}(f^{eab}f^{ecd} + f^{ead}f^{ecb}) \tag{3.48}$$

The derivatives ∂^1_z and ∂^2_z in (3.46) are to be contracted with the gauge boson lines. while p^μ and q^μ in (3.47) are the momenta of the A_5 lines. There are no 3 or 4 A_5 vertices because of the antisymmetry of f^{abc}. The Feynman rules for other bulk fields can be derived analogously, with due regard for the factors of metric at the vertices. For example. a ϕ^4 vertex would have a factor of $(kz)^{-5}$, while a $\phi^2 A_\mu A^\mu$ vertex would go like $(kz)^{-3}$. Ghosts, which are scalars, technically come from terms compensating for the gauge invariance of $F_{MN}F^{MN}$, so they have $(kz)^{-1}$ vertices.

3.6 Limits of the Green's functions

Before we evaluate the quantum effects, we will study the propagator in various limits. To do this, we find it convenient to work with Euclidean momentum. Recall that the Green's function for the massless vector boson is:

$$G_q(u,v) = kuv\frac{[\mathcal{I}_0(q/k)\mathcal{K}_1(qu) + \mathcal{K}_0(q/k)\mathcal{I}_1(qu)][\mathcal{I}_0(q/T)\mathcal{K}_1(qv) + \mathcal{K}_0(q/T)\mathcal{I}_1(qv)]}{\mathcal{I}_0(q/k)\mathcal{K}_0(q/T) - \mathcal{K}_0(q/k)\mathcal{I}_0(q/T)}. \tag{3.49}$$

One advantage of this form is that the modified Bessel functions, $\mathcal{I}$ and $\mathcal{K}$. have limits which are exponentials, while the ordinary Bessel functions oscillate.

The first regime we consider is $q \ll T$:

$$G_{q \ll T}(u, v) \to -\frac{1}{R q^2}.$$ (3.50)

This is what we expect: at low energy, only the zero mode of the gauge boson is accessible. Its profile is constant so G is naturally independent of u and v. The factor of R is absorbed in the conversion from 5D to 4D couplings.

It is also useful to consider the next term in the small q expansion of G_q at a point u in the bulk. This will tell us the size of the 4-Fermi operator which comes from integrating out the excited KK modes.

$$G_0(u, u) = -\frac{k}{4T^2 k^2 R^2} + \frac{1}{2}u^2 \left(k + \frac{1}{R}\right) - \frac{1}{R}u^2 \log ku + \frac{1}{4k^3 R^2}.$$ (3.51)

On the Planck brane and TeV branes respectively, it is:

$$G_0\left(\frac{1}{k}, \frac{1}{k}\right) \approx -\frac{k}{4T^2 k^2 R^2}$$ (3.52)

$$G_0\left(\frac{1}{T}, \frac{1}{T}\right) \approx -\frac{k}{2T^2}.$$ (3.53)

The additional $(kR)^2$ suppression on the Planck brane over the TeV brane can be understood from the KK picture. The light modes have greater amplitude near their masses, which are near the TeV brane. While there are the same number of heavier modes, localized near the Planck brane, these are additionally suppressed by the square of their larger masses. This has been made quantitative by Davoudiasl et al. in [37]. They calculated the equivalent of G_0 using the KK mode propagators:

$$G_0(z, z) = -k \sum \chi_i(z)^2 \frac{1}{m_i^2}$$ (3.54)

and showed how this number is constrained by precision tests of the Standard Model:

$$V = -M_W^2 \cdot R G_0 < 0.0013.$$ (3.55)

Using our formula, this forces $T > 8.9$ TeV if fermions are on the TeV brane and $T > 197$ GeV if fermions are on the Planck brane (for $kR \approx 32$).

With mass terms. the Green's function satisfies:

$$\left[\partial_z^2 - \frac{1}{z}\partial_z + p^2 - \frac{1}{k^2z^2}\left(a^2k\delta(z - \frac{1}{T}) + \bar{a}^2k\delta(z - \frac{1}{k}) + m^2k^2\right)\right]G_p^{1,m}(z,z') =$$
$$= zk\delta(z - z').$$
(3.56)

We can work through the same analysis as in the massless case. We find that if m or $\bar{a}$ is not zero, or if a is very large, then the propagator is constant at low momentum. For a non-zero bulk mass:

$$G_0^{1,m}(u,v) = -\frac{kuv}{(ku)^\nu(kv)^\nu}\frac{(1 + \nu + (\nu - 1)(ku)^{2\nu})(1 + \nu + (\nu - 1)(Tv)^{2\nu})}{2\nu(\nu^2 - 1)}.$$
(3.57)

This tells us the strength of the 4-Fermi operators generated by integrating out heavy fields. For example. if we have a unified model where the X and Y bosons get a bulk mass of order k, then on the Planck and TeV branes ($\nu > 1$):

$$RG_0^{1,m}\left(\frac{1}{k},\frac{1}{k}\right) \approx -\frac{kR}{k^2(\nu - 1)}$$
(3.58)

$$RG_0^{1,m}\left(\frac{1}{T},\frac{1}{T}\right) \approx -\frac{kR}{T^2(\nu + 1)}.$$
(3.59)

We know that constraints from proton decay force this number to be smaller than $1/(10^{16}\text{GeV})^2$. In particular. we are safe on the Planck brane if $k > 10^{16}$ GeV for any non-zero bulk mass $m > k/T$. On the TeV brane, however, there is no value of the bulk mass which will sufficiently suppress proton decay; it is suppressed by at most $1/T^2$. We clearly need to prevent this contribution. We discuss this later in the unification section.

Increasing q, we find that for $q \gg T$, but $qu \ll 1$ and $qv \ll 1$:

$$G_q(u,v) \to -\frac{k}{q^2\left(\log(2k/q) - \gamma\right)},$$
(3.60)

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant. This is valid on the Planck brane at $u = v = 1/k$ for $q < k$. In particular, it confirms results of [13, 82] that there is a tree level running of the coupling with q. For $q \gg T$ on the TeV brane,

$$G_q\left(\frac{1}{T},\frac{1}{T}\right) \to -\frac{k}{qT}.$$
(3.61)

That the propagator goes as $1/q$ instead of $1/q^2$ in this regime is evidence of what we noted in the effective theory: there are q/T effectively massless modes which contribute at energy q. It is also evidence that this propagator is not valid for $q > T$ on the TeV brane.

Next, we look at $q \gg T$ and $qu, qv > 1$. Here the propagator looks like:

$$G_q(u, v) \to -\frac{k\sqrt{uv}}{2q}e^{-q(v-u)}. \tag{3.62}$$

The $1/q$ dependence has the same explanation as (3.61). Note that the propagator vanishes unless u and v are nearly coincident in the fifth dimension. Finally, we can consider very large energy, $q \gg k$:

$$G_q(u, v) \to -\frac{k\sqrt{uv}}{q}\frac{\cosh(q(u - 1/k))\cosh(q(1/T - v))}{\sinh(q(1/T - 1/k))}. \tag{3.63}$$
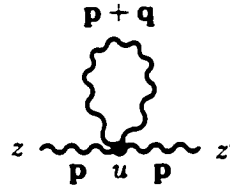
Since we are at energies much higher than the curvature scale, $q \gg k$, we get a result very similar to the propagator in flat space with the fifth dimension bounded at $1/k$ and $1/T$:

$$G_q^{flat}(u, v) \to -\frac{1}{q}\frac{\cosh(q(u - 1/k))\cosh(q(1/T - v))}{\sinh(q(1/T - 1/k))}. \tag{3.64}$$

3.7 Regulating 5D loops

From studying the 5D propagator in the previous section, we have learned that it cannot be trusted for $qu \gg 1$. This is to be expected if the physical cutoff is around k, since the cutoff on the momentum integral will scale with position in the fifth dimension. So we understand that we need a position-dependent cutoff on four-dimensional momentum. This is in accord with the holographic investigations of Chapter 2. The obvious way to implement this cutoff is to integrate up to momentum $q = \Lambda/(kz)$ at a point z in the bulk. But we will now show that the Green's function must also be renormalized. The correct procedure is to recompute the Green's function at an energy q with boundary conditions from an effective IR brane at $z = \Lambda/(kq)$, and then perform the integral.

Consider the following diagram, which contributes to the gauge boson self energy:

$$\times \quad g_{5D}^2 \int d^4q \int \frac{du}{ku} G_p(z,u) G_p(u,z') G_{p+q}(u,u) \qquad (3.65)$$

We will eventually be concerned with the correction to the zero mode propagator, so we set the external momentum $p = 0$. The low energy propagator is given by (3.50) which is independent of u and v. Since the tree level potential is proportional to $g_5^2/(Rp^2)$ we make the identification $g_{5D}^2 = g_{4D}^2 R$ as before. Also, we will assume the Ward identities are still satisfied and pull out a factor of q^2/p^2. Then the integral reduces to:

$$\frac{1}{Rp^2} g_{4D}^2 \int q\,dq \int \frac{du}{ku} G_q(u,u). \qquad (3.66)$$

The $1/(Rp^2)$ in front corresponds to the tree level propagator we are modifying. That factor of $1/R$ gets absorbed when we cap the ends with a g_{5D}^2 in a full S-matrix calculation.

Later on, we will calculate this integral exactly, but for now, we only which to elucidate the regularization scheme. So for a toy calculation, we will pretend that there is only a zero mode, and so $G_p(u,u)$ has the $p \ll T$ form (3.50) at all energies. First, suppose we have a flat cutoff, at $q = \Lambda$. We know this is wrong, but if we just have the zero mode, it should give precisely the 4D result. Indeed,

$$\int_\mu^\Lambda q\,dq \int_{1/k}^{1/T} \frac{du}{ku} \frac{1}{q^2 R} = \log\left(\frac{\Lambda}{\mu}\right) \qquad (3.67)$$

which is just what we want. Now, suppose we cut off u at $\Lambda/(kq)$, with this Green's function. Then we have:

$$\int_\mu^\Lambda q\,dq \int_{1/k}^{\Lambda/(kq)} \frac{du}{ku} \frac{1}{q^2 R} = \frac{1}{2kR} \log^2\left(\frac{\Lambda}{\mu}\right) \approx \frac{1}{2} \log\left(\frac{\Lambda}{\mu}\right). \qquad (3.68)$$

where we have taken $\mu = \Lambda T/k$ in the last step. Only half the contribution of the zero mode shows up because at an energy q, we are only including $\Lambda T/(kq)$ of it.

Now suppose the IR brane were at $u = \Lambda/(kq)$ instead of $u = 1/T$. Then, at any energy, there would always be an entire zero mode present. The Green's function would not have a sharp cutoff, but would get renormalized with Neumann boundary conditions

appropriate to its energy scale. Of course, the physical brane is still at $u = 1/T$, but the Green's function sees the cutoff as an effective brane. With this regularization, our integral is:

$$\int_\mu^\Lambda q \, dq \int_{1/k}^{\Lambda/(kq)} \frac{du}{ku} \frac{k}{q^2 \log (\Lambda/q)} = \log \left(\frac{\Lambda}{\mu} \right).$$
(3.69)

For further illustration, we can work with the full propagator, instead of just the zero-mode approximation, using our new regularization scheme. It is natural to split the integral into two regions, where the propagator can be well-approximated. We can use (3.60) for small qu and (3.62) for large qu. The small region gives, cutting off at $qu = c$:

$$\int_\mu^{ck} q \, dq \int_{1/k}^{c/q} \frac{du}{ku} \frac{k}{q^2 (\log (2k/q) - \gamma)} \approx \log \left(\frac{ck}{\mu} \right).$$
(3.70)

This is the contribution of one gauge boson, although it is not exactly the ground state. The large qu region gives:

$$\int_\mu^{ck} q \, dq \int_{c/q}^{\Lambda/(kq)} \frac{du}{ku} \frac{ku}{2q} + \int_{ck}^\Lambda q \, dq \int_{1/k}^{\Lambda/(kq)} \frac{du}{ku} \frac{ku}{2q} =$$
$$= \frac{\Lambda - ck}{2k} \log \left(\frac{ck}{\mu} \right) + \frac{\Lambda}{2k} \log \left(\frac{\Lambda}{ck} \right) - \frac{\Lambda - ck}{2k}.$$
(3.71)

This represents, roughly, the additional contribution from the excited modes. In total, there is a log piece, similar to the 4D log but enhanced by a factor of $(\Lambda/k - c)/2$, and a constant piece proportional to Λ. For relatively low values of Λ, the logarithm will dominate, and theory looks four-dimensional. The constant piece contributes to threshold corrections. Later on, when we calculate the β-function exactly in section 3.10, we will find similar qualitative understanding to this rough analytic approximation.

Our regularization scheme applies just as well to higher-loop diagrams. We can define the Green's function $G_p(u, v)$ as normalized with a brane at $\Lambda/(pk)$, and zero for $v > \Lambda/(kp)$. This automatically implements the cutoff, and we don't have to worry about how to associate the z of a vertex with the momentum of a line.

As a final justification of our regularization scheme, we can look at a renormalization group interpretation through AdS/CFT [110, 15]. It is well known that scale transformations in the CFT correspond to translations in z. But a scale transformation in a quantum

field theory is implemented by a renormalization group flow. It follows that integrating out the high-momentum degrees of freedom in the 4D theory should correspond to integrating out the small z region of the 5D theory. Suppose our 5D Lagrangian is defined at some scale M. This scale is associated not only with the explicit couplings in the Lagrangian, but also the boundary conditions with which we define the propagators. The high energy degrees of freedom are not aware of the region with $z > (\Lambda M/k)^{-1}$, which includes the TeV brane. Therefore, we are forced to normalize the propagators with an effective virtual brane at $\Lambda M/k$. In this way, we derive the low-energy Wilsonian effective action in five dimensions. If we follow this procedure down to energies of order TeV, we can then integrate over the fifth dimension to derive the four-dimensional effective theory.

3.8 Corrections to the radion potential

As a sample calculation, we compute a two-loop contribution to the vacuum energy that determines the radion potential [54, 34]. Consider the following diagram contributing to the vacuum energy of a scalar:

$$
\propto \quad \lambda_{5D} \int_0^{\Lambda T/k} q^3 dq \int_0^{\Lambda T/k} p^3 dp \int_{1/k}^{1/T} \frac{du}{(ku)^5} S_q(u, u) S_p(u, u) \,.
$$

Here λ_{5D} is the 5D ϕ^4 coupling, which has dimensions of length. It is related to the 4D coupling by $\lambda_{4D} = \lambda_{5D}(2k^3/(k^2 - T^2)) \approx 2k\lambda_{5D}$. We have cut off the momenta at $q = \Lambda T/k$. The region of integration with q or p greater than $\Lambda T/k$ has no T dependence, and hence cannot contribute to stabilizing the extra dimension. At low energies,

$$
S_q(u, u) \approx -\frac{2k}{q^2} + ku^2 - \frac{k^3}{4}u^4 - \frac{\log(k/T)}{k} + \mathcal{O}(q^2) \,. \tag{3.72}
$$

Note the enhanced u dependence of the scalar over the vector propagator (compare (3.51)). This expansion, which is quite a good approximation of the full propagator in the region of

integration we are considering, gives a vacuum energy (ignoring the numerical constants):

$$V \approx \lambda_{4D} \lambda^4 T^4 + \mathcal{O}(k^{-2} T^6 \log T). \tag{3.73}$$

This expression has the same T-dependence as zero point energy presented in [53].

3.9 Gauge boson self-energy

In order to address the question of unification. we will now look at how various bulk fields contribute to the 1-loop β-functions. There are 6 diagrams that contribute at 1-loop: 2 gauge boson. 1 ghost, and 3 involving A_5. One of these diagrams is particularly ugly. involving ∂_z acting on the A_μ propagator. After fixing the gauge. evaluating, and summing all these diagrams, we should get a correction to the gauge boson propagator which is transverse. But there is an easier way: the background field method. The idea is to compute the effective action directly, which at 1-loop only involves evaluating functional determinants. Furthermore. we have the freedom to choose the external field to be whatever we like. In particular. we can choose it to be the piece of A_μ which is independent of z. We will see that the quantum fields for A_μ, A_5 and ghosts effectively decouple. The Ward identities will be explicitly satisfied. as the diagrams containing each type of particle will separately produce a transverse correction to the A_μ propagator.

3.9.1 Background field Lagrangian

First, we separate the gauge field into a constant external piece and a fluctuating quantum piece:

$$A_M^a \rightarrow \frac{1}{g_{5D}} \left(A_M^a + \mathcal{A}_M^a \right). \tag{3.74}$$

We have also renormalized out the coupling. Note that A_M now has mass dimension 1 as in four-dimensions. If we let D_M be the covariant derivative with respect to A_M only then

$$F_{MN}^a \rightarrow F_{MN}^a + D_M \mathcal{A}_N^a - D_N \mathcal{A}_M^a + [\mathcal{A}_M, \mathcal{A}_N]^a. \tag{3.75}$$

We also must take our gauge fixing functional to be A_M-covariant:

$$\Delta L = -\frac{1}{2g^2\xi kz}G(A)^2 = -\frac{1}{2g^2\xi kz}\left[D^\mu A_\mu^a - \xi z D_z\left(\frac{1}{z}A_5^a\right)\right]^2.$$ (3.76)

The Lagrangian is then:

$$L = -\frac{1}{4g^2kz}(F_{\mu\nu}^a + D_\mu A_\nu^a - D_\nu A_\mu^a + [A_\mu, A_\nu]^a)^2 +$$
$$+\frac{1}{2g^2kz}(F_{\mu5}^a + D_\mu A_5^a - D_z A_\mu^a + [A_\mu, A_5]^a)^2 -$$
$$-\frac{1}{2g^2\xi kz}\left[D^\mu A_\mu^a - \xi z D_z\left(\frac{1}{z}A_5^a\right)\right]^2.$$ (3.77)

At 1-loop, we only need to look at terms quadratic in the quantum fields A_M. After an integration by parts, the quadratic Lagrangian is:

$$L_2 = -\frac{1}{2g^2kz}\left(A_\mu^a\left[-(D^2)^{ab}\eta^{\mu\nu} + (D^\nu D^\mu)^{ab}\eta^{\mu\nu} - \frac{1}{\xi}(D^\mu D^\nu)^{ab}\right]A_\nu^b +\right.$$
$$\left.+ F^{a\mu\nu}[A_\mu, A_\nu]^a\right) + \frac{1}{2g^2kz}A_5^a\left(-(D^2)^{ab}\right)A_5^b +$$
$$+\frac{\xi}{2g^2kz}A_5^a D_z\left(zD_z\left(\frac{1}{z}A_5^b\right)\right) - \frac{1}{g^2kz}A_5^a[D_\mu D_z - D_z D_\mu]^{ab}A_\mu^b -$$
$$-\frac{1}{2g^2k}A_\mu D_z\left(\frac{1}{z}D_z A_\mu\right) + \frac{1}{g^2kz}F^{a\mu5}[A_\mu, A_5]^a.$$ (3.78)

We derive the ghost Lagrangian from variations of $G(A)$:

$$\frac{\delta G}{\delta\alpha} = D^\mu(D_\mu + f^{abc}A_\mu^b) - \xi z D_z\left[\frac{1}{z}(D_z + f^{abc}A_5^b)\right].$$ (3.79)

Combining (3.78) with the ghost Lagrangian, using the relation

$$A_M^a[D_R, D_S]^{ab}A_N^b = -F_{RS}^a[A_M, A_N]^a$$ (3.80)

the final quadratic Lagrangian for the pure non-Abelian gauge theory in AdS_5 is:

$$L_2 = -\frac{1}{2g^2kz}\left(A_\mu^a\left[-(D^2)^{ab}\eta^{\mu\nu} + \left(1-\frac{1}{\xi}\right)(D^\mu D^\nu)^{ab}\right]A_\nu^b + zA_\mu^a D_z\left(\frac{1}{z}D_z A_\nu^b\right)\eta^{\mu\nu} +\right.$$
$$\left.+2F^{a\mu\nu}[A_\mu, A_\nu]^a\right) + \frac{1}{2g^2kz}\left(A_5^a(-D^2)^{ab}A_6^b + \xi A_5^a D_z\left(zD_z\left(\frac{1}{z}A_5^b\right)\right)\right) +$$
$$+\frac{1}{g^2kz}\bar{c}^a\left[-(D^2)^{ab} + \xi z D_z\left(\frac{1}{z}D_z\cdot\right)\right]c^b.$$ (3.81)

Observe that the cross terms between $\mathcal{A}_\mu$ and $\mathcal{A}_5$ vanish, as expected. At this point, we will specialize to the β-function calculation we are interested in. It involves an external zero mode of A_μ, whose profile is constant in the fifth dimension. We simply the external field to be the piece of the original field which is independent of z. We also set $A_5 = 0$, as there is no external A_5 component. This lets us write ∂_z instead of D_z. Then in the Feynman gauge, $\xi = 1$, the Lagrangian is:

$$L_2 = -\frac{1}{2g^2 kz}\left(\mathcal{A}_\mu^a[-(D^2)^{ab}\eta^{\mu\nu}]\mathcal{A}_\nu^b + z\mathcal{A}_\mu^a D_z\left(\frac{1}{z}D_z\mathcal{A}_\nu^b\right)\eta^{\mu\nu} + 2F^{a\mu\nu}[\mathcal{A}_\mu.\mathcal{A}_\nu]^a\right) +$$

$$+\frac{1}{2g^2 kz}\left(\mathcal{A}_5^a(-D^2)^{ab}\mathcal{A}_6^b + \mathcal{A}_5^a D_z\left(zD_z\left(\frac{1}{z}\mathcal{A}_5^b\right)\right)\right) +$$

$$+\frac{1}{g^2 kz}\bar{c}^a\left[-(D^2)^{ab} + zD_z\left(\frac{1}{z}D_z\cdot\right)\right]c^b. \tag{3.82}$$

We see immediately that the fields have the propagators we derived before (3.30), in the Feynman-'t Hooft gauge. There are no cross terms between A_μ and A_5 and we can evaluate the functional determinant for each field independently. In particular, A_5 is seen as a scalar field transforming in the adjoint representation of the gauge group, with the same Green's function as a vector with bulk mass $m^2 = -1$.

3.9.2 Functional determinants

We can now evaluate the functional determinants using standard textbook techniques [78]. There are two diagrams which contribute, one is spin-dependent (from the $F^{a\mu\nu}$ vertex in (3.82)) and vanishes for scalars, and the other is spin-independent. (The third diagram from the quartic interaction does not contribute as $d \to 4$ in dimensional regularization, so we will ignore it for simplicity.) Both diagrams are independently transverse in the external momentum. This is evidence that the Ward identity for the 5D gauge invariance is working.

The spin-dependent diagram gives:

$$z \sim\!\!\!\!\!\!\!\!\!\left\{\!\!\!\!\!\begin{array}{c} p+q \\ u\ v \\ q \end{array}\!\!\!\!\!\right\}\!\!\!\!\!\!\!\!\!\sim z' \quad = -\frac{1}{2}\int \frac{d^4p}{(2\pi)^4}A^a_\mu(-p)A^b_\mu(p)\int \frac{d^4q}{(2\pi)^4}\frac{p^2\eta^{\mu\nu}-p^\mu p^\nu}{q^2(p+q)^2}4C_r C(j)\ \times$$

$$q^2(p+q)^2 2\int_{1/k}^{\Lambda/(kq)}\frac{du}{ku}\int_u^{\Lambda/(kq)}\frac{dv}{kv}G_q(u,v)G_{q-p}(u,v) \qquad (3.83)$$

where $C(j) = 2$ for vectors and zero for scalars and C_r is Dynkin index for the appropriate representation. Since we are interested in the vacuum polarization in the $p \to 0$ limit, and the transverse projector is already manifest, we can simply set $p = 0$ in the uv integrals. Now change variables to $y = qu$ and $z = qv$ and set $k = 1$. Then the second line above becomes

$$I(\Lambda, q) = q^4 2\int_q^\Lambda \frac{dy}{y}\int_y^\Lambda \frac{dz}{z}G_q\left(\frac{y}{q},\frac{z}{q}\right)^2. \qquad (3.84)$$

The point of doing this is that the integrand now contains the square of:

$$G_q\left(\frac{y}{q},\frac{z}{q}\right) = \frac{yz}{q^2}\frac{[\mathcal{I}_0(q)\mathcal{K}_1(y)+\mathcal{K}_0(q)\mathcal{I}_1(y)][\mathcal{I}_0(\Lambda)\mathcal{K}_1(z)+\mathcal{K}_0(\Lambda)\mathcal{I}_1(z)]}{\mathcal{I}_0(q)\mathcal{K}_0(\Lambda)-\mathcal{K}_0(q)\mathcal{I}_0(\Lambda)}. \qquad (3.85)$$

The $1/q^4$ in G_q^2 cancels the q^4 prefactor in (3.84), leaving a dimensionless number multiplying the standard 4D integral. This is not strictly true, as $I(\Lambda, q)$ still has a weak dependence on q. But quite generally, we can write $I(\Lambda, q) = I_0(\Lambda) + I_1(\Lambda)\frac{q}{k} + \cdots$, so for $q \ll k$, $I = I_0$ is a fine approximation. Anyway, the background field method at 1-loop cannot give us reliable information about additional divergences, or threshold corrections. The best we can do is to use the degree to which I is not constant as a rough measure of the size of the additional corrections.

The other diagram, which is spin-independent, contributes:

$$-\frac{1}{2}\int \frac{d^4p}{(2\pi)^4}A^a_\mu(-p)A^a_\mu(p)\int \frac{d^4q}{(2\pi)^4}\frac{(2q+q)^\mu(2q+p)^\nu}{q^2(p+q)^2}C(r)d(j)\ \times$$

$$\times q^2(p+q)^2 2\int_{1/k}^{\Lambda/(kq)}\frac{du}{ku}\int_u^{\Lambda/q}\frac{dv}{kv}G_q(u,v)G_{q+p}(u,v), \qquad (3.86)$$

where $d(j)$ is the number of spin components. The second line is exactly the same integral expression. $I(\Lambda, q)$. as before. While the tensor structure of the first line is not explicitly transverse in the external momentum. it is in fact transverse after the q integral is performed (in dimension regularization as $d \to 4$).

Each particle will have a different value for I_0. We must replace G_q with the appropriate propagator (cf. section 3.4.2) for each particle and redo the integrals in each case. We shall call the result $I_0^{\sigma,m}$. corresponding to $G_p^{\sigma,m}$ from section 3.4.2. Observe that as with the R_ξ gauges. A_5 and the ghosts have kinetic terms corresponding to spin 1. so they will both have $\sigma = 1$. In particular, I_0 for ghosts is identical to I_0 for A_μ.

The reason these diagrams are relevant is that they directly produce F^2 terms in the effective action. Indeed, the Fourier transform of the quadratic terms in F^2 is:

$$-\frac{1}{4g^2} \int d^4x \frac{dz}{kz} F_{\mu\nu} F^{\mu\nu} = -\frac{R}{2g^2} \int \frac{d^4p}{(2\pi)^4} A_\mu^a(-p) A_\nu^b(p)(p^2 \eta^{\mu\nu} + p^\mu p^\nu). \qquad (3.87)$$

Since A_μ is independent of z, we have performed the z-integral explicitly. So we see that we have calculated a correction to the dimensionless coupling $g_{5D} R^{-\frac{1}{2}}$. Equivalently, we have calculated the running of g_{5D} itself. once we absorb the factor of R into the coefficient of the logarithm. The result for the 1-loop β-function is:

$$\beta(g_{5D}) = -\frac{g_{5D}^3}{4\pi^2} \frac{1}{R} C_2(G) \left(\frac{11}{3} I_0^{1,0} - \frac{1}{6} I_0^{1,i}\right) + matter. \qquad (3.88)$$

It may appear that we have calculated the running for only the zero mode of the 5D gauge boson. But gauge invariance implies that there can only be one 5D coupling. at any energy scale. So we have in fact calculated the β-function of every mode. If we define $g_{4D} \equiv g_{5D} R^{-\frac{1}{2}}$, then the 4D β-function is:

$$\beta(g_{4D}) = -\frac{g_{4D}^3}{4\pi^2} C_2(G) \left(\frac{11}{3} I_0^{1,0} - \frac{1}{6} I_0^{1,i}\right) + matter. \qquad (3.89)$$

Note that the sign of the A_5 contribution is opposite to that of the ghosts (which contribute $\frac{1}{3} I_0^{1,0}$ to the above expression). This is because they have opposite statistics which changes the sign of the exponent of the functional determinant. It is easy to understand this result: the ghosts remove the two unphysical polarizations of A_μ. and A_5 adds back one of them.
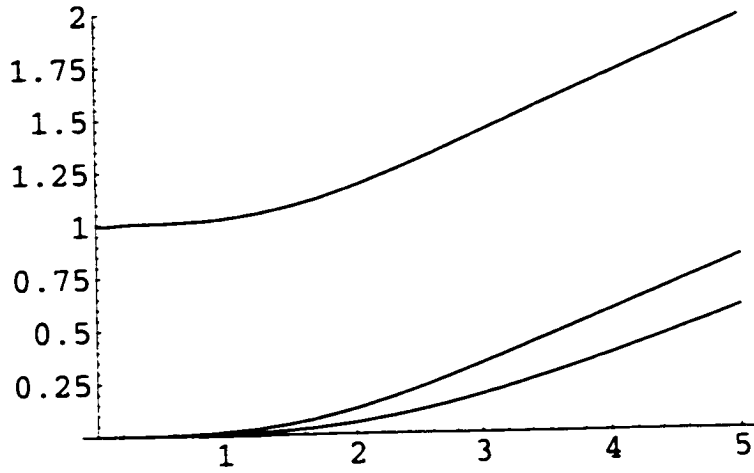
Figure 3.1: I_0 as a function of Λ/k: from top to bottom: massless vector, massive vector ($m = 1$), and Dirichlet vector.

3.9.3 Numerical results for $I(\Lambda, q)$

The function $I_0(\Lambda)$ is shown in figure (3.1), and $I(\Lambda, q)$ is shown in figures 3.2 and 3.3. Since we need need to perform a Wick rotation to evaluate the 4D integrals, we used the Euclidean propagators in calculating $I(\Lambda, q)$. Of course, the constant piece, I_0, is independent of $q \to iq$, and we have confirmed this numerically. It turns out that the integrals converge faster in Euclidean space. We can see from figure 3.1 that I_0 is roughly proportional to the number of KK modes running around the loop. Because of our boundary conditions, the effective spacing between the KK modes at an energy q is $\pi qk/\Lambda$. For the massless 5D vector, at energy q there is a massless mode, plus approximately $q/(\pi qk/\Lambda) = \Lambda/(\pi k)$ other modes visible. This fits roughly with figure 3.1. The discrepancy is due to the fact that the spacing is not precisely π for the lowest modes, and that the sum of the higher modes is not completely negligible. Notice that the I_0's of massive and Dirichlet cases, which have no massless zero mode, are about 1 less than I_0 for the massless vector. When Λ is bigger than k, and the β-function is correspondingly higher, the unification scale will be lower. A theory of accelerated unification was also considered in [4] in a different scenario.

Figures 3.2 and 3.3 can be understood similarly. As $q \to k$, the branes approach each other. For the massless case, there is always one complete mode, the zero mode. In this
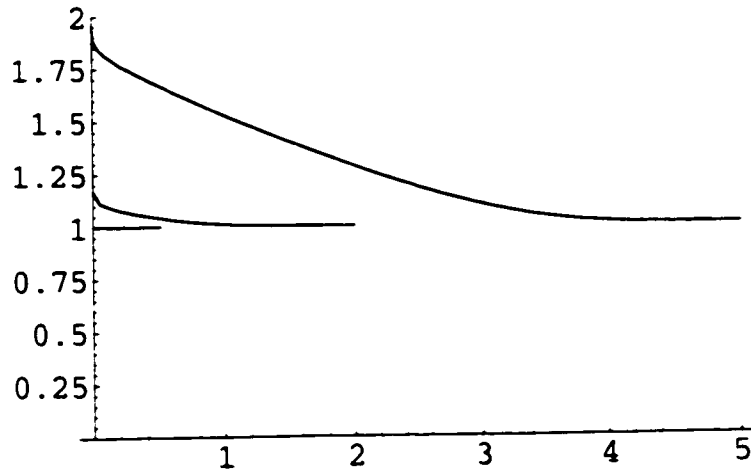
Figure 3.2: $I(\Lambda.q/k)$ as a function of q/k with Λ fixed. From top to bottom. $\Lambda = 5k. 2k.$ and $0.5k$.

$I_0^{\sigma,m}(\Lambda)$	massless vector	massive vector $m = 1$	massive vector $m = 5$	massive vector $m^2 = -1$	massless Dirichlet vector	massless scalar	massive scalar $m = 1$
$\Lambda = 0.5$	1.007	0.001	0.000	0.018	0.000	1.000	0.000
$\Lambda = 1$	1.024	0.013	0.001	0.147	0.005	1.005	0.004
$\Lambda = 5$	1.954	0.820	0.178	1.411	0.581	1.581	0.525

Table 3.2: Values of $I_0(\Lambda)$ for fields of various spin and mass.

limit the theory looks 4-dimensional. For the massive case, the zero mode exists as well. and so it approximates the massless case. With Dirichlet conditions, the zero mode is eliminated, so the function goes to zero. Note that if we had left the boundary conditions at $1/T$, this function would have gone to zero as $q \to \Lambda$ for any of the cases. We list the values of $I_0(\Lambda)$ for various cases in table 3.2.

Recall that in the Feynman gauge. ghosts and A_5 get the I_0 of vectors and A_5 has an effective mass $m = 1$. The numbers in this table can be predicted approximately from the KK masses in table 3.1. We can see that $I_0(\Lambda)$ is approximately the number of modes with mass less than $\Lambda T/k$. This gives us a very useful intuition for seeing how changing the field content affects unification, as we will now illustrate.
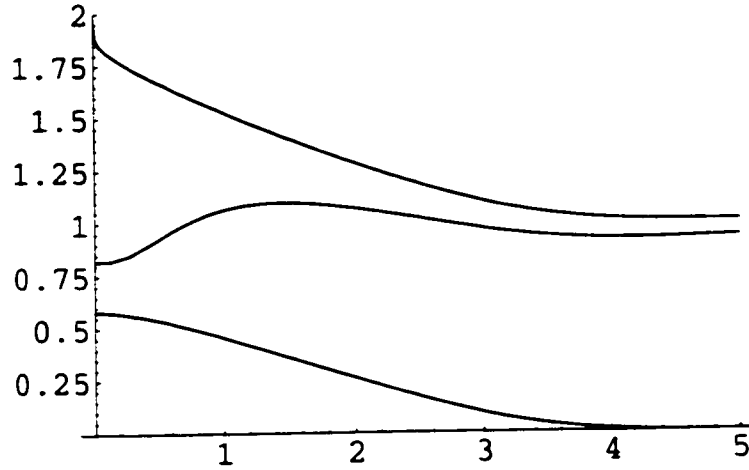
Figure 3.3: $I(\Lambda = 5k, q/k)$. From top to bottom: massless vector, massive vector (with bulk mass $m = 1$), and Dirichlet vector.

3.10   Coupling constant unification

In order to study coupling constant unification, we need to choose a particular model. Because the main motivation of this work is to solve the hierarchy problem using the warp factor, all weak-scale masses should generated from Higgs scalars confined to the TeV brane. We will consider three possible scenarios. The first is that there is no unified group. Indeed, the generic prediction of fundamental theories is only that there should be one coupling constant at high energy, not that there should be a unified group. We put the 3-2-1 gauge bosons in the bulk, and the Higgs and fermions on the TeV brane. From the CFT point of view, the TeV brane fields are to be viewed as condensates. So we can expect there to be a number of bulk fermions or scalars transforming as electroweak doublets, which condense to form the Higgs. In this case, We have to assume there is at a fundamental level a reason to assume the U(1) is normalized in a way consistent with a GUT model. For this, additional physics assumptions are necessary.

A second possibility is that there is a unified group, such as SU(5). If the doublet Higgs is part of a larger multiplet, such as a vector **5**, then the triplet will necessarily have TeV scale excitations, leading to proton decay. This is the standard doublet-triplet problem [85]. One 4D solution is to couple the triplet to a missing partner, which gives it a large mass

and decouples it from the Standard Model. However. this solution will not work with a TeV brane triplet. because its mass can be at most TeV. Instead. one can for example implement the pseudoGoldstone boson method. as in [85. 30]. Briefly. the idea is to postulate a weakly gauged global symmetry. such as $SU(6) \times SU(6)$. This is broken by an adjoint Σ and two fundamentals H and $\bar{H}$. The doublet Higgs arises as a pseudoGoldstone boson. and there is no triplet at all. This sort of condensation also seems likely from the CFT point of view. where all TeV brane fields are composites.

Dimension 6 operators that violate baryon number pose a potential problem. As mentioned before, the X and Y bosons should not couple on the TeV brane. This can be done by having the gauge symmetry not commute with the orbifold transformation [68. 57] so that the WZ have positive parity under Z_2. but the XY have negative parity and have vanishing amplitude on the TeV brane. An additional baryon number symmetry should be imposed on the brane to forbid dangerous operators. Notice that the $U(1)$ on the brane might have a kinetic term. and therefore a coupling. not determined by unification. If this is the case. one would hope the brane couplings for $U(1)$. $SU(2)$. and $SU(3)$ are all big so that a mechanism such as the one in [113] would apply.

A third possibility is that we don't use the TeV brane to generate the weak scale. as in [82]. The hierarchy problem must be solved some other way. such as using supersymmetry.

Now return to the first scenario. with no unified group. The TeV-brane particles will contribute to running only up to $q = \Lambda T/k$. After this, they contribute like the bulk fields which represent their preonic constituents in the CFT. There are many possibilities for what these can be, but for the sake of illustration, we will assume they are either fermions or scalars which have the 3-2-1 quantum numbers of the Standard Model Higgs. Then the

1-loop β-functions lead to the following running:

$$\alpha_1^{-1}(M_{GUT}) = \alpha_1^{-1}(M_Z) - \frac{2}{\pi}\left(\frac{n_g}{3} + \frac{3}{5}\frac{n_s}{24}\right)\log\left(\frac{M_{GUT}}{M_Z}\right) \tag{3.90}$$

$$\alpha_2^{-1}(M_{GUT}) = \alpha_2^{-1}(M_Z) - \frac{2}{\pi}\left(-\frac{11}{6}I_0^{1.0}(\Lambda) + \frac{1}{12}I_0^{1.i}(\Lambda) + \frac{n_g}{3} + \frac{n_s}{24}\right)\log\left(\frac{M_{GUT}}{M_Z}\right) \tag{3.91}$$

$$\alpha_3^{-1}(M_{GUT}) = \alpha_3^{-1}(M_Z) - \frac{2}{\pi}\left(-\frac{11}{4}I_0^{1.0}(\Lambda) + \frac{1}{8}I_0^{1.i}(\Lambda) + \frac{n_g}{3}\right)\log\left(\frac{M_{GUT}}{M_Z}\right). \tag{3.92}$$

The $I_0^{1.i}$ terms come from the contribution of A_5 which has an effective bulk mass $m = 1$. $I_0(\Lambda)$ is roughly equal to the number of KK modes with mass below $\Lambda T/k$. It is defined exactly in the previous section. There are additional terms in the above equations proportional to $I_1(\Lambda)\frac{M_{GUT}}{k}$, which we will assume to be small. For $I_0 = 1$ (which occurs as $\Lambda \to 0$), these are just the standard equations for 4 dimensional running. For bulk scalars, the effect is $n_s \to n_s I_0^{2.0}(\Lambda)$. If the bulk preons are n_f Majorana fermions, which would prevent them from picking up a large bulk mass, then this term should be $n_s \to 2n_f I_0^{1/2.0}(\Lambda)$, where $I_0^{1/2.0}(\Lambda)$ comes from 5D massless fermion loops, and the factor of 2 is because fermions contribute twice as much as complex scalars to the gauge boson self-energy. The same modifications should be made for the n_g term, but these do not affect unification, so we will ignore them.

To show that unification can be improved, we pick a specific model. We choose $\Lambda = 1$ and put 4 Majorana fermions in the bulk. We leave $n_g = 3$, to facilitate the comparison with the Standard Model. Then we use the numerical values: $I_0(1) = 1.024$, $I_0^{1.i}(1) = 0.013$, $I_0^{1/2.0}(1) = 1.009$. We will use the observed values [72] of $\alpha_3(M_Z) = 0.1195$, $\alpha_e^{-1}(M_Z) = 127.934$, and $\sin^2\theta_w = 0.23107$ at the Z-boson mass $M_Z = 91.187$ GeV. The couplings are shown for this case in figure 3.4. The Standard Model is shown for comparison.

As Λ is increased, $I_0(\Lambda)$ and $I_0^{1/2.0}(\Lambda)$ grow at roughly the same rate. The net effect is that the β-functions basically scale uniformly with Λ. This will not have much of an effect on whether unification occurs, but it can drastically change the scale. For example, if we take $\Lambda = 5$, the scale drops from 10^{14} to 10^8. So, if we expect unification near the string scale, we must have $\Lambda \approx 1$. We assumed that $I(\Lambda, q)$ was constant. As we mentioned
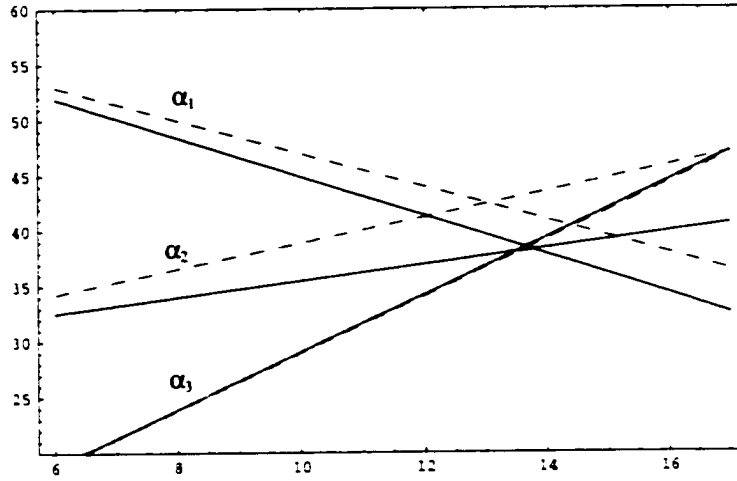
Figure 3.4: α^{-1} as a function of $\log_{10}(M_{GUT}/M_Z)$. Unification of couplings for $\Lambda = k$ (solid lines). The standard model is shown for comparison (dashed lines).

before, the additional effect from the first order term. I_1 is suppressed by M_{GUT}/k. So if $M_{GUT} \ll k$ it is negligible. but if $M_{GUT} \approx k$, it can be significant. Even though our regularization scheme cannot tell us the precise effect from the 1-loop calculations. we can easily determine the sign. $I_1(\Lambda)$ is the the slope of the curves in figure 3.2, and is always negative. So for $M_{GUT} \approx k$, these corrections will lower the unification scale.

Now consider the second scenario, where the XY bosons are decoupled from the Standard Model by changing their Z_2 parity. Then the coefficient of the log picks up an additional piece, proportional to $I_0^d(\Lambda)$. as listed in the massless Dirichlet vector column of table 3.2. Since complete multiplets do not contribute to unification. we can simplify equations (3.90)–(3.92) by substituting:

$$I_0(\Lambda) \to I_0(\Lambda) - I_0^d(\Lambda).$$
(3.93)

The main effect of this is that it allows us to go to higher values of Λ without lowering the unification scale too much. For example, $I_0(5) = 1.954$, but $I_0(5) - I_0^d(5) = 1.473$. This makes $M_{GUT} \approx 10^{11}$ rather than 10^8 as it would be without these additional states. We can also put in fields transforming as adjoints or fundamentals under the GUT group with Dirichlet or Neumann components. There are too many possibilities for us to examine them here, but it is fairly straightforward to work out how they affect unification.

Finally consider the third scenario, where matter is on the Planck brane. Here SU(5) might be broken by a massive adjoint in the standard way, and the triplet might be coupled to some heavy missing partners. Proton decay is suppressed by at least k^{-2}, as we can see from (3.58). Unification is similar to the second scenario, but we must make the replacement $I_0(\Lambda) \to I_0(\Lambda) - I_0^{1,m}(\Lambda)$ in equations (3.90)–(3.92). From table 3.2, we can see that if the XY bulk mass is $m = 1$, the relevant value is $I_0(5) - I_0^{1,1}(5) = 1.134$. This yields $M_{GUT} \approx 10^{13}$ even for Λ as big as $5k$. However, if the bulk mass of X and Y is too large, for example $m = 5$, then $I_0(5) - I_0^{1.5}(5) = 1.776$, which leads to $M_{GUT} \approx 10^9$. It is clear that there is a lot of room for detailed model building, which we leave for future work.

3.11 Conclusions

We have shown how to perform Feynman diagram calculations consistently in five-dimensional anti-de Sitter space. Our regularization scheme is inspired by the AdS/CFT duality [74, 115, 56]. There we see that scale transformations in the 4D theory are equivalent to z-translations in the 5D theory. Therefore, we can understand how following the renormalization group flow down to the scale μ corresponds to integrating out the fifth dimension from $z = 1/k$ up to $z = \Lambda/(\mu k)$. The correct implementation of this is to renormalize 5D propagators as if the IR brane were at the relevant energy scale for the computation. Not only does this ensure that at a position z the UV cutoff is mediated by the warp factor, but also that a complete 4D mode of the bulk field is always present. The regulator was shown at length to be compatible with holography in Chapter 2.

The original Randall-Sundrum scenario was presented as a solution to the hierarchy problem. It is now clear that it is also consistent with coupling constant unification. With Standard Model matter confined to the TeV brane, the maximum unification scale is naively seen to be $\Lambda T/k$. From the CFT picture, we know this cannot be true. Now we understand how higher scales are reached in 5D as well. We have briefly described some possible unification scenarios. None of them are perfect, and more detailed model building is called for, but it is clear that unification can be improved from the Standard Model. The key is

that even though gauge bosons are in the bulk. running is effectively four dimensional.

Although we originally intended to tie up a loose end of the AdS/CFT picture. it seems like we have revealed a whole new tangle. There are many directions to go from here. and the work is to a large degree unfinished. There are many threshold corrections that we have not yet included. These include sub-leading terms in the background field calculation. sub-leading terms in $I(\Lambda, q)$. and a higher loop calculation. Furthermore. the answer depends on the details of the model; here it is not only a question of the GUT group. but also the bulk fields that yield the TeV brane matter. It is well known that while supersymmetric GUTs appear to unify beautifully at 1-loop. at 2-loops unification does not occur within experimental bounds (without involved model building). It is important to see in more detail how well unification works in this model. Finally. we have been somewhat lax about the relationships among the various scales in the theory. namely $M_{GUT}. k. \Lambda k. M_{5D}. M_{Pl}. R$ and the string scale. These should ultimately be incorporated. Of course. we would also want to motivate Λ in a particular model. and furthermore understand the origin of unification and its scale at a fundamental level.

Acknowledgments

Chapter 4

Effective Field Theory for Massive Gravitons

4.1 Introduction

In previous chapters, we have seen how effective field theory for gravity can both improve our theoretical understanding of fundamental principles, such as holography, and serve as a tool for constraining phenomenological models, such as Randall-Sundrum. While these applications were all in the space-time continuum, effective field theory is an even more powerful tool in spaces that are defined to be discrete. These discrete spaces explicitly break general coordinate invariance, and effective field theory lets us parameterize the effects of this breaking on low energy physics. Moreover, we can use effective field theory to compare different theories with broken symmetry, and often use it to construct superior models. In this chapter, a effective field theory technique for studying theories with broken general coordinate invariance is introduced and used to study the massive graviton. In Chapter 5 this formalism is applied to lattice gravity and the construction of discrete gravitational dimensions.

This work was partially motivated by recent work on dynamically generating non-gravitational extra dimensions from fundamentally four-dimensional gauge theories [5, 61,

65

31]. These theories can be represented by a graph or "theory space" consisting of sites and links [7]. In some cases, at low energies the link fields become non-linear sigma model Goldstone fields, which are eaten to yield a spectrum of massless and massive gauge bosons. This spectrum may match the Kaluza-Klein tower of a compactified higher-dimensional theory and be phenomenologically indistinguishable from an extra dimension. More interestingly, theory space generalizes the notion of higher-dimensional locality. This has allowed for the construction of purely four-dimensional models that reproduce apparently higher-dimensional mechanisms in a simple context. It has also produced powerful new tools for physics beyond the standard model with no higher-dimensional interpretation whatsoever.

In previous work, gravity has been tacked on to these theories simply by minimally coupling four-dimensional Einstein gravity to the four-dimensional fields. This is a consistent thing to do, but then gravitational interactions do not respect locality in theory space. In order to have gravitational interactions be local, we would like to make gravity propagate in theory space. This is what we consider in this chapter. Actually, applying it to large discrete dimensions will be the subject of Chapter 5. I should emphasize that we are only interested in a low-energy description with sites and Goldstone link fields, which in a unitary gauge reproduces a finite spectrum of massless and massive gravitons. This is not "deconstruction", in the sense that we are not, for the moment, interested in a full UV completion of these theories. Instead, we are only trying to make sense of gravity living in discrete spaces, and will be content with understanding the structure of the low-energy effective theory describing such a scenario.

If we try to put gravity into theory space in the most straightforward way, following the gauge theory example of [5], we have no trouble at the level of the sites. We can easily endow each site with its own metric and general coordinate invariance symmetry. The difficulty arises when we try to write down the analog of the Goldstone boson link fields which must transform non-trivially under two general coordinate invariances. For instance, a four-dimensional field can only depend on one set of coordinates, not two, so how can it vary under two separate general coordinate transformations? Moreover, in a unitary

gauge where the links are eaten, we have a massless diagonal graviton and a finite tower of massive gravitons. Such theories have many strange properties, and there are doubts in the literature about whether they are consistent [24, 45, 3, 75, 26].

In this chapter, I will show that there actually is a very natural way of introducing link fields. It allows us to see that the theory of a massless graviton coupled to a finite number of massive gravitons makes sense as a consistent effective field theory valid up to energies parametrically above the particle masses. The construction of the Goldstone link fields for gravity is easiest to understand in analogy with the gauge theory case which is reviewed in detail. For example, right multiplication by an element of a gauge group in gauge theory translates to composition with a coordinate transformation in the gravity case. This allows us to define links, with the quantum numbers of four-dimensional vectors. These links have simple transformation laws under pairs of general coordinate invariances, and allow the construction of interacting Lagrangians with multiple general coordinate invariances.

After describing the general formalism, we study the case of a single graviton of mass m_g in detail. This can be understood as a two-site model in the limit where the Planck scale on one of the sites is taken to infinity. Just as for massive gauge theories, the Goldstone description is extremely useful in understanding the properties of the longitudinal polarizations of the graviton. Massive gravitons are well-known to have a number of strange properties. The mass term must have a specific Fierz-Pauli structure [47, 109]. The propagator around flat space suffers from the famous van Dam-Veltman-Zakharov (vDVZ) discontinuity [108, 116], though this disappears in anti-de Sitter (AdS) backgrounds [66, 83, 70, 69, 71]. Finally, there is the observation first made by Vainshtein [107] and further explored recently in [84, 38, 73], that the gravitational field outside of a massive object breaks down at a peculiar macroscopic distance scale much larger than the naive gravitational radius. We will understand all of these properties in a transparent way, and see that they are different reflections of a single underlying cause: the scalar longitudinal mode of the graviton becomes strongly coupled at far lower energies than naively expected. Nevertheless, we will see that we have a sensible effective field theory valid up to a cutoff Λ parametrically higher than

m_g, with $\Lambda \sim (m_g^4 M_{\text{Pl}})^{1/5}$ for the Fierz-Pauli theory. This cutoff can be pushed up to $\Lambda \sim (m_g^2 M_{\text{Pl}})^{1/3}$ in certain non-linear extensions of the Fierz-Pauli Lagrangian. We will also be able to determine, by simple power-counting, the form of all the corrections to the Fierz-Pauli Lagrangian which are generated radiatively.

With the assurance that effective theories with multiple interacting spin two fields make sense, many potential applications can be envisioned. The most obvious application is to building gravitational extra dimensions, which is discussed in Chapter 5. Other possible applications are discussed in the conclusion to this chapter, the remainder of which intimately parallels [12].

4.2   Review of gauge theory

In this section we review some aspects of the effective theory of massive spin one fields. We emphasize advantages of the effective field theory formalism [33, 27] and show why introducing a "fake" gauge invariance can be a very useful thing to do.

Consider for definiteness an $SU(n) \times SU(n)$ gauge theory, which is broken to the diagonal subgroup. At energies beneath the Higgsing scale, the theory consists only of a massive and massless gauge multiplet, or equivalently of two gauge multiplets together with the Goldstone bosons which are eaten to make the massive gauge boson heavy. In theory space language, we have a two-site model with gauge symmetry $SU(n)_1 \times SU(n)_2$, with a single bi-fundamental non-linear sigma model link field U:

$$\text{O} \xrightarrow{\ U\ } \text{O} \tag{4.1}$$

U transforms linearly under the gauge symmetries as $U \to g_2^{-1} U g_1$ and induces non-linear transformations on the Goldstone boson fields when we expand $U = e^{i\pi}$. The Lagrangian can be written as

$$\mathcal{L} = -\frac{1}{g_1^2} \text{tr} F_1^2 - \frac{1}{g_2^2} \text{tr} F_2^2 + f^2 \text{tr} |D_\mu U|^2 + \cdots \tag{4.2}$$

where $D_\mu U = \partial_\mu U + i A_{1\mu} U - i U A_{2\mu}$ is the covariant derivative. In the unitary gauge, we

can set $U = 1$ and the Lagrangian becomes

$$\mathcal{L} = -\frac{1}{g_1^2}\text{tr}F_1^2 - \frac{1}{g_2^2}\text{tr}F_2^2 + f^2\text{tr}(A_1 - A_2)^2 + \cdots \tag{4.3}$$

In the limit as we take, say, $g_1 \to 0$, the surviving massless gauge boson becomes all A_1 and completely decouples, and we are left with a single massive gauge boson $A \equiv A_2$. In this limit, the Lagrangian is simply

$$\mathcal{L} = -\frac{1}{g^2}\text{tr}F^2 + f^2\text{tr}|D_\mu U|^2 + \cdots \tag{4.4}$$

which in unitary gauge becomes

$$\mathcal{L} = -\frac{1}{g^2}\text{tr}F^2 + f^2\text{tr}A^2 + \cdots \tag{4.5}$$

So we can identify the mass of the gauge boson as $m_A = gf$.

Now, the physics described by the the Lagrangian with the Goldstone bosons included is identical to the unitary gauge Lagrangian without the Goldstone fields. The unitary gauge Lagrangian (4.5) does not have any gauge invariance, while the Goldstone boson Lagrangian (4.4) does. It is clear that this symmetry is a complete fake: we can always go to the unitary gauge where it is not there! This is *always* true for local symmetries – they are not symmetries but redundancies of description. If a theory is not gauge invariant, we can introduce Goldstone fields with appropriate transformation properties to make it gauge invariant.

However, there are important advantages to introducing the Goldstone bosons: at energies far above m_A, the Goldstones (π) become the longitudinal component of the massive gauge boson (A^L):

$$\tag{4.6}$$

From the Goldstone description it is obvious that the interactions of these longitudinal modes become strongly coupled at a scale $\sim 4\pi f \sim 4\pi m_A/g$, since this is dimensionful scale that appears in the non-renormalizable, non-linear sigma model. Since the physics

is exactly that of the unitary gauge theory, this could also have been inferred directly in unitary gauge, though the analysis would be more cumbersome and less illuminating. For instance, we can evaluate the Feynman diagrams for tree-level longitudinal gauge boson scattering. Since the polarization vector for the longitudinal gauge boson is $\epsilon_\mu \sim k_\mu / m_A$ at high energy, there is a danger that these amplitudes can become large. The tree-level amplitude for $A^L A^L \to A^L A^L$ could grow as rapidly as $g^2 (E/m_A)^4$ from this consideration, since there are four polarization vectors. However, there is a cancellation between the direct 4-point gauge interaction and exchange diagrams for this process, and the amplitude only grows as $\sim g^2 (E/m_A)^2$, which becomes strongly coupled at $\sim 4\pi m_A / g \sim 4\pi f$. This example illustrates why the Goldstone boson formalism is so powerful: it focuses on precisely the degrees of freedom that limit the regime of validity of the effective theory by becoming strongly coupled. These degrees of freedom are obscured in the unitary gauge.

The Goldstone description also allows us to determine the structure of the higher-order terms in the effective Lagrangian by simple power-counting. We expect at quantum level that in addition to the leading two-derivative terms in the non-linear sigma model, we generate higher derivative terms such as

$$\sim \frac{1}{16\pi^2} \mathrm{tr} |D_\mu U|^4, \; \frac{1}{16\pi^2} \mathrm{tr} |D^2 U|^2, \cdots \tag{4.7}$$

In unitary gauge these correspond to operators of the form

$$\frac{1}{16\pi^2} \mathrm{tr} A^4, \; \frac{1}{16\pi^2} \mathrm{tr} (\partial A)^2, \cdots \tag{4.8}$$

However, the natural size for these operators would have been hard to understand directly in unitary gauge. These operators all explicitly break gauge invariance. How would we know how to organize the various terms that break gauge invariance? Why, for instance, since the theory is not gauge invariant, do we still use the gauge invariant $F_{\mu\nu}^2$ kinetic term but only break gauge invariance through the mass term? Normally, this is justified because a general kinetic term gives rise to ghosts in the propagator. But surely we must expect that non-gauge invariant kinetic terms are generated. The standard power-counting analysis shows that non-gauge invariant kinetic terms $\sim (\partial A)^2$ are generated, as well as

non-gauge invariant A^4 type interactions. But these are down by weak-coupling factors $\sim g^2/(16\pi^2)$ relative to the leading terms, as long as the theory is treated as an effective theory with cutoff $\sim 4\pi m_A/g$. These sizes insure that all the dangerous effects associated with the non-gauge invariant propagators, such as the appearance of ghosts, are deferred to the cutoff, parametrically above m_A. Furthermore, it is easy to determine the natural size of all other gauge-violating operators effects in a systematic way.

Thus, while the Goldstone description is *physically identical* to the unitary gauge description of a massive gauge boson, it is vastly more powerful in elucidating the physics. The Goldstone description makes it clear that the theory of a massive gauge boson is sensible up to a cutoff $\sim 4\pi m_A/g$. Above this scale, an ultraviolet completion is needed. This is actually where the Goldstone description gains its full power, since finding a UV completion only necessitates finding a good UV theory that leads to the Goldstones at low energies. This can be done straightforwardly, for instance via linear sigma models (as in the standard Higgs mechanism), or QCD-like completions (as in technicolor). But none of these UV completions could have easily been guessed from the unitary gauge picture.

Summarizing, the Goldstone boson description offers three advantages to thinking about theories with massive gauge bosons: (1) It transparently encodes the interactions of the longitudinal components of the gauge bosons at high energy, and determines the cutoff of the effective theory. (2) Simple power-counting determines the natural size of all non-gauge invariant operators. (3) It helps point the way, from the bottom-up, to possible UV completions of the physics.

4.3   Building blocks for gravity in theory space

We would like to do the same kind of analysis for gravity. We will begin with the building blocks for a general theory with many "sites" endowed with different 4D general covariances. We will show how to define "link" fields with suitable non-linear transformation properties, in analogy with the gauge theory case. We will also discuss the gravitational analog of "plaquette" operators, needed to realize higher-dimensional theory spaces.

4.3.1 Sites and Links

We start with a collection of sites j each of which has its own general coordinate invariance symmetry GC_j. We will label a given set of coordinates on the site j with x_j. The GC_j symmetry is generated by $x_j^\mu \to f_j^\mu(x_j)$, where we assume the functions f_j are smooth and invertible. In this paper we will only be concerned with local physics, and therefore ignore any issues relating to the global topology of the sites.

We would now like to introduce link fields that allow us to compare objects on different sites, which transform under different GC symmetries. In order to do this in a transparent way completely parallel to the gauge theory example, let us discuss transformation properties under GC symmetries with a simple notation. A field $\phi(x)$ is a scalar if it transforms under GC as

$$\phi(x) \to \phi'(x) = \phi(f(x)) \tag{4.9}$$

We can write this in a more suggestive way in terms of functional composition:

$$\phi \to \phi \circ f \tag{4.10}$$

Similarly a vector field $a_\mu(x)$ transforms under GC as

$$a_\mu(x) \to \frac{\partial f^\alpha}{\partial x^\mu}(x) a_\alpha(f(x)) \tag{4.11}$$

We can write this more compactly by treating a as a form: $\mathbf{a} = a_\mu(x)\mathrm{d}x^\mu$. Then we just write

$$\mathbf{a} \to \mathbf{a} \circ f \tag{4.12}$$

where it is understood that $\mathrm{d}x^\mu \to \mathrm{d}f^\mu = \partial_\alpha f^\mu \mathrm{d}x^\alpha$. This clearly generalizes to all tensor fields. For example, the transformation properties of the metric $g_{\mu\nu}(x)$ are encoded in $\mathbf{g} = g_{\mu\nu}(x)dx^\mu \otimes dx^\nu$ as $\mathbf{g} \to \mathbf{g} \circ f$. Written in this way, coordinate transformations look very similar to gauge transformations on fields in a gauge theory, e.g. $\phi^\dagger \to \phi^\dagger g$, except we have composition in the place of group multiplication.

Now, suppose we have two different sites i, j, with two different general coordinate invariances $GC_{i,j}$. We would like to be able to compare fields on the two sites, which
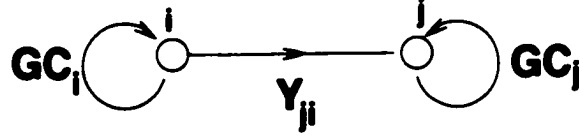
Figure 4.1: Sites i and j with corresponding $GC_i \times GC_j$ symmetries. connected by a link Y_{ji}. Under $GC_i \times GC_j$, $Y_{ji} \to f_j^{-1} \circ Y_{ji} \circ f_i$.

are charged under the different groups. In the gauge theory case. this is accomplished by introducing a link field U_{ji} which transforms under gauge transformations as $U_{ji} \to g_j^{-1} U_{ji} g_i$. The object $(\phi_j^\dagger U_{ji})$ transforms only under g_i as $(\phi_j^\dagger U_{ji}) \to (\phi_j^\dagger U_{ji}) g_i$. Similarly. in our case. we introduce a link field Y_{ji} which transforms under $GC_i \times GC_j$ as

$$Y_{ji} \to f_j^{-1} \circ Y_{ji} \circ f_i \qquad (4.13)$$

Explicitly, the link field $Y_{ji}^\mu(x_i)$ is a mapping from the site i to the site j (see figure 4.1). It associates a point on the site i with coordinate x_i^μ with a point on the site j with coordinate $Y^\mu(x_i)$: and under $GC_i \times GC_j$, we have

$$Y_{ji}^\mu(x_i) \to (f_j^{-1})^\mu(Y_{ji}(f_i(x_i))) \qquad (4.14)$$

Y_{ji} is a pullback map from site j to site i.

It is now clear how we can compare fields on different sites using the link fields. Suppose we have a field ψ_i on site i and a field ψ_j on site j which transform under GC_i and GC_j respectively as:

$$\psi_i \to \psi_i \circ f_i. \quad \psi_j \to \psi_j \circ f_j. \qquad (4.15)$$

We can construct an object Ψ out of ψ_j which transforms under GC_i by forming

$$\Psi = \psi_j \circ Y_{ji} \Rightarrow \Psi \to \Psi \circ f_i \qquad (4.16)$$

Let us see how this works explicitly for various tensor fields. Starting with a scalar $\phi_j(x_j)$ under GC_j, we can form the field

$$\Phi(x_i) = \phi_j(Y_{ji}(x_i)) \qquad (4.17)$$

which transforms as a scalar under GC_i. Similarly, out of a vector $a_{j\mu}(x_j)$ or a metric $g_{j\mu\nu}(x_j)$, we can form the objects

$$A_\mu(x_i) = \frac{\partial Y^\alpha}{\partial x_i^\mu}(x_i)a_{j\alpha}(Y_{ji}(x_i)). \quad G_{\mu\nu}(x_i) = \frac{\partial Y^\alpha}{\partial x_i^\mu}(x_i)\frac{\partial Y^\beta}{\partial x_i^\nu}(x_i)g_{j\alpha\beta}(Y_{ji}(x_i)) \quad (4.18)$$

which transform respectively as a vector and a metric under GC_i.

Note also that these expressions have the structure of induced tensors familiar from brane dynamics. We can view the site i as a space-time filling "brane" embedded in the world j. $Y_{ji}(x_i)$ is the location of a given point x_i on the brane, in the coordinates of the space-time in which the brane is embedded.

The important point is that we can low write down a Lagrangian like:

$$\mathcal{L} = \sqrt{g_i}g_i^{\mu\nu}(g_{i\mu\nu} - G_{\mu\nu})g_i^{\rho\sigma}(g_{i\rho\sigma} - G_{\rho\sigma}) \quad (4.19)$$

which is invariant under both GC_i and GC_j. We can often fix a "unitary gauge" where Y=id and this Lagrangian becomes:

$$\mathcal{L} = \sqrt{g_i}g_i^{\mu\nu}(g_{i\mu\nu} - g_{j\mu\nu})g_i^{\rho\sigma}(g_{i\rho\sigma} - g_{j\rho\sigma}) \quad (4.20)$$

which is exactly what we need to create mass terms for gravitons. (Lagrangians of the form (4.20) have been considered in [35, 36].) Introducing the $Y's$ has allowed us to introduce these mass terms in a way that is formally fully generally covariant. Just as in the analysis of the massive gauge boson, introducing the $Y's$ with this "fake" general covariance will nevertheless be extremely useful in understanding the structure of the theory.

4.3.2 Explicit Goldstone boson expansion

While the above construction is somewhat abstract, it is straightforward to expand Y and G in terms of pions and see how the two general coordinate invariances are realized explicitly. The unitary gauge has Y=id, that is $Y_{ji}^\mu(x_i) = x_i^\mu$. The transformations that leave $Y_{ji} =$ id are the diagonal subgroup of $GC_i \times GC_j$, where $f_i = f_j$. Even in situations with many Y fields where we cannot gauge fix all the Y's to the identity, it is useful to expand the

Y' around a common x as this corresponds to the small fluctuations around a common background space. We therefore expand Y as

$$Y^\alpha(x) = x^\alpha + \pi^\alpha(x) \tag{4.21}$$

where here and in what follows we have dropped the ij indices on the Y and the i index on x to avoid notational clutter.

Then, the object $G_{\mu\nu}$, from (4.18), can be expanded as:

$$
\begin{aligned}
G_{\mu\nu} &= \frac{\partial Y^\alpha(x)}{\partial x^\mu}\frac{\partial Y^\beta(x)}{\partial x^\nu}g^j_{\alpha\beta}(Y(x)) = \frac{\partial(x^\alpha + \pi^\alpha)}{\partial x^\mu}\frac{\partial(x^\beta + \pi^\beta)}{\partial x^\nu}g^j_{\alpha\beta}(x+\pi) \\
&= (\delta^\alpha_\mu + \pi^\alpha_{,\mu})(\delta^\beta_\nu + \pi^\beta_{,\nu})(g^j_{\alpha\beta} + \pi^\mu g^j_{\alpha\beta,\mu} + \frac{1}{2}\pi^\mu\pi^\nu g^j_{\alpha\beta,\mu,\nu} + \cdots) \\
&= g^j_{\mu\nu} + \pi^\lambda g^j_{\mu\nu,\lambda} + \pi^\alpha_{,\mu}g^j_{\alpha\nu} + \pi^\alpha_{,\nu}g^j_{\alpha\mu} + \frac{1}{2}\pi^\alpha\pi^\beta g^j_{\mu\nu,\alpha,\beta} \\
&\quad + \pi^\alpha_{,\mu}\pi^\beta_{,\nu}g^j_{\alpha\beta} + \pi^\alpha_{,\mu}\pi^\beta g^j_{\alpha\nu,\beta} + \pi^\alpha_{,\nu}\pi^\beta g^j_{\mu\alpha,\beta} + \cdots
\end{aligned} \tag{4.22}
$$

Now we will look at the transformation properties of g, G, Y and π, under infinitesimal general coordinate transformations generated by $f_i(x) = x + \varepsilon_i(x)$ and $f_j(x) = x + \varepsilon_j(x)$. The metrics on the sites transform as:

$$\delta g^i_{\mu\nu} = \varepsilon^\lambda_i g^i_{\mu\nu,\lambda} + \varepsilon^\lambda_{i,\mu}g^i_{\lambda\nu} + \varepsilon^\lambda_{i,\nu}g^i_{\mu\lambda} \tag{4.23}$$

$$\delta g^j_{\mu\nu} = \varepsilon^\lambda_j g^j_{\mu\nu,\lambda} + \varepsilon^\lambda_{j,\mu}g^j_{\lambda\nu} + \varepsilon^\lambda_{j,\nu}g^j_{\mu\lambda} \tag{4.24}$$

$$\delta\sqrt{g_i} = \varepsilon^\lambda_{i,\lambda}\sqrt{g_i} + \varepsilon^\lambda_i(\sqrt{g_i})_{,\lambda}, \quad \delta\sqrt{g_j} = \varepsilon^\lambda_{j,\lambda}\sqrt{g_j} + \varepsilon^\lambda_j(\sqrt{g_j})_{,\lambda} \tag{4.25}$$

Thus a Lagrangian like

$$\mathcal{L} = \sqrt{g_i}R(g_i) + \sqrt{g_j}R(g_j) \tag{4.26}$$

trivially has two general coordinate invariances. If we make the replacements (4.23)-(4.25), with independent ε_i and ε_j this Lagrangian is unchanged. Note, however, that the "hopping" Lagrangian in (4.20) is only invariant under the diagonal subgroup for which $\varepsilon_i = \varepsilon_j$.

Now, the transformation laws of the pions come from transformation of the link Y. First, under GC$_i$:

$$
\begin{aligned}
Y(x) \to Y(x') &= x + \varepsilon_i + \pi(x + \varepsilon_i) \equiv x + \pi + \delta\pi \\
\Rightarrow \delta\pi^\mu &= \varepsilon^\mu_i + \varepsilon^\alpha_i\pi^\mu_{,\alpha}
\end{aligned} \tag{4.27}
$$

Under GC_j

$$Y \;\rightarrow\; Y - \varepsilon_j(Y) = x + \pi - \varepsilon_j(x + \pi) \equiv x + \pi + \delta\pi$$

$$\Rightarrow \delta\pi^\mu \;=\; -\varepsilon_j^\mu(x + \pi) = -\varepsilon_j^\mu - \pi^\alpha\varepsilon_{j,\alpha}^\mu - \frac{1}{2}\pi^\alpha\pi^\beta\varepsilon_{j,\alpha,\beta}^\mu + \cdots \qquad (4.28)$$

So the pions transform under the two transformations as:

$$\delta\pi^\mu \;=\; \varepsilon_i^\mu + \varepsilon_i^\beta\pi_{,\beta}^\mu - \varepsilon_j^\mu - \pi^\beta\varepsilon_{j,\beta}^\mu - \frac{1}{2}\pi^\alpha\pi^\beta\varepsilon_{j,\alpha,\beta}^\mu - \cdots \qquad (4.29)$$

Note that in the global symmetry limit , where the ε's are constant, we have

$$\pi^\mu \rightarrow \pi^\mu + \varepsilon_i^\nu\pi_{,\nu}^\mu + \varepsilon_i - \varepsilon_j = \pi(x + \varepsilon_i) + \varepsilon_i - \varepsilon_j \qquad (4.30)$$

This is just a translation in x_i by ε_i, together with a shift symmetry. Note that in this global limit the symmetry is Abelian. The shift symmetry is the analog of the shift symmetry acting on scalar Goldstone bosons that keeps them exactly massless.

The transformations of the the pions (4.29) are non-linear and messy. But $G_{\mu\nu}$ has simple transformation properties which come from the simple transformations of Y. By plugging (4.24) and (4.29) into (4.22) we find that:

$$\delta G_{\mu\nu} = \varepsilon_i^\lambda G_{\mu\nu,\lambda} + \varepsilon_{i,\mu}^\lambda G_{\lambda\nu} + \varepsilon_{i,\nu}^\lambda G_{\mu\lambda} \qquad (4.31)$$

$G_{\mu\nu}$ transforms like a tensor under GC_i and is invariant under GC_j. This is exactly what we wanted. We can now see that the Lagrangian (4.19) possesses two separate general coordinate symmetries: if we make the transformations (4.23) and (4.31) with independent ε_i and ε_j the Lagrangian is unchanged.

We can also see that is easy to make $G_{\mu\nu}$ invariant under both GC_i and GC_j. We just redefine the non-linear transformation laws of π to not depend on ε_i by setting $\varepsilon_i = 0$ in (4.29). Then, without changing the expansion of G in terms of π and g_j, (4.22), G will be invariant for any ε_i and ε_j. This is a handy way to add general coordinate invariance to *any* Lagrangian, even one without a symmetry under the diagonal subgroup. Again, this symmetry is a complete fake, but it can still be useful.
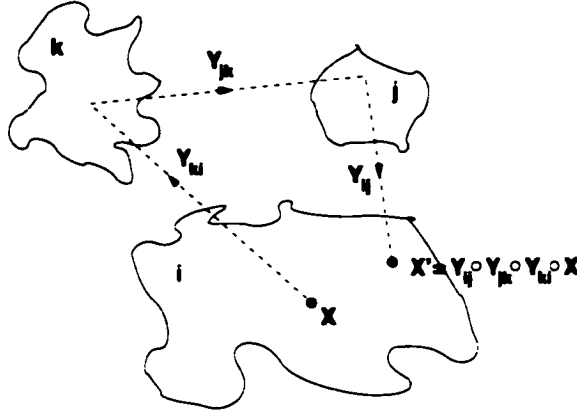
Figure 4.2: Making a plaquette from $Y_{ji} \circ Y_{ik} \circ Y_{kj}$.

4.3.3 Plaquettes

In the gauge theory case, higher-dimensional theory spaces can be constructed with the appropriate mesh of sites and links. In the case of more than one extra dimension, not all of the link fields can be gauged away. The classical spectrum includes the usual KK tower of spin one particles, but also a large number of massless scalars corresponding to the uneaten link fields. It is possible to give mass to these extra massless fields through the addition of "plaquette" interactions. For example, going in a closed circle of links from i to j to k back to i, we can form the plaquette $(U_{ij}U_{jk}U_{ki})$, which is conjugated under the action of g_i. Taking the trace the yields an invariant potential that can be added to the Lagrangian.

We can construct plaquettes in the gravitational case as well. Suppose we have three sites i, j, k with links Y_{ij}, Y_{jk}, Y_{ki}. We can form the functional product

$$\Psi = Y_{ij} \circ Y_{jk} \circ Y_{ki} \tag{4.32}$$

which transforms as

$$\Psi \to f_i^{-1} \circ \Psi \circ f_i \tag{4.33}$$

How can we build an invariant out of this quantity that can be added to the action? What is the analogue of the trace in the gauge theory case? Note that Ψ maps a point X on site i to a different point $X' = \Psi(X)$ (see figure 4.2). Then the geodesic distance between X and X' is coordinate independent and transforms like a scalar under GC_i. The simplest

invariant we can build which is analytic in the fields is then:

$$\int d^4x \sqrt{g_i}\, l^2(x, \Psi(x); g_i)$$

(4.34)

where $l[x, y; g] = \int_x^y ds$ is the geodesic distance between the points x, y with the metric g. (We are imagining that all the π's are perturbatively close to zero, so that there is no ambiguity in which is the shortest geodesic between x and $\Psi(x)$.) This is the analog of the simplest single trace plaquette operator in the gauge theory. Expanding around flat space, this operator gives a mass to the uneaten combination of Goldstones $\pi_{ij}^\alpha + \pi_{jk}^\alpha + \pi_{ki}^\alpha$ and provides additional non-linear interactions needed to preserve the general covariances. This is all in complete analogy with the gauge theory case.

4.4 Massive gravitons

We will now show how the effective field theory formalism of the previous section makes studying a massive graviton embarrassingly easy. These fields have been sporadically studied for many years, and are known to have several peculiar properties. Almost all of the studies have been based on deforming GR by the addition of the (already somewhat peculiar) Fierz-Pauli mass term, $(h_{\mu\mu})^2 - h_{\mu\nu}h_{\mu\nu}$, where $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$ is the metric linearized around flat space. This specific linear combination is needed for a unitary propagator [47, 109]. Then there is the famous van Dam-Veltman-Zakharov discontinuity [108, 116] in the graviton propagator as $m_g \to 0$, which seems to indicate that the an arbitrarily small mass graviton yields different predictions than Einstein's theory. Recently, it has been observed [66, 83, 84, 70, 69, 71], that this discontinuity disappears in Anti-de-Sitter space and in de Sitter space, though there is a ghost instead in the de Sitter theory. Finally, there is the observation of Vainshtein [107], that the discontinuity may not be relevant for physical sources, because the linearized approximation to gravity outside a source of mass M breaks down at a much larger distance than the gravitational radius $R_g = l_{Pl}^2 M$; at a distance $(m_g^{-4} R_g)^{1/5}$.

With our Goldstone boson description, we will understand these peculiarities trivially,

and see that they are all associated with a single underlying cause. The scalar longitudinal component of the graviton becomes strongly coupled at a far lower energy scale than we may have expected by analogy with the familiar gauge theory case. For a massive spin one field, the cutoff is $\sim m_A/g$. This would translate to a cutoff $\sim \sqrt{m_g M_{Pl}}$ in the gravity case. However, we will find that while there is a sensible effective theory for interacting massive gravitons, the cutoff is parametrically far lower than this. Beginning with the Fierz-Pauli Lagrangian, it is $\sim (m_g^4 M_{Pl})^{1/5}$, while a slightly more clever starting point can push the cutoff higher to $\sim (m_g^2 M_{Pl})^{1/3}$.

4.4.1 Two site model

Following the gauge example, it is straightforward to isolate a single massive graviton in an arbitrary background from a two site model. We start with an action is of the form

$$S = S_{grav} + S_{mass} \tag{4.35}$$

where

$$S_{grav} = \int d^4x \sqrt{-g}\left(M_{Pl}^2 R[g] + \cdots\right) + \int d^4x_0 \sqrt{-g_0}\left(-\Lambda_0 + M_0^2 R[g_0] + \cdots\right) \tag{4.36}$$

represents the action for the gravitons on the sites. For simplicity we have not put in a cosmological constant term on the first site. And

$$S_{mass} = \int d^4x \sqrt{-g}\, g^{\mu\nu} g^{\alpha\beta}\left(a H_{\mu\nu} H_{\alpha\beta} + b H_{\mu\alpha} H_{\nu\beta}\right) + \cdots \tag{4.37}$$

denotes the "hopping" action that will give one combination of gravitons a mass. Here,

$$H_{\mu\nu}(x) \equiv g_{\mu\nu}(x) - \partial_\mu Y^\alpha(x)\partial_\nu Y^\beta(x) g_{0\alpha\beta}(Y(x)) \tag{4.38}$$

We can go to a unitary gauge where $Y = $ id and there is one manifest general coordinate invariance under which both g and g_0 transform as tensors. The spectrum contains one massless graviton and one massive graviton. In the limit where we send $M_0 \to \infty$, this massless graviton is all g_0 and becomes non-dynamical, and so we are left with a theory of a single massive graviton described by g, in a non-dynamical background geometry g_0.

(W. Siegel has informed us that this way of introducing general coordinate invariance for a massive graviton was considered in [94]. For other early work see [55].) The Goldstone formulation will be invaluable in elucidating the interactions of the longitudinal components of the massive gravitons and determining the structure of the effective field theory.

4.4.2 Linearized analysis

Let us begin by analyzing our action to quadratic order in the fields. To wit. $H_{\mu\nu}$ is expanded as in (4.22) with $g_j = g_0$:

$$H_{\mu\nu} = h_{\mu\nu} + g_{0\mu\alpha}\nabla^0_\nu \pi^\alpha + g_{0\nu\alpha}\nabla^0_\mu \pi^\alpha + \cdots \qquad (4.39)$$

where

$$h_{\mu\nu} \equiv g_{\mu\nu} - g_{0\mu\nu} \qquad (4.40)$$

and ∇^0_μ is the covariant derivative with the background metric g_0.

Our Goldstones are a vector field. which has 3 polarizations. These are eaten by the massless graviton. which has 2 polarizations, to produce a massive graviton with a total of 5 physical polarizations. We can decompose the π^α into the transverse spin one and scalar mode by expressing

$$\pi^\alpha(x) = g_0^{\alpha\beta}(A_\beta + \partial_\beta \phi) \qquad (4.41)$$

This allows us to introduce a fake $U(1)$ gauge symmetry under which

$$A_\beta \to A_\beta + \partial_\beta \Lambda, \ \phi \to \phi - \Lambda \qquad (4.42)$$

The graviton mass term will turn into the kinetic term for these Goldstones. Of course. as usual, the new fake general covariance and $U(1)$ symmetries must be gauge fixed by the addition of suitable gauge fixing terms for $h_{\mu\nu}$ and A_α, for instance fixing to Feynman-like gauges for both. The precise form of this gauge fixing will not be relevant for our discussion. Note that as defined A_α has mass dimension -1 and ϕ has mass dimension -2.

By the usual logic leading to the equivalence theorem for gauge theories. we expect that the physics of A and ϕ is that of the vector longitudinal (g^{vL}) and scalar longitudinal (g^{sL})

polarizations of the massive graviton field, at energies much higher than the mass of the graviton:

$$\tag{4.43}$$

Let us now consider the case where the background $g_{0\mu\nu} = \eta_{\mu\nu}$ is flat. We can immediately notice a peculiarity that will be at the heart of the difference the gravity and gauge theory cases. In the global symmetry limit where we send $M_{\mathrm{Pl}} \to \infty$ (equivalently set $h_{\mu\nu} \to 0$), there is a shift symmetry on the π^α that means that π^α only appears with a derivative acting on it. This implies that ϕ only ever appears with two derivatives, so ϕ^2 must have four derivatives. Therefore, ϕ cannot have a normal kinetic term in this limit. Indeed, for general coefficients a and b in S_m, ϕ will have a four-derivative kinetic term as

$$S_{mass} \supset \int d^4x \, 4a\phi_{,\mu,\nu}\phi_{,\mu,\nu} + 4b\Box\phi\Box\phi \tag{4.44}$$

$$= \int d^4x \, 4(a+b)\Box\phi\Box\phi \tag{4.45}$$

where to get to the second line we have integrated by parts. In order to eliminate this pathological kinetic term, which implies ghosts and violations of unitarity, we must choose $a + b = 0$. This is precisely the Fierz-Pauli mass term in the unitary gauge,

$$\mathcal{L}_{FP} = f^4 \left(h_{\mu\nu}h_{\mu\nu} - h^2 \right) \tag{4.46}$$

corresponding to a graviton mass

$$m_g^2 = \frac{f^4}{M_{\mathrm{Pl}}^2}. \tag{4.47}$$

We will assume throughout that the graviton mass is parametrically much smaller than the Planck scale, $m_g \ll M_{Pl}$; this is essential for any sensible effective theory.

Note that with the Fierz-Pauli choice $a = -b$, ϕ has *no* kinetic term in the decoupling limit. On the other hand, A_α has a perfectly healthy kinetic term $f^4(A_{\mu,\nu} - A_{\nu,\mu})^2$ in this limit. Now, ϕ *does* acquire a normal two-derivative kinetic term, but only via mixing with the graviton $h_{\mu\nu}$. Indeed, in the expansion of S_{mass}, there is is a mixing term

$$f^4 \left(h_{\mu\nu}\phi_{,\mu,\nu} - h\Box\phi \right) \tag{4.48}$$

It is useful to express this term in a more familiar form. After integrating by parts, this kinetic mixing term is $f^4\phi R_{\text{lin}}$ where $R_{\text{lin}} = h_{\mu\nu,\mu,\nu} - \Box h$ is the Ricci scalar at linear order in h. Thus, at quadratic order in the fields the kinetic Lagrangian is the same as

$$\sqrt{g}M_{Pl}^2(1 + m_g^2\phi)R \tag{4.49}$$

We can eliminate the kinetic mixing between ϕ and h by by a Weyl rescaling of the metric, which is well-known to generate a kinetic term for ϕ with the correct sign. At quadratic order this amounts to is redefining

$$h_{\mu\nu} = \hat{h}_{\mu\nu} - \eta_{\mu\nu}(m_g^2\phi) \tag{4.50}$$

and the induced kinetic term for ϕ is

$$M_{Pl}^2 m_g^4(\partial\phi)^2 = \frac{f^8}{M_{Pl}^2}(\partial\phi)^2 \tag{4.51}$$

It will be also be convenient to add a gauge-fixing term directly for $\hat{h}$. Then at quadratic order, we have usual kinetic and mass terms for $\hat{h}_{\mu\nu}$, and a kinetic term for ϕ, with no kinetic mixing between them.

So ϕ does acquire a normal kinetic term, but this term disappears in the limit where f is held fixed and M_{Pl} is sent to infinity. This is in stark contrast with the gauge theory case, where the Goldstone scalar kinetic term is $\sim f^2(\partial\pi)^2$ and survives as $g \to 0$.

Now, we certainly do not expect our choice $a = -b$ to be exactly radiatively stable at quantum level, so let us get a better idea of what would happen if $a \neq -b$. The ϕ kinetic term would have the structure in momentum space

$$\frac{f^8}{M_{Pl}^2}p^2\phi^2 + (a + b)p^4\phi^2 \tag{4.52}$$

and this would lead to ghosts or tachyons at a momentum scale $p^2 \sim \frac{f^8}{(a+b)M_{Pl}^2} \sim m_g^2\frac{f^4}{a+b}$. So we see why we need to have $(a + b) \ll f^4$, because otherwise our effective theory would break down right around the mass of the particle m_g and would be completely useless. This is the analog of the reason why in gauge theory we do not include a $\frac{1}{g^2}(\partial A)^2$ kinetic term. As we saw, such a term is generated at quantum level, but with a small enough coefficient

so that its harmful effects are deferred to the cutoff. We will see that exactly the same thing happens in our case; with a suitable cutoff. a small $(a + b)$ is generated (along with a whole slew of other terms). but with small enough sizes to allow the effective theory to make sense to energies parametrically above m_g.

Notice that in order to go to canonical normalization for ϕ. we define

$$\phi = \frac{M_{\text{Pl}}}{f^4}\phi^c = \frac{1}{m_g^2 M_{\text{Pl}}}\phi^c \tag{4.53}$$

The M_{Pl} in the numerator implies that the interactions of ϕ will become strongly coupled at an energy far beneath f. again in contrast with the gauge theory case. On the other hand, the A_α kinetic term is proportional to $\sim f^4$. and the canonically normalized field is

$$A_\alpha = \frac{1}{f^2}A_\alpha^c = \frac{1}{m_g M_{\text{Pl}}}A_\alpha^c \tag{4.54}$$

These conclusions are changed in a general curved background, which introduces another scale into the problem. In flat space, with the choice $a = -b = f^4$ there is no kinetic term for ϕ without mixing through h: the contributions to the ϕ kinetic term proportional to a and b cancel exactly after integrating by parts. In a general curved background. we have instead at quadratic order

$$f^4 \int d^4x\sqrt{-g_0} \left(\nabla^0_\mu\nabla^0_\nu\phi\nabla^{0\mu}\nabla^{0\nu}\phi - (\nabla^0)^2\phi(\nabla^0)^2\phi\right) \tag{4.55}$$

$$= f^4 \int d^4x\sqrt{-g_0}\,\phi_{,\mu}\left[\nabla^{0\mu}.\nabla^{0\nu}\right]\phi_{,\nu} \tag{4.56}$$

after integration by parts. Since the commutator of the covariant derivatives is proportional to the Riemann tensor and is non-vanishing in a curved background, there is an induced kinetic term for ϕ (and a corresponding mass term for A_α) proportional to the background curvature. For a maximally symmetric space like AdS the $[\nabla^{0\mu}.\nabla^{0\mu}]$ term can be replaced by $\frac{1}{L^2}g^{0\mu\nu}$ in (4.56), where L is AdS radius of curvature. Therefore this contribution to the ϕ kinetic term is

$$\frac{f^4}{L^2}(\partial\phi)^2 = \frac{m_g^2}{L^2}M_{\text{Pl}}^2(\partial\phi)^2 \tag{4.57}$$

This should be compared with the kinetic term coming from mixing with h, $\sim m_g^4 M_{\text{Pl}}^2(\partial\phi)^2$. We see that for $1/L \gg m_g$, the new contribution to the kinetic term dominates. It is easy

to check that it has the good sign in AdS space and the bad sign leading to ghosts in dS space. Note that for $1/L \gg m_g$, . the kinetic term survives as M_{Pl} is taken to infinity with f fixed, and in this limit the canonically normalized ϕ field is

$$\phi = \frac{L}{f^2}\phi^c = \frac{1}{m_g L M_{\mathrm{Pl}}}\phi^c \quad \left(\text{for}\frac{1}{L} \gg m_g\right) \tag{4.58}$$

4.4.3 The vDVZ discontinuity in general backgrounds

We are now in a position to understand easily the presence/absence of the vDVZ discontinuity in general space-times. Let us first work in flat space. The coupling of the graviton $h_{\mu\nu}$ to the energy momentum tensor is

$$T^\mu_\nu h^\nu_\mu = T^\mu_\nu(\hat{h}^\nu_\mu + m^2_g\delta^\nu_\mu\phi) = T^\mu_\nu\hat{h}^\nu_\mu + \frac{1}{M_{\mathrm{Pl}}}T\phi^c \tag{4.59}$$

where in the last equality we have gone to canonical normalization for ϕ. We see that independent of the graviton mass m_g, there is a coupling of the trace of the energy momentum tensor to ϕ with gravitational strength. Thus, for $m_g \to 0$ matter and radiation would couple with different relative strengths than they would if ϕ were absent, that is if m_g were strictly massless. This is exactly the vDVZ discontinuity, and we have traced its origin to the strongly coupled nature of the scalar ϕ. Despite the fact that $h_{\mu\nu}$ only has a small admixture $m^2_g\eta_{\mu\nu}\phi$ of ϕ which appears to go to zero as $m_g \to 0$, because the ϕ kinetic term vanishes in the same limit, when we go to canonical normalization the scalar couples with gravitational strength independent of m_g.

Then we can see immediately that this discontinuity vanishes in an AdS background. There is still a coupling $m^2_g\phi T$, however in the limit where $1/L \gg m_g$, in going to canonical normalization $\phi = L/(m_g M_{\mathrm{Pl}})\phi^c$ we obtain the coupling

$$\frac{m_g L}{M_{\mathrm{Pl}}}T\phi^c \tag{4.60}$$

which *does* vanishes as $m_g \to 0$. There is therefore no new gravitational strength force and no vDVZ discontinuity in this limit.

4.4.4 Strong coupling scale and power-counting in the effective theory

Let us now return to flat space but consider the fully interacting theory. We are in particular interested in the interactions of the longitudinal components of the graviton at energies far above m_g, which are the interactions of the ϕ and A_α. These are the strongest interactions and signal when the theory breaks down.

In flat space, our expression for $H_{\mu\nu}$ in terms of the Goldstone bosons becomes

$$H_{\mu\nu} = h_{\mu\nu} + \pi_{\mu,\nu} + \pi_{\nu,\mu} + \pi_{\alpha,\mu}\pi_{\alpha,\nu} \qquad (4.61)$$

where $\pi_\mu = \eta_{\mu\nu}\pi^\nu$. Since we are not interested in the usual helicity two polarization, we can set $\hat{h}_{\mu\nu} = 0$. Furthermore, since $\hat{h}_{\mu\nu}$ and $h_{\mu\nu}$ only differ by an amount that is $\sim 1/M_{\rm Pl}\phi^c\eta_{\mu\nu}$, for the amplitudes of interest which will be getting large far beneath $M_{\rm Pl}$ we can simply set $h_{\mu\nu} = 0$ in all interaction terms. Thus, to obtain the interactions for our Goldstones, it suffices to replace everywhere

$$H_{\mu\nu} \to \pi_{\mu,\nu} + \pi_{\nu,\mu} + \pi_{\alpha,\mu}\pi_{\alpha,\nu} \qquad (4.62)$$

The Fierz-Pauli mass term can then be seen to contain cubic and quartic interactions for ϕ and A. The only interactions that can become anomalously large involve ϕ, and are schematically

$$f^4\left[(\partial^2\phi)^3 + (\partial^2\phi)^4 + \partial^2\phi\partial A\partial A\right] \qquad (4.63)$$

(Note that it is impossible to have a term with a single ∂A and $\partial^2\phi$'s, because by the $U(1)$ gauge invariance (4.42), the ∂A piece would have to involve $F_{\mu\nu}$ which is antisymmetric, and vanishes when contracted with anything made out of $\phi_{,\mu,\nu}$ which is symmetric in μ,ν.)

Going to canonical normalization, these interactions become

$$\frac{1}{m_g^4 M_{\rm Pl}}(\partial^2\phi^c)^3 + \frac{1}{m_g^6 M_{\rm Pl}^2}(\partial^2\phi^c)^4 + \frac{1}{m_g^2 M_{\rm Pl}}\partial^2\phi^c\partial A^c\partial A^c \qquad (4.64)$$

The cubic scalar interaction is the strongest coupling and becomes large at an energy scale

$$\Lambda_5 \sim (m_g^4 M_{\rm Pl})^{1/5} \qquad (4.65)$$

Correspondingly, the amplitude $\mathcal{A}(\phi\phi \to \phi\phi)$ from ϕ exchange grows as $\sim E^{10}/\Lambda_5^{10}$ and gets strongly coupled at Λ_5. This means that the scattering amplitude for the scalar longitudinal polarization of the massive graviton gets strongly coupled at the scale Λ_5:

$$ \text{(diagram)} + \text{(diagram)} = \text{(diagram)} + \text{ smaller amplitudes} \qquad (4.66) $$

This could have actually been guessed directly in unitary gauge. The polarization vector for the scalar longitudinal polarization of the graviton at high energies is $\epsilon_{\mu\nu} \sim k_\mu k_\nu/m_g^2 \sim (E^2/m_g^2)$. Naively, each one of the graviton diagrams in (4.66) grow as

$$ \left(\frac{E^2}{m_g^2}\right)^4 \times \frac{E^2}{M_{\rm Pl}^2} \sim \frac{E^{10}}{\Lambda_5^{10}} \qquad (4.67) $$

In the gauge theory case, there is a cancellation of the leading term between the two diagrams; one may have expected a similar cancellation here. However, without the need to perform this very hairy perturbative massive gravity computation, our Goldstone description tells us that no such cancellation occurs, and the amplitude gets strongly coupled at Λ_5. Perhaps this is not surprising, since our starting point, the Fierz-Pauli Lagrangian, is somewhat arbitrary. That is, there is nothing special about terms quadratic in h given that h is dimensionless. Higher order interactions may help cancel the strongest divergences in (4.66). We will come back to this in section 4.4.5.

Let us proceed to determine the structure of the operators generated at quantum level in this effective theory, taken to have a cutoff Λ_5. We must include all operators consistent with the symmetries, suppressed by the cutoff Λ_5. The shift symmetry guarantees that the leading operators are of the form

$$ \frac{\partial^q (\partial^2 \phi^c)^p}{\Lambda_5^{3p+q-4}} \qquad (4.68) $$

In order to find what operators these correspond to in unitary gauge, we can go back to the original normalization for $\phi^c = m_g^2 M_{\rm Pl}\phi$ and recall that $\phi_{,\mu,\nu}$ always comes from an $h_{\mu\nu}$. Thus, in unitary gauge, we have operators of the form

$$ c_{p,q}\partial^q h^p \qquad (4.69) $$

where the coefficients $c_{p,q}$ have a natural size

$$c_{p,q} \sim \Lambda_5^{-3p-q+4} M_{\text{Pl}}^p m_g^{2p} = \left(m_g^{16-4q-2p} M_{\text{Pl}}^{2p-q+4} \right)^{1/5} \tag{4.70}$$

Note that, for example, the term with $p = 2, q = 0$ is a general mass term for h, not necessarily of the Fierz-Pauli form. However, its coefficient, $c_{2,0} = (m_g^{12/5} M_{\text{Pl}}^{8/5})$, is parametrically much smaller than the Fierz-Pauli coefficient $f^4 = m_g^2 M_{\text{Pl}}^2$ (4.46), and the unitarity violation/tachyons are postponed to energies above the cutoff Λ_5.

We can summarize by saying that there is a natural effective theory with an action

$$\int d^4x \sqrt{-g} (M_{\text{Pl}}^2 R + \cdots) + m_g^2 M_{\text{Pl}}^2 (h_{\mu\nu}^2 - h^2) + \sum_{p,q} c_{p,q} \partial^q h^p \tag{4.71}$$

with a cutoff $\Lambda_5 = (m_g^4 M_{\text{Pl}})^{1/5}$. The "pollution" from all the higher order terms do not give rise to any pathologies until above Λ_5. Needless to say, it would have been hard to guess the structure of this effective theory directly from a unitary gauge analysis.

4.4.5 Adding interactions to raise the cutoff

We can easily find another natural effective theory where the cutoff is parametrically higher than Λ_5. By adding higher-order terms of the form $f^4(h^3 + h^4 + \cdots)$, we can remove all the ϕ self-couplings from the action. There is no unique procedure, but one way is as follows. Since $H_{\mu\nu} \sim \phi_{,\mu,\nu} + O((\partial^2 \phi)^2)$, at any order we can cancel all terms of the form $(\partial^2 \phi)^n$ by appropriately choosing the coefficient of H^n terms. These in turn only generate higher order ϕ self-interactions, which are canceled at the next step. Having eliminated all the ϕ self-interactions with terms of the form $f^4 H^n$, we are left with interactions of the form

$$f^4 (\partial A)^p (\partial^2 \phi)^q = \frac{1}{m_g^{p+2q-2} M_{\text{Pl}}^{p+q-2}} (\partial A^c)^p (\partial^2 \phi^c)^q \tag{4.72}$$

for $p > 1$. These become strongly coupled at a scale

$$\left(m_g^{p+2q-2} M_{\text{Pl}}^{p+q-2} \right)^{\frac{1}{3q+2p-4}} \tag{4.73}$$

It is easy to see that the lowest this scale can ever be is $\Lambda_3 = (m_g^2 M_{\text{Pl}})^{1/3}$, which is achieved for $p = 2, q = 1$, and also asymptotically as $q \to \infty$. Therefore the cutoff of

this effective theory is Λ_3. Note that with this choice, the leading contribution to the $\phi\phi \to \phi\phi$ amplitude is absent, which means that there is a partial cancellation between the two unitary gauge diagrams on the left side of (4.66). The largest amplitude is that of AA scattering through ϕ exchange. where the amplitude again grows as naively expected in unitary gauge $\sim E^6/(m_g^4 M_{Pl}^2)$, becoming strongly coupled at $\sim \Lambda_3$. It is easy to find that the natural size for operators of the form $c_{p,q}\partial^q h^p$ in unitary gauge is now

$$c_{p,q} \sim \Lambda_3^{4-q} = \left(m_g^2 M_{Pl}\right)^{\frac{4-q}{3}} \tag{4.74}$$

Again, the choice of coefficients needed to eliminate all the ϕ self-interactions is technically natural; since $\Lambda_3 \ll f = \sqrt{m_g M_{Pl}}$. the pollution from the operators not of the special form is small and pathologies are postponed to the cutoff.

This makes it tempting to try to push the cutoff higher by adding other interactions. We will address this question in detail elsewhere.

4.4.6 Breakdown of the effective theory around heavy sources

We have seen that our effective field theory breaks down at high energy scales $\Lambda \sim \Lambda_5$ or Λ_3, and there are infinitely many higher dimension operators suppressed by Λ that encode our ignorance of the short-distance UV completion of these theories. In high-energy scattering experiments. these effects only become important near the cutoff scale. But quite generally, effective theories can also break down at large distance scales in the presence of large background fields that make the higher dimension operators important. For instance. the Euler-Heisenberg Lagrangian which describes electrodynamics at energies beneath the electron mass contains higher-dimension operators of the form $\sim F^4/m^4, F^6/m^8 \cdots$. The effective theory certainly breaks down at short distances of order m^{-1}. but it can also break down in backgrounds with spatially homogeneous but large electric fields. where F/m^2 becomes ~ 1. Schwinger pair production becomes important and the structure of the short-distance physics becomes relevant even at large distances. Similarly. in our case there are higher-dimension operators which become important when $\partial^2 \phi^c/\Lambda^3$ becomes ~ 1, and this

quantity can indeed become large in the presence of heavy sources with masses M much larger than M_{Pl}.

Consider for instance the potential field for ϕ^c. Since ϕ^c couples with gravitational strength, it will affect the motion of test particles around a heavy source. The potential set up for ϕ^c by a source of mass M can be diagrammatically represented as

$$(4.75)$$

Here the blobs denote the heavy external source. The first diagram represents the potential set up for ϕ^c at linearized level, which will be a good approximation at sufficiently large distances from the source. The coupling to the source is proportional to M/M_{Pl} so

$$\phi^{c(1)} \sim \frac{M}{M_{Pl}} \frac{1}{r} \qquad (4.76)$$

The remaining diagrams can be easily estimated. Each vertex gives us a factor of (M/M_{Pl}), while from n point vertices of the form $\partial^q (\partial^2 \phi_c)^n / \Lambda^{3n+q-4}$ there is a factor of $1/\Lambda^{3n+q-4}$. The remaining $1/r$ factors are fixed by dimensional analysis, and we find the contribution to ϕ_c is

$$\phi_c^{(n,q)} \sim (\frac{M}{M_{Pl}})^{n-1} \frac{1}{\Lambda^{3n+q-4}} \frac{1}{r^{3n+q-3}} \qquad (4.77)$$

The distance r_n at which the n'th order contribution to ϕ_c becomes comparable to the lowest order contribution is then

$$r_{n,q} \sim \Lambda^{-1} (\frac{M}{M_{Pl}})^{\frac{n-2}{3n+q-4}} \qquad (4.78)$$

Note interestingly that this distance *increases* with n, and asymptotes to

$$r_* \sim \Lambda^{-1} (\frac{M}{M_{Pl}})^{1/3} \qquad (4.79)$$

This distance is precisely where $\partial^2 \phi^{c(1)}/\Lambda^3$ becomes ~ 1. In the action, terms with the minimal number of derivatives $q = 0$, but with any n, all become important at the distance r_*.

This tells us two things. First, the effective theory around heavy sources breaks down at parametrically much larger distances than the short distance cutoff Λ^{-1}, by a factor of $(M/M_{Pl})^{1/3}$, since infinitely many higher dimension operators become important at this scale. Second, there is no range of distances for which the linear approximation breaks down but non-linear effects can be reliably computed in the effective theory, i.e. for which only a finite number of the higher dimension operators become important. This is because it is the *highest* dimension operators that contribute to the onset of non-linear effects at large distances. We therefore directly transition from the linear regime to one where the effective theory breaks down. This is in contrast with non-linear effects in Einstein Gravity. There, at a distance of order the gravitational radius, the linearized approximation breaks down, and all the operators with two derivatives and any number of h's become important. However, general covariance dictates that all these higher dimension operators are packaged together into the Ricci scalar R, and therefore all their coefficients are known. We can therefore trust the non-linear gravity solution to much smaller distances, and the effective theory only breaks down when the curvatures become Planckian. In our case, we do not have any symmetry principle analogous to general covariance to determine the coefficients of all operators with $q = 0$ for any n, and therefore without a UV completion we are unable to compute non-linear effects around heavy sources consistently within the effective theory.

It is interesting to compare our conclusions here with the observations of Vainshtein [107] that foreshadowed some of these results. Working with the Fierz-Pauli theory, purely at the classical level, Vainshtein found that the linear approximation for gravity breaks down at a macroscopic distance $r_V \sim (G_N M m_g^{-4})^{1/5}$ from the source. He further argued that the full non-linear solution would have a continuous behavior as m_g is taken to zero. We can understand the origin of the Vainshtein radius trivially. Recall that just with the Fierz-Pauli term, the strongest interaction was the triple-scalar interaction suppressed by Λ_5. Therefore the contribution $n = 3, q = 0$ in our analysis above is the largest one, and the corresponding radius $r_{3,0}$, where this non-linear correction becomes comparable to the lowest-order term, is precisely r_V. However, from the point of view of the effective theory,

there is no reason to only keep the Fierz-Pauli terms. and in fact we must include all other operators with their natural sizes. Doing this we have found even larger non-linear effects. We have seen that the entire effective theory breaks down at the distance r_*. so that it is impossible to make any reliable predictions for gravitational strength forces at distances smaller than r_* without specifying the UV completion of the theory. And we certainly do not have any reason to expect a smooth limit as m_g is taken to zero. since in this limit the distance scale at which the effective theory breaks down goes to infinity.

If we wish to consider the possibility that our four-dimensional graviton has a small mass. then its Compton wavelength should be on order of the size of the universe: $m_g \sim 10^{28}$ cm^{-1}. Then. we find that $\Lambda_5^{-1} \sim 10^{13}$ cm. Even if we modify the theory to raise the cutoff to Λ_3, we would only have $\Lambda_3^{-1} \sim 10^7$ cm. The effective theory breaks down at even larger distances around heavy sources. Both of these scales are far larger than ~ 1 mm. where gravitational effects have been measured. Therefore our effective theory for a massive 4D graviton breaks down at distances larger than the scale we have measured gravity. and we cannot say in any controlled way that this theory is consistent with experiment. In order to avoid conflict with experiment, the short distance cutoff must at least be pushed to $\sim$ mm. which is around $\sqrt{m_g M_{Pl}}$. We will discuss this possibility in more detail elsewhere.

4.5 Summary, Discussion and Outlook

We have shown how to understand massive gravitons within the language of effective field theory. We are now able to write down interacting gravitational Lagrangians in a theory space with multiple copies of general coordinate invariance. The key is to introduce link fields which transform non-linearly under various transformations. In unitary gauge, these links are eaten to make the gravitons massive. Our generally covariant formalism allows us to study the largest interaction in the theory, involving the longitudinal components of the massive gravitons, in a simple way.

As an illustration. we have applied this formalism to study a single graviton of mass m_g. We find that there is a consistent effective theory with a cutoff Λ which can be

taken parametrically higher than m_g. It can be taken as high as $\Lambda \sim (M_{Pl}m_g^4)^{\frac{1}{5}}$ for the simplest case based on the Fierz-Pauli theory, and as high as $\Lambda \sim (M_{Pl}m_g^2)^{\frac{1}{3}}$ if we add additional terms beyond the Fierz-Pauli structure. We have understood a number of strange features of massive gravitons in a transparent way, and seen that they are all consequences of the peculiar behavior of the scalar longitudinal component of the graviton, ϕ. That the mass term must have Fierz-Pauli form to guarantee unitarity follows immediately from eliminating the pathological large four-derivative kinetic term for ϕ. Having done this, around flat space ϕ only acquires a kinetic term by mixing with $h_{\mu\nu}$. This is the origin of the vDVZ discontinuity. However, around curved backgrounds, such as AdS space, ϕ does pick up a normal kinetic term proportional to the background curvature even without mixing, and therefore the vDVZ discontinuity is absent. We have shown how to include all terms beyond the Fierz-Pauli Lagrangian with their natural sizes in the effective theory, and in particular observed that the Fierz-Pauli form of the mass term is radiatively stable. We also saw that around sources of mass M much larger than M_{Pl}, the effective theory breaks down at much larger distances than the short-distance cutoff scale, parametrically at a radius $\sim (M/M_{Pl})^{1/3}\Lambda^{-1}$.

Of course, the purpose of the effective field theory formalism is not just to understand a single massive graviton. Now that we understand the dynamics of theories with multiple interacting gravitons, we can construct large classes of models with gravity in theory space. As a trivial example, we can consider a theory space version of the first Randall-Sundrum model [89]. We have two sites, one with TeV scale gravity and the standard model and the other with Planck scale gravity. A simple link field will give a TeV mass to one combination of gravitons. At low energies there is a massless graviton with ordinary Planck scale couplings, together with a massive graviton with 1/TeV couplings to the Standard Model fields. Note that in our set-up there is no need to introduce and stabilize a radion. It would be interesting to extend the theory to more sites in a way that would dynamically generate the large hierarchy of scales in a natural way. Nevertheless, the two-site model has a low quantum gravity cutoff of $\sim$ TeV which cuts off the Higgs mass quadratic divergence.

More generally, we can construct models which involve gravitationally sequestered sectors weakly coupled to the standard model. It should also be straightforward to extend our methods to understanding supergravity in theory space. An obvious application of these ideas would be the communication of supersymmetry breaking between sites. For instance we could consider the theory space version of anomaly mediation. It will also be interesting to explore cosmological issues in few site models. In addition to simple constructions with a few sites, we can also consider building gravitational dimensions with many sites, which leads to some fascinating physics that will be discussed in detail elsewhere.

The most interesting of all possible applications would be the construction of a UV complete theory of gravity. Recall that the *deconstruction* of non-renormalizable higher-dimensional gauge theories has provided them with a UV completion. The structure of the high energy theory was easily guessed at by attempting to UV complete the low-energy non-linear sigma model fields with spontaneous symmetry breaking. Furthermore, with the addition of extra ingredients, such as supersymmetry and conformal invariance, these models lead to deconstructions of non-gravitational sectors of string theory [8]. It is a tantalizing possibility that by pursuing the analogy with the gauge theory which we have begun to develop in this paper, perhaps with some additional ingredients, we may be led to UV completions of four-dimensional quantum gravity.

Acknowledgments

Chapter 5

Discrete Gravitational Extra

Dimensions

In Chapter 4, a technique was introduced for studying gravitational theories where general coordinate invariance is explicitly broken. Some such theories are concisely described by "theory spaces" [5, 61], which comprise "sites" and "links". Each of the sites has its own four-dimensional metric an associated general coordinate invariance, and the symmetry breaking is done through link fields which map between sites. For example, in the simplest case with one site and one link, the link is "eaten", leading to a single massive graviton.

We saw that the physics of massive gravitons is qualitatively different than that of massive gauge bosons, owing to the peculiar properties of the scalar longitudinal component of the graviton. Nevertheless, there is a sensible effective theory for a single massive graviton which breaks down at energies parametrically above the graviton mass. In this chapter, we study what happens when we string together many sites and links to generate what looks like a gravitational extra dimension. Such models have been considered before [35, 16, 96, 64, 65], but not at the level of a consistent effective field theory. Given the peculiarities associated with massive gravity, we should expect surprises, and we indeed encounter a number of them. This chapter follows very closely to [14].

We will begin by considering the minimal discretizations, with nearest neighbor inter-

94

actions. The discretized dimension can be taken to be either a circle or an interval:
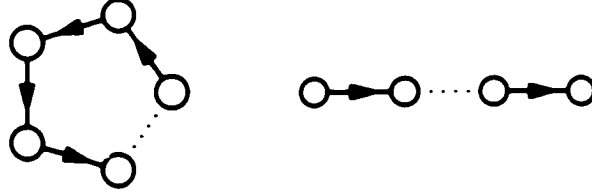
We will concentrate on the simpler case of an interval for now, and return briefly to the circle later on. Each of the N sites is endowed with a four-dimensional metric $g^J_{\mu\nu}(x)$. The action

$$S = S_{\text{site}} + S_{\text{link}} \tag{5.1}$$

contains a part

$$S_{\text{site}} = \sum_J \int d^4 x M^2 \sqrt{g^J} R[g^J] \tag{5.2}$$

which has N copies of general coordinate invariance (GC). We can see this by choosing a new dummy variable x_J for each integral in the above sum. The other part of S involves interactions between neighboring sites, and will produce mass terms of Fierz-Pauli form [47]:

$$S^U_{\text{link}} = \sum_J \int d^4 x \sqrt{g^J} M^2 m^2 (g^J_{\mu\nu} - g^{J+1}_{\mu\nu})(g^J_{\alpha J} - g^{J+1}_{\alpha J})(g^{J\mu\nu} g^{J\alpha J} - g^{J\mu\alpha} g^{J\nu J}) \tag{5.3}$$

The U in S^U_{link} stands for Unitary gauge. Indeed, we can see that a gauge is chosen because S^U_{link} breaks all but one copy of GC.

The action we have constructed is just a naive discretization of compactified 5D Einstein gravity. To see this, start with a 5D metric $G_{MN}(x, z)$, with z the compact extra dimension. Ignoring the radion and graviphoton degrees of freedom (they do not affect the following discussion) it is possible to choose a gauge where the metric takes the form

$$G_{MN} = \begin{pmatrix} g_{\mu\nu}(x, z) & 0 \\ 0 & 1 \end{pmatrix} \tag{5.4}$$

In this gauge, the 5D action $\int d^5 x M^3_{5D} \sqrt{G} R^5[G]$ is

$$S = \int d^4 x dz \sqrt{g} M^3_{5D} \left(R_{4D}[g] + \frac{1}{4} \partial_z g_{\mu\nu} (g^{\mu\alpha} g^{\nu J} - g^{\mu\nu} g^{\rho J}) \partial_z g_{\alpha J} \right) \tag{5.5}$$

Then, a naive discretization with lattice spacing a instructs us to replace

$$\int dz \to a \sum_j , \quad \partial_z g_{\mu\nu} \to \frac{1}{a}(g^j_{\mu\nu} - g^{j+1}_{\mu\nu}) \tag{5.6}$$

which directly produces our unitary gauge action (5.3). And we can read off that the radius R, lattice spacing a, and effective 5D Planck scale M_{5D} are given by:

$$R = Na, \quad a = m^{-1}, \quad M_{5D} = M^2 m \tag{5.7}$$

Now, as explained in [12], we can restore the broken GC symmetries in (5.3) by adding link fields $Y^\mu_j(x)$ between the sites. It is useful to think of the link as a map from a point on site j with coordinate x^μ_j to a point on site $j+1$ with coordinate $Y^\mu_j(x_j)$. Under general coordinate transformations generated by $x_j \to f_j(x_j)$, the link fields transform as:

$$Y_j \to f^{-1}_{j+1} \circ Y_j \circ f_j \tag{5.8}$$

which allows us to compare fields on adjacent sites.

These link fields can be used to construct objects which transform as tensors under GC_j and are invariant under GC_{j+1} out of objects which are invariant under GC_j and tensors under GC_{j+1}:

$$G^{j+1}_{\mu\nu}(x_j) \equiv \partial_\mu Y^\alpha_j \partial_\nu Y^\beta_j g^{j+1}_{\alpha\beta}(Y_j(x_j)) \tag{5.9}$$

$G^{j+1}_{\mu\nu}(x_j)$ can be thought of as a pull-back of the metric on site $j+1$ to site j using the maps Y_j. Thus, the link action becomes:

$$S_{\text{link}} = \sum_j \int d^4x_j \sqrt{g^j} M^2 m^2 (g^j_{\mu\nu}(x_j) - G^{j+1}_{\mu\nu}(x_j))$$
$$\times (g^j_{\alpha\beta}(x_j) - G^{j+1}_{\alpha\beta}(x_j))(g^{j\mu\nu} g^{j\alpha\beta} - g^{j\mu\alpha} g^{j\nu\beta}) \tag{5.10}$$

We can now replace the dummy variables x_j by a common set of coordinates x. By construction this 4D action is explicitly invariant under N copies of GC.

It is useful to expand the link fields about the identity as

$$Y^\mu_j(x) = x^\mu + \pi^\mu_j(x) \tag{5.11}$$

The vector fields π^μ are the Goldstone bosons that are eaten in producing a collection of massive spin two fields. Indeed. in unitary gauge we set $\pi_j = 0$ ($Y_j^\mu(x) = x^\mu$) and (5.10) reduces to (5.3).

The spectum of this minimal theory space is that of standard latticizations: there is a massless 4D graviton. and tower of massive spin two fields. with a characteristic lattice spectrum $m_n = m\sin(n/N)$. For $n \ll N$ the spectrum approaches a KK tower of compactified 5D theory. At the linear level. the exchange of these modes generates the correct 5D graviton propagator up to small corrections. and so. for example. 5D Newtonian gravity is reproduced.

However. at the non-linear level peculiar new features are revealed. which can be traced to the interactions of the longitudinal modes of the massive gravitons. as in [12]. Expanding the metrics about flat space as $g^j_{\mu\nu} = \eta_{\mu\nu} + h^j_{\mu\nu}$. the hopping terms involve

$$g^j_{\mu\nu} - G^{j+1}_{\mu\nu} = h^j_{\mu\nu} - h^{j+1}_{\mu\nu} + \pi^j_{\mu,\nu} + \pi^j_{\nu,\mu} + \pi^{j\alpha}_{,\mu}\pi^j_{\alpha,\nu} + \cdots \tag{5.12}$$

where the $\cdots$ refer to terms involving more powers of the h. As in the study of a single massive graviton. it is useful to decompose the π^μ_j as

$$\pi^j_\mu = A^j_\mu + \partial_\mu\phi^j \tag{5.13}$$

The dynamics of the ϕ^j is that of the scalar longitudinal components of the gravitons. and they produce the amplitudes that grow most dangerously with energy.

Inserting (5.12) into (5.10) and going to momentum space displays the kinetic terms and interactions for all the Goldstone modes. Schematically:

$$S = \int d^4x N M^2 h_n \partial^2 h_n + N M^2 m^2 \left(\frac{n^2}{N^2}h_n^2 \right.$$
$$\left. + \frac{n}{N}h_n\partial^2\phi_n + (\partial^2\phi_n)(\partial^2\phi_m)(\partial^2\phi_{n+m}) + \cdots \right) \tag{5.14}$$

where h_0 is the massless graviton and h_n and ϕ_n are the graviton and scalar Goldstone at the nth mass level. Now. ϕ_n picks up a kinetic term of the form $M^2m^4n^2N^{-1}\phi_n\partial^2\phi_n$ from mixing with h_n and the strongest interactions come from the $\partial^6\phi^3$ vertex for the lowest

modes, just as in [12]. For instance the amplitude for ϕ_1-ϕ_1 scattering goes as:

$$\mathcal{A} = \quad \phi_1 \quad \phi_1 \quad \sim \quad \frac{E^{10}}{\Lambda^{10}} \qquad (5.15)$$

Where, expressed in terms of the low energy 4D Planck scale $M_{\text{Pl}} = \sqrt{N}M$ and the mass of the first KK mode $m_1 = m/N$:

$$\Lambda = (N m_1^4 M_{\text{Pl}})^{1/5} \qquad (5.16)$$

This is higher by a factor of $N^{1/5}$ than the scale that the theory of a single graviton of mass $m_g = m_1$ breaks down.

In terms of 5D variables:

$$\Lambda = \left(\frac{M_{5D}^3}{R^5 a^2} \right)^{1/10} \qquad (5.17)$$

Note that bizarrely, the UV scale at which the theory becomes strongly coupled depends on an IR scale, the *size* of the extra dimension!

Naturally, as we decrease the lattice spacing, Λ increases. However, for a consistent effective theory, we should require that all the states in the theory be lighter than the UV cutoff Λ. In particular, the heaviest KK mode, of mass $\sim a^{-1}$, should be lighter than Λ. This means that the lattice spacing cannot be decreased beyond a certain point, and the highest UV cutoff the theory can possibly have is

$$\Lambda_{\text{max}} \sim a_{\text{min}}^{-1} \sim \left(\frac{M_{5D}^3}{R^5} \right)^{1/8} \qquad (5.18)$$

Again, this strikingly exhibits a UV/IR connection: the highest possible UV cutoff Λ_{max} decreases as the size R of the dimension is increased in such a way that $\Lambda_{\text{max}}^8 R^5 = M_{5D}^3$ stays fixed. Note that for any radius larger than the 5D Planck length, Λ_{max} is always smaller than M_{5D}. In other words, the minimal lattice cannot look like 5D gravity at the non-linear level all the way up to the 5D Planck scale.

It is useful to understand what is going on directly in unitary gauge, where the amplitude (5.15) is that of scattering scalar longitudinal (sL) polarizations of the lightest massive gravitons. The amplitude (5.15) can be written in the instructive form:

$$\mathcal{A} \sim \frac{E^{10}}{\Lambda^{10}} \sim \frac{E^{10} R^5 a^2}{M_{5D}^3} \sim \frac{E^{10}}{M_{\text{Pl}}^2 (1/R)^8} \times \frac{a^2}{R^2} \qquad (5.19)$$

In the case of a single massive graviton of mass m_g and Planck scale M_{Pl}, there are two contributions to the amplitude for graviton scattering, one from graviton exchange and one from the direct 4-point graviton vertex
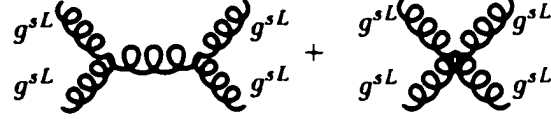
There is no cancellation between these two contributions, so the amplitude grows as $E^{10}/(m_g^8 M_{Pl}^2)$. Our scattering amplitude has exactly the same form, with $m_g \to 1/R$, however there is a suppression factor of a^2/R^2. Evidently, in the continuum theory, there is an *exact* cancellation between the two contributions, ensured by the 5D gravitational Ward identities. However, in the discretized theory, the spectrum and interactions are modified by $\sim (a/R)$ effects, and so the cancellation is imperfect, reflecting the breaking of the 5D general coordinate invariance by our discretization.

Needless to say, this behavior is dramatically different than for gauge theories. The same theory space for a non-Abelian gauge theory would become strongly coupled at an energy scale which is *always* higher than than the mass of all the modes. So the discretized theory can be made to look identical to a higher dimensional gauge theory all the way up to scale where the 5D (non-renormalizable) gauge theory would naturally break down.

Nevertheless, the minimal gravitational model does define a sensible effective field theory of gravity valid to energies parametrically above the compactification radius $1/R$. The strong interactions formally vanish in the limit of zero lattice spacing, they just do not vanish quickly enough to reproduce 5D gravity all the way up to M_{5D} in a consistent effective field theory.

Let us now study this model in position space, where the physics is more transparent. Working directly in continuum language, the action S becomes:

$$S = \int d^4x\, dz\, M_{5D}^3 \left((\partial h)^2 + (\partial_z h)^2 + \frac{1}{a} h \partial^2 \partial_z \phi + \frac{1}{a^2} (\partial^2 \phi)^3 + \cdots \right) \qquad (5.20)$$

We have done an integration by parts to bring the bilinear mixing between h and ϕ to the above form. Note also that because the extra dimension is an interval, the boundary

conditions at the ends are that $\partial_z h$ and π vanish, so that there are no Goldstone zero modes (corresponding to the fact that they are all eaten).

We can remove the kinetic mixing between h and ϕ by defining

$$h_{\mu\nu} = \hat{h}_{\mu\nu} - \eta_{\mu\nu}\psi, \text{ where } \psi = \frac{1}{a}\partial_z\phi \tag{5.21}$$

which generates a kinetic term for ϕ. We add gauge-fixing terms directly for $\hat{h}$, the precise form of which will not be important here. The ϕ action then becomes:

$$S = \int d^4x dz M_{5D}^3 \left((\partial\psi)^2 + (\partial_z\psi)^2 + \frac{1}{a^2}(\partial^2\phi)^3 + \cdots \right)$$

Already we see that this action is strange. The combination $\psi = \frac{1}{a}\partial_z\phi$ has a normal kinetic term. ψ also couples directly to the trace of the energy momentum tensor in the way required to produce the correct tensor structure for the propagator in 5D. So, at the linear level, ψ is the physical propagating field. Observe, however, that the self-interactions involve ϕ and are therefore highly non-local with respect to ψ. To see this, we can formally write $\phi = \frac{a}{\partial_z}\psi$; with the boundary conditions that ϕ vanishes at the ends of the interval. ∂_z is invertible and this can be done unambiguously. The interaction Lagrangian for ψ is then

$$\int d^4x dz M_{5D}^3 a(\frac{\partial^2}{\partial_z}\psi)^3 + \cdots \tag{5.22}$$

which is manifestly non-local in the z direction.

Note that the interaction (5.22) formally vanishes in the zero lattice spacing limit, just as $\Lambda \to \infty$ as $a \to 0$ in (5.17). However for any finite lattice spacing a, at large enough distances this term can become important. On a finite interval of size R, the largest wavelength modes of size $\sim R$ will suffer the strongest interactions. These modes $\sim \psi_R$ correspond to the longitudinal polarizations of the lowest KK modes, and are described by the effective 4D action:

$$\int d^4x \left(M_{5D}^3 R(\partial\psi_R)^2 + M_{5D}^3 aR^4(\partial^2\psi_R)^3 + \cdots \right) \tag{5.23}$$

We can see immediately that the amplitude for $\psi_R\psi_R \to \psi_R \to \psi_R\psi_R$ is the same as (5.15). It is precisely the non-local interactions of ψ which lead to the strong amplitudes which force $\Lambda \ll M_{5D}$.

We have seen that, while our starting point *appears* extremely local, with only nearest neighbor hopping terms for the gravitons on the sites, it actually induces highly non-local interactions in the discretized dimension. Why did this happen? Presumably, these effects are related to the fact that we have broken the full 5D diffeomorphism invariance of the theory by our discretization. Usually, however, when gauge symmetries are explicitly broken, a theory gains new degrees of freedom (which correspond to pure gauge modes in the gauge invariant theory). Here (ignoring the irrelevant radion), the number of physical degrees of freedom beneath a^{-1} match exactly between the discrete and continuum theories. Indeed, despite the absence of the full 5D diffeomorphism invariance and Lorentz invariance, the theory still has an exactly "massless" graviton even in the 5D sense, by which we mean gapless excitations with $\omega \to 0$ as $k \to 0$. However, the *interactions* are qualitatively different in the two theories.

So far we have discussed compactification on an interval. When we compactify on a circle, we have to contend with the Goldstone zero mode, which is no longer killed by the boundary conditions. The zero mode has $\psi_0 = a^{-1}\partial_z\phi_0 = 0$ (equivalently, the $n = 0$ mode in (5.14)) and therefore has no kinetic term at all. However, it does appear in interactions! This means that it is inconsistent not to include a kinetic term for ϕ_0 in the theory, which becomes a mass term for the graviphoton π_0^μ (more specifically, the generally covariant plaquette operator [12] made from the "Wilson line" $Y_1 \circ \cdots \circ Y_N$ is generated, which contains the mass term for π_μ^0 as well as various non-linear interactions). It is not surprising that such a term should be generated. In the continuum, the massless graviphoton is associated with a $U(1)$ gauge symmetry inherited from the 5D reparameterization invariance of the circle. This invariance is clearly broken by the discretization. We can see this concretely, because in the continuum ϕ_0 is a pure gauge mode with no dynamics at all, while in our discretization it does appear in interactions. Therefore, the $U(1)$ gauge invariance is explicitly broken and there is no symmetry to prevent π_μ^0 from picking up a mass.

A similar analysis can be done for the case where we start with 3D sites and go up to a 4D theory. Amusingly, in this case, the 3D gravity on the sites does not have any

propagating degrees of freedom – all of the local degrees of freedom in the 4D theory come from the links. Once again, the ϕ fields only acquire a kinetic term by mixing with the site metrics, and the same sorts of non-local interactions arise. The maximum cutoff for a consistent effective theory in this case is

$$\Lambda_{max} \sim \left(\frac{M_{Pl}^2}{R^5} \right)^{1/7} \tag{5.24}$$

Of course, the discretization of one of three spatial dimensions in our universe leads to the breaking of rotational invariance and there are already very severe limits on such effects. But even ignoring this the above cutoff is still incredibly low; taking $R \sim 10^{28}$ cm to be about the size of the universe today, $\Lambda_{max}^{-1} \sim 10^{12}$ cm! Clearly such discretizations are not sensible to describe our 4D world.

Are there discretizations that avoid inducing non-local interactions and correctly reproduce local continuum physics? Simply adding next-to-nearest neighbor interactions, or anything similar which is essentially local on the lattice, cannot possibly work. We can get a hint of what is needed by looking at a discretization that is *guaranteed* to work, at least at tree-level. Take a continuum theory on an interval and simply *truncate* the KK tower, keeping only N of the modes. Of course, we do not need an infinite number of states [45] for an effective theory. By KK momentum conservation, the tree-level scattering amplitudes for the lowest KK modes cannot involve the truncated modes, and therefore their scattering amplitudes coincide with the (healthy) ones of the continuum theory. The strongest tree level amplitudes come from the scattering of the modes between $\sim N/2$ and N near the top of the tower, of mass $\sim N/R$. As is discussed in detail in [92] the strong coupling scale in this case is determined by the Λ_3 scale associated with this mass:

$$\Lambda \sim \left((\frac{N}{R})^2 M_{Pl} \right)^{1/3} \sim \left(\frac{M_{5D}^3 R}{a^4} \right)^{1/6} \tag{5.25}$$

So, for any lattice spacing a, the model makes sense as an effective theory up to energies parametrically higher than the top of the KK tower. Therefore in contrast with our minimal discretization, we can take a limit where the new strong amplitudes induced by the discretization are no more important than those associated with the 5D Planck scale M_{5D}.

At least at tree-level, this model can make sense all the way up to M_{5D}, reproducing local 5D physics.

Starting with these N modes in momentum space, we can go back to an N-site theory in position space by a discrete Fourier transform. The sharp momentum truncation induces interactions between sites in position space that are not strictly local:
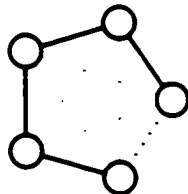
The Fourier transform of a step function has a rapid oscillatory behavior, but dies off as a *power* with distance in position space [92], rather than exponentially as in genuinely local theories.

The non-local interactions we found in the minimal discretization have the same origin as the strong-coupling effects for a single massive graviton discussed in [12]. These strong coupling effects were softened in curved backgrounds (such as AdS), so it would interesting to look for similar improvements in the context of discrete dimensions, where the sites are taken to have AdS geometries. Another straightforward exercise is to discretize warped geometries (such as AdS_5). Since quantum gravity in AdS_5 is dual to a 4D field theory, it would be interesting to see whether a naive discretization can do a better job in describing the continuum physics in this case, at least at scales larger than the AdS_5 length.

We have seen that a naive local discretization of gravity induces non-local interactions at long distances, in sharp contrast to gauge theories, where the minimal discretization is perfectly well-behaved. This is apparent already at tree-level for a single discrete dimension, and makes it impossible to reproduce the correct continuum physics within a consistent effective field theory. We have also seen that a less local discretization, comprising very specific interactions between distant sites in position space, does better than the minimal discretization, and can successfully reproduce local physics at low energies at least at tree-level. This discretization follows from a simple truncation of the continuum Kaluza-Klein tower. In gauge theory, such a truncation is artificial and the corresponding non-local

interactions in position space give rise to various pathologies: it is interesting that in the gravity case this "artificial" discretization is superior to the minimal one.

Ultimately, it is of interest to consider full space or space-time lattices that may provide a sensible definition of quantum gravity in four dimensions. However, despite the many hints of discrete structures underlying quantum gravity, lattice approaches have all suffered from the inability to reproduce the correct continuum physics at low energies. Our simple examples suggest that a successful discretization of gravity cannot look local, but should involve a special set of non-local interactions between distant sites. It is tempting to speculate that this is a reflection of a UV/IR connection in quantum gravity, and that the overly-local nature of most approaches to lattice gravity is responsible for their failures at low energy. Of course, further progress requires finding a principle that dictates the structure of the required non-local interactions.

Acknowledgements

Chapter 6

Constructing Gravitational Dimensions

6.1 Introduction

There are many compelling reasons to study discrete gravitational dimensions. The ultimate goal. of course. is to construct a space-time lattice which reproduces general relativity at low energies. A more practical application would be towards phenomenological extensions of the standard model. Here we use them to characterize what type of low energy effective theories might have arisen from the compactification of a continuous extra dimensional space. Until recently, the best approach to this problem seemed to be to take an extra dimensional model and work out the low energy theory by explicitly integrating out the extra dimensions. Because the KK tower of such theories can be truncated at very high energy without harming the low energy theory, any such model can be interpreted as a discrete theory space by a simple Fourier transform. So the question becomes: which theory spaces produce low energy effective theories with an extra dimensional interpretation? Because we have seen how to study such gravitational theory spaces directly in Chapter 4 we now have a very general approach to the problem.

Normally, we would expect that a discrete extra dimension should look continuous for

105

small enough lattice spacing. This is true for gauge theories, where any haphazardly constructed theory space that looks continuous at the linear level can be made to look continuous at the non-linear level if the discretization is taken sufficiently fine. That is, violations of unitarity from the gauge boson interactions can be pushed above the natural cutoff of the higher dimensional theory by simply shrinking the lattice spacing. In Chapter 5, it was shown that the same simple intuition does not apply for discrete dimensions involving gravity. For example, the continuum limit of a discretization with only nearest neighbor hopping terms must have interactions, apparent at low energy, which are highly non-local in the extra dimension. The problem is that for gravity self-consistent effective field theory imposes a limit on how weak we can make the unitarity violating effects. In Chapter 5, the origin of this impediment was traced to the crazy scalar longitudinal mode of a massive graviton which propagates only after mixing with the transverse modes. It was also argued in Chapter 5 that the truncated KK theory of a single compact extra dimensional model *does* have a local continuum limit. However, no explanation was given about how the scalar longitudinal mode is dealt with from the point of view of the low energy effective theory.

In this chapter, we begin to explore how to construct theory spaces from the ground up. We elaborate on the results of Chapters 4 and 5 and close the book on a number of issues left unresolved by those investigations. First, we attempt to improve the minimal nearest neighbor discretization by adding non-linear, but still nearest neighbor, interactions among the site gravitons. If there were a extension of the Fierz-Pauli Lagrangian for a massive graviton which controlled the dangerous scalar longitudinal mode, this could be replicated around the minimal model and the continuum limit would be drastically improved. In Section 6.3 we show that no such extension exists. In fact, we completely characterize all non-linear extensions Fierz-Pauli and show conclusively that a theory for a single graviton of mass m_g must break down by $(m_g^2 M_{Pl})^{1/3}$. Next, we explore a truncated KK theory for the case of a circle. We study the interacting Lagrangian of this theory in great detail and compute all of the strongest tree-level amplitudes. The most dangerous amplitudes involving the troublesome scalar longitudinal mode of the lightest massive graviton are

..

softened by the exchange of heavier gravitons and the massless graviphoton. The radion also contributes, but not to the strongest diagrams. Finally, we make some comments about various broken symmetries and discuss some implications of this work. Much of the technical details are removed to Appendices A and B; all of the important qualitative results are presented in the main text. The chapter follows very closely to [92].

6.2   Goldstone Bosons and the Minimal Discretization

We begin with a review of Chapters 4 and 5 with a few added niceties. The theories we will consider all contain at least one massive graviton, and therefore involve Lagrangians of Fierz-Pauli [47] form:

$$\mathcal{L} = M_{\text{Pl}}^2 \sqrt{g} R(g) + M_{\text{Pl}}^2 m_g^2 (g_{\mu\nu} - \eta_{\mu\nu})(\eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\mu\nu}\eta^{\rho\sigma})(g_{\rho\sigma} - \eta_{\rho\sigma}) + \cdots \tag{6.1}$$

The mass term explicitly breaks general coordinate invariance (GC) and leads to the propagation of the longitudinal modes of the graviton. It is helpful to project out these modes directly as separate fields which can be interpreted as the Goldstone bosons for the breaking of the GC symmetry. This is done by applying the coordinate transformation $x^\alpha \to y^\alpha(x) = x^\alpha + \pi^\alpha(x)$ to the Lagrangian. The dependence of the new Lagrangian on the Goldstone bosons π^α conveys all the effects of the broken symmetry. More explicitly, we apply the following replacement to (6.1):

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} \equiv \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial y^\beta}{\partial x^\nu} g_{\alpha\beta}(y) = (x^\alpha + \pi^\alpha)_{,\mu}(x^\alpha + \pi^\beta)_{,\nu} g_{\alpha\beta}(x + \pi) \tag{6.2}$$

It also follows that the Lagrangian which results from this replacement is generally coordinate invariant. After all, the π^α represent *all* the symmetry violating effects. Of course, π^α must transform non-linearly, but its transformation law is simply induced from the transformation of y^α and given in Chapter 4.

At this point, it is useful to expand the metrics around flat space $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and the Goldstone bosons as $\pi^\alpha = \eta^{\alpha\mu} A_\mu$. Then, after an integration by pars, the Lagrangian

(6.1) becomes:

$$\mathcal{L} = \frac{1}{4}M_{Pl}^2(-h_{\mu\nu,\alpha}^2 + 2h_{\mu\nu,\mu}^2 - 2h_{\mu\nu,\mu}h_{,\nu} + h_{,\mu}^2) + M_{Pl}^2 m_g^2(F_{\mu\nu}^2 + 2A_{\mu,\nu}h_{\mu\nu} - 2A_{\mu,\mu}h) + \cdots \quad (6.3)$$

The first part is the standard kinetic term for a massless graviton, and the second part contains the standard kinetic term for A_μ. Note that if we had chosen a tensor structure for the mass term in (6.1) different from Fierz-Pauli, A_μ would have non-standard kinetic term signalling unitary violation at $\sim m_g$.

A_μ is an interacting vector boson, which for lack of any gauge symmetry, contains three propagating degrees of freedom. We can separate out its longitudinal mode, which corresponds to the scalar longitudinal mode of $h_{\mu\nu}$ by substituting $A_\mu \rightarrow A_\mu + \phi_{,\mu}$. This establishes an artificial $U(1)$ symmetry for which ϕ is the Goldstone boson. We will return to this symmetry in Section 6.3. Using a more schematic notation, the Lagrangian becomes:

$$\mathcal{L} = M_{Pl}^2 h\Box h + M_{Pl}^2 m_g^2(A\Box A + h\Box\phi) + \cdots \quad (6.4)$$

In this Lagrangian, ϕ only gets a kinetic term from mixing with h. Naturally, because ϕ always appears with two derivatives, the only way it could get a proper kinetic term is through mixing. Nevertheless, this feature is the source of all the bizarre features of massive gravitons discussed in Chapters 4 and 5 and expounded here.

To study the interacting theory, we need to canonically normalize the fields: $h_{\mu\nu}^c = M_{Pl}h_{\mu\nu}$, $A_\mu^c = m_g M_{Pl}A_\mu$, and $\phi^c = m_g^2 M_{Pl}\phi$. Thus, each interaction will have an associated scale which we can read directly off the Lagrangian. Because all the strong interactions, involving the Goldstone fields, come out of the mass term in (6.1), we can derive a general formula:

$$m_g^2 M_{Pl}^2 A_\mu^{n_A}\phi^{n_\phi}h^{n_h} = \Lambda_\lambda^{4-n_A-n_\phi-h_h}A_\mu^{cn_A}\phi^{cn_\phi}h^{cn_h} \quad (6.5)$$

where

$$\Lambda_\lambda = \left(m_g^{\lambda-1}M_{Pl}\right)^{1/\lambda}, \qquad \lambda = \frac{3n_\phi + 2n_A + n_h - 4}{n_\phi + n_A + n_h - 2} \quad (6.6)$$

This implies, for example, that the strongest vertex is ϕ^3 which has the scale $\Lambda_5 = (m_g^4 M_{Pl})^{1/5}$. The amplitude for a simple exchange diagram involving this vertex will grow as $\mathcal{A} \sim E^{10}/\Lambda_5^{10}$.

Incidentally, it may seem strange that the Lagrangian (6.1) should be based on $\sqrt{g}R$ when general coordinate invariance is explicitly broken by the mass term. But this partial GC symmetry guarantees that all the interactions coming from the $\sqrt{g}R$ term involve only transverse polarizations. If this were not true, and a term like $M_{Pl}^2\partial^2 h^3$ were present with arbitrary tensor structure, it would produce interactions of ϕ which a simple calculation shows get strong at Λ_7. So, the GC symmetry in the $\sqrt{g}R$ term, which has *all* the interactions in unitary gauge, actually raises the scale of strong coupling. While this is not a qualitative improvement, it does demonstrate that Λ_5 is not the lowest possible scale where a theory for single massive graviton based on Fierz-Pauli could break down. In fact, the whole point of introducing Goldstone bosons as a symmetry breaking effect is that we can start at Λ_5: the cancellation of the Λ_7 diagrams, which would be obscure in unitary gauge, is given for free.

We continue our review by looking at the minimal lattice explored in Chapter 5. The theory space picture looks like:
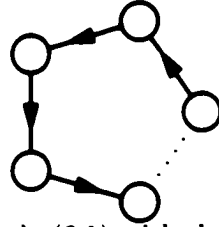
$$(6.7)$$

The associated Lagrangian is simply (6.1) with the mass terms replaced by hopping terms:

$$\mathcal{L}_{\min} = \sum_j M^2\sqrt{g^j}R[g^j] + M^2m^2\sqrt{g^j}(g^j_{\mu\nu} - g^{j+1}_{\mu\nu})(g^{\mu\rho}_j g^{\nu\sigma}_j - g^{\mu\nu}_j g^{\rho\sigma}_j)(g^j_{\rho\sigma} - g^{j+1}_{\rho\sigma}) \quad (6.8)$$

The hopping terms break all but one of the general coordinate invariances. So we restore these symmetries by by replacing:

$$g^{j+1}_{\mu\nu} \to \frac{\partial y^\alpha_j}{\partial x^\mu}\frac{\partial y^\beta_j}{\partial x^\nu}g^{j+1}_{\alpha\beta}(y_j) \quad (6.9)$$

Next, we expand metrics around flat space and the y_j in terms of vector and scalar Goldstones a^j_μ and ϕ^j (using lower case and j for the site basis). Then the Lagrangian looks like:

$$\mathcal{L}_{\min} = M^2 h_j\Box h_j + M^2m^2\{(h_j - h_{j+1})^2 + (h_j - h_{j+1})\Box\phi_j + a_j\Box a_j + \phi_j\phi_j\phi_j + \phi_j a_j a_j\} + \cdots \quad (6.10)$$

To diagonalize the mass matrix, we take the standard linear combinations: $h_j = e^{2\pi i \frac{jn}{N}} G_n$, $a_j = e^{2\pi i \frac{jn}{N}} A_n$, and $\phi_j = e^{2\pi i \frac{jn}{N}} \Phi_n$ (uppercase and n for the momentum basis). Then, summing over j, and using the approximation $m_n \sim m\frac{n}{N}$ the Lagrangian becomes:

$$\mathcal{L}_{\min} = NM^2 G_n \Box G_{-n}$$

$$+ NM^2 m^2 \{ \frac{n^2}{N^2} G_n G_{-n} + \frac{n}{N} G_n \Box \Phi_{-n} + A_n \Box A_{-n} + \Phi_n \Phi_m \Phi_{-n-m} + \Phi_n A_m A_{-n-m} + \cdots \}$$

Just as with a single massive graviton, we can read of the strength of the interactions after going to canonical normalization: $G_n = \frac{1}{\sqrt{N}M} G_n^c$, $A_n = \frac{1}{\sqrt{N}\sqrt{m}} A_n^c$, and $\Phi_n = \frac{\sqrt{N}}{n M m^2} \Phi_n^c$. In terms of the the physical scales $M_{\rm Pl} = M\sqrt{N}$ and $m_1 = m/N$ the strongest interactions look like:

$$\mathcal{L} = \cdots + \frac{1}{NM_{\rm Pl}m_1^4} \Phi_1^c \Phi_1^c \Phi_{-2}^c + \frac{1}{NM_{\rm Pl}m_1^2} \Phi_1^c A_1^c A_{-2}^c + \cdots \tag{6.11}$$

We then read off that the strong coupling scale, set by the Φ^3 vertex is:

$$\Lambda_{\min} = (Nm_1^4 M_{\rm Pl})^{1/5} \tag{6.12}$$

This scale seems reasonable. Formally, $\Lambda_{\min}$ goes to ∞ as $N \to \infty$, and so we can reproduce linearized 5D gravity at low energy. However, within a consistent effective field theory, we can never take $\Lambda_{\min}$ higher than the mass of the heaviest modes in the theory $m_N \sim Nm_1$. This constraint, can be written as:

$$\Lambda_{\min} < \Lambda_{\max} = M_{\rm 5D}(RM_{\rm 5D})^{-5/8} \tag{6.13}$$

where $R = 1/m_1$ is the size of the discrete dimension and $M_{\rm 5D} = (m_1 M_{\rm Pl}^2)^{1/3}$ is the 5D Planck scale. Since $\Lambda_{\max}$ must be less than $M_{\rm 5D}$ this theory has no hope of looking like 5D gravity in the continuum limit.

Nevertheless, there is nothing wrong with taking $N \to \infty$ keeping $M_{\rm Pl}$ fixed in the minimal discretization. The resulting continuum theory will be a consistent effective field theory, even if it cannot be interpreted as having a smooth extra dimension. The argument in Chapter 5 for why the continuum limit will be non-local can be paraphrased as follows. The interactions in (6.10) are in terms of ϕ_j, but ϕ_j gets a kinetic terms from coupling to

$h_j - h_{j+1} = \Delta_z h_j$. Equivalently, $\Phi = \Delta_z \phi_j = \phi_j - \phi_{j+1}$ is the physical, propagating field. So the dangerous interactions are really $\phi_j^3 \sim \frac{1}{\Delta_z^3} \Phi^3$ which have a non-local continuum limit.

6.3 Improving the Minimal Model

The simplest improvement on the minimal discretization would be a model which still only has nearest neighbor interactions, but whose unitary gauge Lagrangian is a more complicated function of $g_{\mu\nu}^j - g_{\mu\nu}^{j+1}$. These models are particularly easy to study, because all of their features can be understood from simply looking at non-linear extensions of the Fierz-Pauli Lagrangian for a single massive graviton. Of course, it is unlikely that this approach will provide a significant improvement over the minimal discretization, because these modifications are still strictly local, and we are trying to cure a non-local disease. Nevertheless, this fairly clean set of models will help us understand the locality problem within the low energy field theory. And if they were to succeed (which they won't) we would have all the freedom to construct gravitational theory spaces that we have for gauge theory spaces.

To begin, we should address the question of what property of the effective theory guarantees locality in the continuum limit. Recall that the obstacle to taking a smooth continuum limit of the minimal discretization is that all the modes are not necessarily lighter than the cutoff, a *sine qua non* of a consistent effective field theory. No such restriction exists for a weakly coupled gauge theory because the cutoff $\Lambda \sim 4\pi N m_1/g$ is *always* above the top of the tower: $\Lambda > m_N \sim N m_1$. For gauge theory, the guarantee follows from the fact that Λ depends on N and m_1 only through the product $N m_1 \sim m_N$. In contrast, for gravity the scale $\Lambda_{\min} = (N m_1^4 M_{Pl})^{1/5}$ does not depend on N and m_1 in an auspicious combination. However, if we could find a Lagrangian for a single massive graviton which breaks down at $\Lambda = \Lambda_2 = \sqrt{m_g M_{Pl}}$ then the discretization based on this Lagrangian would have $\Lambda = \sqrt{N m_1 M_{Pl}}$ and would automatically satisfy the consistency constraint. This theory would have a local continuum limit. So our task becomes simply: extend the Lagrangian for a massive graviton so that it gets strong at Λ_2 (or higher).

Note, in passing, that the Λ_2 scale for a single massive graviton is the geometric mean between M_{Pl} and m_g. In particular, if we take the graviton to have a Hubble mass $m_g \sim H$, then $\Lambda_2 \sim \text{mm}^{-1}$, which happens to be the current limit to which gravity has been experimentally probed. Of course, the graviton could not have a Hubble mass because of other constraints from non-linear effects around large massive sources, as discussed at length in Chapter 4. But if we look only at short distance constraints, raising the scale for strong coupling of a single massive graviton to Λ_2 would be absolutely necessary to avoid obvious contradiction with experiment.

Now, any Lagrangian we consider must start with a quadratic term of Fierz-Pauli form:

$$\mathcal{L}_{FP} = h_{\mu\nu}^2 - h^2 \tag{6.14}$$

Since we already understand kinetic mixing, and are not presently interested in the relatively weakly coupled transverse modes, let us introduce the Goldstones as in (6.2) and then set $h_{\mu\nu} = 0$. We will be making this transformation often for the rest of the Chapter and denote it $\rightsquigarrow$. It is equivalent to replacing:

$$h_{\mu\nu} \rightsquigarrow A_{\mu,\nu} + A_{\nu,\mu} + A_{\alpha,\mu} A_{\alpha,\nu} \tag{6.15}$$

Thus, after an integration by parts

$$\mathcal{L}_{FP} \rightsquigarrow -F_{\mu\nu}^2 - 4 A_{\mu,\mu} A_{\mu,\nu} A_{\mu,\nu} + 4 A_{\mu\nu} A_{\nu\alpha} A_{\mu\alpha} + \cdots \tag{6.16}$$

The appearance of $F_{\mu\nu}^2$ in (6.16) is suggestive. Recall that ϕ, which we have not yet introduced into (6.16), is the longitudinal mode of the vector field A_μ. It is the Goldstone boson for the breaking of a "fake" $U(1)$ symmetry which (6.16) has already at quadratic level. Of course, we cannot expect the entire Lagrangian to have a $U(1)$ symmetry, because ϕ is necessary to reproduce 5D gravity at the linear level. But we might hope that by adding cubic and higher order terms to (6.14), we can achieve a gauge invariance in the Goldstone Lagrangian, that is, the Lagrangian after the $\rightsquigarrow$ transformation. In other words, the Fierz-Pauli structure may be the first part of an expansion fixed by $U(1)$ gauge invariance of the vector longitudinal modes. Moreover, we can see from (6.6) that all the interactions we

are trying to get rid of, the ones which get strong below Λ_2, involve the field ϕ. So this symmetry condition is sufficient for the construction of discretizations with local continuum limits.

Alas, it turns out that the $U(1)$ is a complete red-herring. We will now see not only that the $U(1)$ invariance embedded in the Fierz-Pauli structure is restricted to the quadratic terms, but, more strongly, that there is no way to raise the scale of strong coupling for a single massive graviton higher than Λ_3.

6.3.1 Extending Fierz-Pauli

At this point, it is handy to introduce some notation. The vector of Goldstones, A_μ, will always come with a derivative, so we can represent $A_{\mu,\nu}$ as a matrix:

$$A \equiv A_{\mu,\nu} \Rightarrow A^T = A_{\nu,\mu} \tag{6.17}$$

$$F \equiv A_{\mu,\nu} - A_{\nu,\mu} = A - A^T \tag{6.18}$$

$$\Phi \equiv \phi_{,\mu,\nu} = \Phi^T \tag{6.19}$$

$$1 \equiv \eta_{\mu\nu} \tag{6.20}$$

Projecting out the longitudinal modes from a given unitary gauge Lagrangian amounts to replacing:

$$h_{\mu\nu} \rightsquigarrow A + A^T + A^T A \tag{6.21}$$

Also, since, by Lorentz invariance, we will always be taking traces of such matrices, we define the $[\cdots]$ notation by:

$$[A \cdots A] \equiv \text{Tr}\,[A \cdots A] = A_{\mu,\nu} \cdots A_{\alpha,\mu} \tag{6.22}$$

So, in the new notation, the Fierz-Pauli term becomes:

$$
\begin{aligned}
\mathcal{L}_{\text{FP}} &= [h^2] - [h]^2 \\
&\rightsquigarrow [(A + A^T + A^T A)^2] - (2[A] + [AA^T])^2 \\
&= 2[AA^T] + 2[AA^T] - 4[A]^2 + 4[A^2 A^T] - 4[A][AA^T] + [A^T A A^T A] - [AA^T]^2 \quad (6.23)
\end{aligned}
$$

The third line involved an integration by parts. We can express this in terms of the symmetric and antisymmetric parts of A: Φ and F. That is. we set $A = \Phi + F$ and $A^T = \Phi - F$. Then.

$$\mathcal{L}_{\mathrm{FP}} \sim -4[F^2] + 4[\Phi^3] - 4[\Phi][\Phi^2] + [F^4] - [F^2]^2 + [\Phi^4] - [\Phi^2]^2 - 4[\Phi^2 F^2] + 2[\Phi F \Phi F] + 2[F^2][\Phi^2]$$

With this notation. it will be much easier to study extensions of Fierz-Pauli. The $U(1)$ symmetry we are searching for implies that the Goldstone Lagrangian should depend only on F and not on Φ. First. observe that

$$h \sim 2\Phi + \Phi^2 + \Phi F - F\Phi - F^2 \tag{6.24}$$

Since the only first order term in the expansion of h is Φ. we can always eliminate the Φ self-couplings from the Lagrangian. For example. we cancel the third order terms by adding

$$\mathcal{L}_3 = -\frac{1}{2}[h^3] + \frac{1}{2}[h][h^2] \tag{6.25}$$

Thus the lowest order Φ self couplings in $\mathcal{L}_{FP} + \mathcal{L}_3$ will be Φ^4. These can be eliminated by adding an $\mathcal{L}_4$ with quartic terms. and so on. By induction. it is easy to see that all the self-interactions of the scalar can be eliminated in this way.

The next order gauge-violating terms look like $F^2\Phi^2$. Up to fourth order. there are 12 terms we must eliminate:

$$[\Phi^3], [\Phi]^2[\Phi], [\Phi]^3, [\Phi^4], [\Phi^3][\Phi]. [\Phi^2]^2. [\Phi^2][\Phi]^2. [\Phi]^4. [F^2\Phi^2], [F\Phi F\Phi]. [F^2][\Phi^2]. [F^2][\Phi]^2 \tag{6.26}$$

These are related by two equations that come from integration by parts:

$$[\Phi^3] = 2[\Phi]^2 + [\Phi]^3 \quad \text{and} \quad [\Phi^4] = [\Phi]^3[\Phi] + [\Phi^2][\Phi^2] + [\Phi^2][\Phi]^2 + [\Phi]^4 \tag{6.27}$$

So there are 10 independent terms which must vanish. However. the most general Lagrangian up to fourth order in h has only 8 terms:

$$\mathcal{L}_\lambda = c_1[h^3] + c_2[h^2][h] + c_3[h]^3 + q_1[h^4] + q_2[h^2][h^2] + q_3[h^3][h] + q_4[h^2][h]^2 + q_5[h]^4 \tag{6.28}$$

We might also consider terms with derivatives acting on h, but these cannot produce terms of the form (6.26). Therefore. unless there is some special arrangement. we do not have enough freedom to fabricate a $U(1)$ symmetry.

Still, it may be possible that although the Lagrangian is gauge dependent. all the physical scattering processes involving the ϕ fields vanish. This could be understood as a non-linearly realized $U(1)$ symmetry which is obscured by our choice of the transverse and longitudinal modes of A_μ. For example, the field redefinition $A_\mu \rightarrow B_\mu + B_3 B_{\mu,3}$ will produce interactions in the non-interacting Lagrangian $(A_{\mu,\nu} - A_{\nu,\mu})^2$. These interactions will not vanish by integration by parts. but all the physical scattering amplitudes involving the new B_μ fields will be zero. We could certainly try to classify all non-linear field redefinitions. and all other reasons that the amplitudes may vanish while the interactions do not. But it is more straightforward just to compute the dependence of the strongest scattering amplitudes on the coefficients in (6.28). This is done in Appendix A. The conclusion is that it is simply impossible to extend the Fierz-Pauli Lagrangian so that unitarity is preserved for a single massive graviton above Λ_3.

6.4 Truncated KK Theory

We have seen that it is impossible to eliminate all the dangerous amplitudes for scattering of the scalar longitudinal modes of a massless graviton by a non-linear extension of the Fierz-Pauli Lagrangian. Had this been possible. we could have used the extended Lagrangian as a template for the link structure in a discretization with only nearest neighbor interactions, and the resulting theory would have had none of the problems of the minimal model discussed in Chapter 5. This failure is disappointing, but could have been anticipated from the fact that the continuum sickness of the minimal model is non-locality in the extra dimension. so its cure should involve non-locality on the lattice.

Despite the discouraging results of Section 6.3, we know that the scale of strong coupling for a discrete extra dimension can be raised above $\Lambda_{min} = (N m_1^4 M_{Pl})^{1/5}$. As was mentioned in Chapter 5, a truncated KK theory *does* have a local continuum limit. in contrast to

the minimal discretization. This simply follows from the observation that a truncation performed at high energy cannot effect low energy physics without undermining the general assumptions of effective field theory. Somehow the low KK modes are not interacting strongly at Λ_5 or even at Λ_3.

In this section, we study the truncated KK theory in great detail. We work out the Lagrangian in theory space, including the radion, graviphoton, and all the tensor structure. Its non-locality, including the power-law decay of the interactions with distance in the discrete dimension, is apparent. There are a number of subtle issues about the introduction of Goldstone bosons which are also explored. Finally, the amplitudes for all the dangerous diagrams involving the scalar longitudinal polarizations of the lowest KK modes are given. Due to exchange of heavier modes, and, somewhat surprisingly, the radion and graviphoton as well, all the dangerous amplitudes cancel at tree level.

Start with a 5D metric G_{MN}. We will label the compact fifth direction as z and the non-compact directions collectively as x. Then we gauge fix as much as possible, so the metric takes the form

$$g_{5D} = \begin{pmatrix} g_{\mu\nu}(x,z) + e^{2r(x)}V_\mu(x)V_\nu(x) & e^{2r(x)}V_\mu(x) \\ e^{2r(x)}V_\nu(x) & e^{2r(x)} \end{pmatrix} \quad (6.29)$$

r is the radion and V_μ is the graviphoton. In this gauge, neither r nor V_μ depends on z and the Lagrangian becomes:

$$\mathcal{L} = M_{5D}^3 \sqrt{g_{5D}} R_{5D}(g_{5D}) \quad (6.30)$$

$$= M_{5D}^3 \sqrt{g(x,z)} \left\{ e^r R_{4D}(g) + \frac{1}{4}e^{-r}(\partial_z g_{\mu\nu}(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\nu}g^{\rho\sigma})\partial_z g_{\rho\sigma}) \right. \quad (6.31)$$

$$\left. - \frac{1}{4}(V_{\mu,\nu} - V_{\nu,\mu})g^{\mu\nu}g^{\rho\sigma}(V_{\rho,\sigma} - V_{\sigma,\rho}) + \mathcal{L}_V \right\} \quad (6.32)$$

Here $R_{5D}(g_{5D})$ is the five dimensional Ricci scalar constructed from 5D metric. $R_{4D}(g)$ is the 4D Ricci scalar constructed out of $g_{\mu\nu}(x,z)$ which we treat as a 4D metric which happens to depend on a continuous parameter z. $\mathcal{L}_V$ contains interactions of the graviphoton with the other fields, some of which are presented in Appendix B.

Now, we assume the compact dimension is a circle and expand the metric in KK modes: $g_{\mu\nu}(x,z) = \sum G_{\mu\nu}^j(x)e^{2\pi i \frac{jz}{R}}$. We continue the convention that lower-case fields are in the

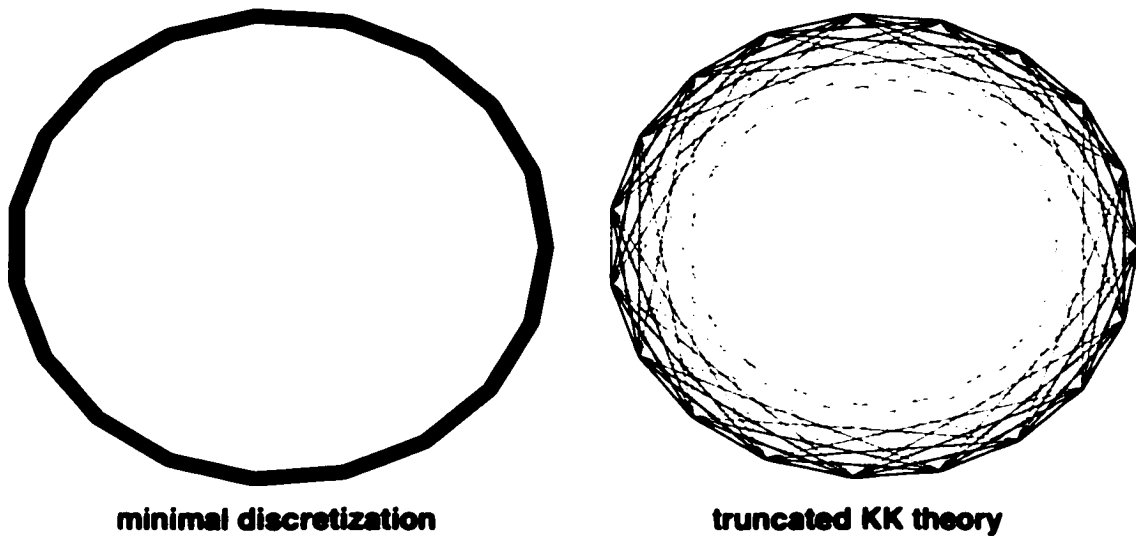**minimal discretization**   **truncated KK theory**

Figure 6.1: In the minimal discretization, links are only between nearest neighbors. In the site basis for the truncated KK theory, there are links between every pair of sites, but the strength of the link dies off with distance. These links which are non-local in theory space remain in the limit of a large number of sites.

site basis and uppercase fields are in the KK basis. Then $\mathcal{L}$ contains the mass terms:

$$\mathcal{L} = M_{\text{Pl}}^2 m_1^2 n^2([G_n G_{-n}] - [G_n][G_{-n}]) + \cdots \tag{6.33}$$

The scale for the masses is set by the radius: $m_1 = 2\pi/R$; and we have introduced the effective 4D Planck scale of the low energy theory: $M_{\text{Pl}}^2 = M_{\text{5D}}^3 R$. We see that the spectrum comprises a massless graviton, a doubly degenerate tower of massive gravitons, and the massless radion and graviphoton.[1]

Now we truncate the theory to N modes, and go back to position space via the discrete Fourier transform $G_n = \frac{1}{N}e^{2\pi i \frac{na}{N}}g_a$. Then, the mass terms become:

$$\mathcal{L}_2 = n^2 G_n G_{-n} = \frac{n^2}{N^2}e^{2\pi i n \frac{\Delta a}{N}}g_a g_{a+\Delta a} = \frac{2n^2}{N^2}\cos\frac{2\pi n \Delta a}{N}g_a g_{a+\Delta a} \tag{6.34}$$

Evidently, there are links between distant sites. As $N \to \infty$

$$\mathcal{L}_2 \to \frac{1}{6\pi^2}N(2g_a^2 - \frac{1}{\Delta a^2}g_a g_{a+\Delta a}) \tag{6.35}$$

---

[1]It might seem that the theory would have been simpler if we had compactified on an interval instead of a circle, thereby removing the mass degeneracy and the graviphoton. In fact, because the KK wavefunctions for the interval are sines instead of exponentials, the interactions in the circle are much easier to work with. This point is discussed further in Appendix B.

Even in this limit. the interactions which are non-local in theory space are only power-law suppressed. This is contrasted to the case of the minimal discretization in Figure 6.1.

Now let us look at the interactions. We know that in a continuum limit. the theory would break down at the 5D Planck scale $M_{5D} = \Lambda_{3/2} = (m_1 M_{Pl}^2)^{1/3}$. We also know that if we truncate to two modes, a massless and a massive one. the theory would break down *at least* by $\Lambda_3 = (m_1^2 M_{Pl})^{1/3} \ll \Lambda_{3/2}$. This is due to scattering of the first massive mode. with mass m_1. So the additional modes must somehow cancel the strongest diagrams involving mode 1. Moreover. we know which fields may contribute to this cancellation. Momentum conservation in the fifth direction translates to KK number conservation in the 4D theory. So the only fields which can contribute to tree-level scattering of the gravitons at the first mass level are gravitons at the second mass level (with mass $m_2 = 2m_1$). the massless graviton. the graviphoton. and the radion. Therefore, we are interested in the tree level processes:

$$\tag{6.36}$$

The Goldstone boson formalism lets us analyze the diagrams in (6.36) very efficiently. We can classify the amplitudes according to their energy dependence. *i.e.* whether they get strong at Λ_5. Λ_4. Λ_3, *etc.*. Because of the result from Section 6.3. that a single massive graviton must break down by Λ_3, it is illuminating to see how the Λ_3 and stronger diagrams in the truncated KK theory cancel. This is demonstrated in explicit detail in Appendix B. As a brief summary, we find that there are no vertices that contribute to Λ_5: the quartic vertices which contribute to Λ_4 are canceled by exchange of the graviphoton; And both the graviphoton and the vector polarization of mode 2 are necessary to cancel the Λ_3 diagrams involving external vectors. This is very important, because it implies that we are very unlikely to find a compactification in which the second lightest massive KK mode is parametrically heavier than the lightest one. It also implies that the first tree level diagrams which the truncation affects involve the Λ_3 scale appropriate external modes at mass level $N/2$. That is, the truncated KK theory gets strong at $\Lambda \sim (N^2 m_1^2 M_{Pl})^{1/3}$. It is worth

pointing out that the calculations in the appendix are a highly non-trivial check on the goldstone boson equivalence theorem for gravity and the entire effective field theory for massive gravitons.

6.5 Discussion and Outlook

In this Chapter, a number of important features of theories with massive gravitons have been analyzed. An arbitrary theory space which has only nearest neighbor interactions will lead. in the continuum limit, to a theory which looks extra-dimensional at the linear level, but has highly non-local interactions. This was understood to be a result of the anomalously strong scattering amplitudes for the light gravitons in the theory. We attempted to eliminate these amplitudes with a non-linear extension of the Fierz-Pauli Lagrangian, but found that no such extension could push the strong coupling scale above Λ_3. We would need to push it as high as Λ_2 to improve the continuum limit of nearest neighbor discretizations. In contrast, a theory space based on the truncation of a KK tower coming from an honest extra dimension does have a local continuum limit. We saw that the scale for the scattering of the light KK modes was pushed above their Λ_3 through the exchange of additional massive gravitons and the massless graviphoton. The truncated KK theory breaks down at the Λ_3 scale of the modes whose tree-level scattering is affected by the truncation. namely those halfway up the tower.

The truncated KK theory gives us much insight into what theory spaces might come from an extra dimensional compactification. We saw that exchange of gravitons on the second mass level canceled strong four-point interactions of gravitons on the first mass level. So it is unlikely that any compactification will have a KK tower with the second lightest graviton *parametrically* heavier than the first. More generally, we should not expect to see a parametrically large mass gap between any set of modes. It would be interesting to see what happens on AdS backgrounds, where the strong coupling problem is ameliorated and there is evidence that at least one graviton mode can be made extremely light [67, 91]. Also, note that the radion was not necessary for cancellation of the strongest diagrams.

and in fact couples with the same strength as the transverse graviton modes. Although it contributes to cancel amplitudes which blow up at $\Lambda_3 < M_{5D} = \Lambda_{3/2}$, we have some freedom to manipulate the radion in the effective theory. For example, simply giving the radion a mass by hand will only affect diagrams at the Λ_2 level applied to the lowest mode. We did not explore the radion interactions in depth, and in fact it may be necessary to perform a wave-packet analysis to understand these sub-dominant amplitudes.

Incidentally, the significance of the Λ_3 scale is not at all clear. It seems to be more fundamental than the Λ_5 scale, which as was pointed out in Section 6.2 is only significant in the somewhat contrived Fierz-Pauli Lagrangian. In contrast, Λ_3 appears as the ultimate upper limit on the strong coupling scale of a theory with a single massive graviton. It is even more intriguing that in a theory constructed to have this strong coupling scale, the natural size for all the operators in unitary gauge is set by Λ_3 and dimensional analysis, a particularly clean situation (see equation (4.74)). Also, Λ_3, applied to the middle modes, determines the scale where the truncated KK theory breaks down. And we saw that it is the strongest scale to which exchange of transverse polarizations of the KK modes, and the radion, can contribute. However, Λ_3 is not enough to use for the construction of nearest neighbor discretizations which are guaranteed to be local in the continuum limit. Certainly, one important area for future investigation is the apparent coincidence of these various results.

Returning to the main theme of this Chapter, it would be nice if the truncated KK theory had some exact symmetry that the minimal discretization lacked. The natural candidate is the general coordinate reparameterization of z. But because z is the discretized dimension, this acts on KK towers as the Virasoro algebra, which does not have any N-dimensional representations. So the symmetry is broken in both theories. In contrast, the GC symmetries of the non-discretized dimensions act faithfully on fields in the site basis. Of course, these GC symmetries are a fake, broken in unitary gauge, and re-established formally with the Goldstone bosons. But we cannot add Goldstone bosons to restore the Virasoro symmetry in either model because the fields belong to a truncated infinite dimensional

multiplet. The $U(1)$ symmetry for which the ϕ fields are Goldstone bosons is also irrelevant. We saw in Section 6.3 that it cannot be made exact by extending Fierz-Pauli. Moreover, we know the ϕ fields must exist in any discretization because we need all the polarizations of the 4D massive gravitons to get the five propagating modes of a massless graviton in 5D. So the truncated KK theory does not seem to have any extra exact symmetries at all.

In fact, there is nothing fundamentally better about the truncated KK theory than the minimal discretization. Both provide consistent low energy effective theories. While locality in the fifth dimension seems nice, there is certainly no experimental evidence to support it. And from the model building point of view, it is likely that there are applications of gravity in theory space for which locality is just irrelevant. Recall that the theory space technology was originally developed in gauge theory to reproduce the phenomenology of an extra dimension [5, 61], but it was soon seen to be well-adapted for the construction of models with a naturally light "little" Higgs [6]. Similarly, gravitational theory spaces may produce applications which have no extra dimensional interpretation at all.

Nevertheless, locality goes hand in hand with improved UV properties. We have seen this already in Section 6.2 where non-locality in the minimal discretization was traced to a low cutoff in the effective theory. And certainly part of the motivation for trying to construct gravitational dimensions comes from string theory, which is both local and UV finite. While the nearest neighbor discretizations have all the exact symmetries of the truncated KK theories, the latter seem to have qualitatively superior UV properties. Now that we understand the appropriate issues, we can work towards establishing a more precise relation between apparent locality and a higher cutoff. This may lead to a better understanding of quantum gravity, and perhaps even a new class of UV completions.

Acknowledgements

Appendix A: Raising the Scale for a Single Massless Graviton

In this appendix it we show that a theory with a single massive graviton must break down by Λ_3. We demonstrate this by considering all possible extensions to the Fierz-Pauli Lagrangian and calculating of all the tree level amplitudes which must cancel. Because of subtleties with a possible non-linearly realized $U(1)$ invariance. this brute-force approach proves to be more convincing than possible symmetry based arguments.

Start with the Fierz-Pauli Lagrangian:

$$\mathcal{L}_{\text{FP}} = \sqrt{g}R(g) + \frac{1}{4}([h^2] - [h]^2) \tag{6.37}$$

The strongest diagrams generated by this Lagrangian blow up at Λ_5, but there are also diagrams which get strong at Λ_4 and Λ_3. More explicitly, for diagrams with four external lines. the Λ_5 is scalar scattering through scalar exchange; Λ_4 is scalar scattering; and Λ_3 is either vector scattering through scalar exchange or scalar vector scattering. In general. vector exchange contributes at the same order as the corresponding quartic vertex.

The only terms which may help cancel these tree-level scattering processes are cubic and quartic in h:

$$\Delta\mathcal{L} = c_1[h^3] + c_2[h^2][h] + c_3[h]^3 + q_1[h^4] + q_2[h^2][h^2] + q_3[h^3][h] + q_4[h^2][h]^2 + q_5[h]^4 \tag{6.38}$$

It is not necessary to consider terms with space-time derivatives acting on h. because they contribute to amplitudes with different momentum dependence than the ones we are trying to cancel.

Now, we want to study the interactions coming from these terms. As usual, we do this by introducing Goldstone bosons, via $h \rightsquigarrow A + A^T + A^T A$ and invoking the equivalence theorem. We work at high energy, where any gauge dependent mass the Goldstone may have is irrelevant, and taken to be zero. This implies that $[A] = [\Phi] = 0$ if these fields correspond to external lines (*i.e.* on-shell, massless particles). In particular, quartic terms which involve $[h]$ and cubic terms which involve $[h]^2$ do not contribute. at first order, to diagrams with four external longitudinal modes. So, q_3, q_4, q_5 and c_3 contribute only above Λ_3 and we can ignore them.

Thus, the interactions which may contribute up to fourth order in $\mathcal{L}_{FP} + \Delta\mathcal{L}$ are:

$$\mathcal{L}_A = (1 + 6c_1)[A^2 A^T] + (-1 + 4c_2)[A][AA^T] + 2c_1[A^3] + 4c_2[A^2][A]$$

$$+ 2q_1[A^4] + (6c_1 + 8q_1)[A^3 A^T] + (3c_1 + 4q_1)[AAA^T A^T] + (\frac{1}{4} + 3c_1 + 2q_1)[AA^T AA^T]$$

$$+ (-\frac{1}{4} + 2c_2 + 4q_2)[AA^T][AA^T] + (2c_2 + 8q_2)[A^2][AA^T] + 4q_2[A^2][A^2] \quad (6.39)$$

The first processes we consider involve scalar exchange. Replacing $A \to \Phi$ once for each of the cubic terms in (6.39) shows that the relevant interactions are:

$$\mathcal{L}_A \supset \phi((1 + 6c_1 + 8c_2)A_{\nu,a,\mu}A_{\mu,a,\nu} + (-1 + 8c_2)A_{a,\nu,\mu}A_{a,\mu,\nu}) \quad (6.40)$$

We have integrated by parts to remove the derivatives from ϕ. Amplitudes coming from these vertices are strong at either Λ_5, if the external A_μ are longitudinally polarized (*i.e.* $A_\mu \to \phi_{,\mu}$), or at Λ_3, if the external A_μ are transverse. Either way, the amplitude will be proportional to the scalar current in (6.40) squared and so (6.40) must exactly vanish. So $c_2 = \frac{1}{8}$ and $c_1 = -\frac{1}{3}$.

Next, we will look at the Λ_3 diagrams from the process $AA \to \Phi\Phi$, which get a contribution from vector exchange:
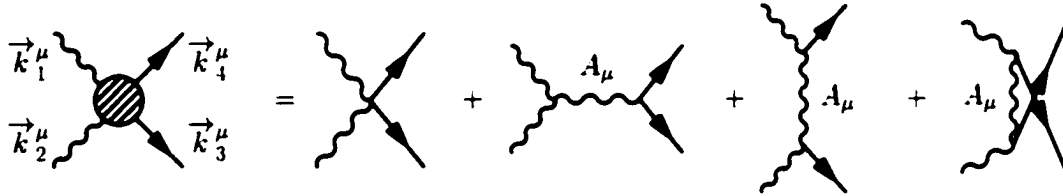
The total amplitude, with the values for c_1 and c_2 we just derived, is:

$$= (4q_1 - 1)[16(k_3\varepsilon_1)(k_3\varepsilon_2)tu + 8(k_1\varepsilon_2)(k_3\varepsilon_1)t^2 + 8(k_1\varepsilon_2)(k_3\varepsilon_1)tu + 8(k_2\varepsilon_1)(k_3\varepsilon_2)tu + 8(k_2\varepsilon_1)(k_3\varepsilon_2)u^2]$$

$$+ (1 + 4q_1 + 8q_2)[8(k_1\varepsilon_2)(k_2\varepsilon_1)(t^2 + u^2)] + (1 + 8q_1 + 16q_2)[-4(\varepsilon_1\varepsilon_2)(t^3 + u^3)]$$

$$+ (3 + 4q_1 + 24q_2)[-8(\varepsilon_1\varepsilon_2)(t^2 u + u^2 t) + (3 + 16q_1 + 32q_2)[-4(k_2\varepsilon_1)(k_3\varepsilon_2)(t^2 + u^2)]$$

$$+ (1 + 24q_1 + 32q_2)[-4(k_3\varepsilon_1)(k_3\varepsilon_2)(t^2 + u^2)] + (3 + 24q_1 + 32q_2)[4(k_1\varepsilon_2)(k_2\varepsilon_1)tu] \quad (6.41)$$

We have used $k_4 = k_1 + k_2 - k_3$ and the Mandelstam variables $s \equiv k_1 k_2 = k_3 k_4$, $t \equiv k_1 k_4 = k_2 k_3$ and $u \equiv k_1 k_3 = k_2 k_4$. The terms are grouped to illustrate that no choice of q_1 and q_2 will make the entire amplitude vanish. Therefore, there is no Lagrangian for a single massive graviton which breaks down above Λ_3.

Appendix B: Cancellations in Truncated KK Theory

In this appendix we calculate the tree level scattering of the light graviton modes in a Kaluza-Klein theory of 5D gravity. Along the way, some subtle issues about setting up the calculation and using the theory space formalism are addressed. In particular, we now defend the choice of a circle, which initially seems more complicated than the interval, because of the extra graviphoton degree of freedom and the degenerate spectrum. For a somewhat similar gauge theory calculation the reader is referred to [32, 93].

For a general compactification, the KK wavefunctions $\chi^J(z)$ depend on the the geometry and the boundary conditions. This lets the Lagrangian be expressed, before the truncation, as the infinite sum

$$\mathcal{L} = \sum o_{ij}^2 G_i G_j + o_{ijk}^3 G_i G_j G_k + \cdots \tag{6.42}$$

The o^n are constants determined by overlap integrals of the χ^J. For example, if we compactify on a circle of radius R, then $\chi_n(z) = e^{2\pi i \frac{nz}{R}}$ and $o_{ij} = \int \chi_i \chi_j dz = \delta(i+j)$.

We now need a prescription for going from the KK basis to the theory space basis. That is, we need to choose discrete wavefunctions $g_i = c_{ij} G_j$. Of course, the obvious choice is a discrete Fourier sum, but we have to be very careful. The essential requirement for a nice theory space representation is that interactions which take place at fixed z in the 5D theory (that is, those which do not depend on ∂_z) should transform to interactions which take place on a single site j. That way, the R_{4D} part of $\mathcal{L}$ (6.29) will have a separate general coordinate invariance for each site, and all the interactions of the longitudinal modes will come from the ∂_z part. This has the advantage that all the interactions will be proportional to $M_{Pl}^2 \partial_z^2 \sim M_{Pl}^2 m^2$ instead of $M_{Pl}^2 \partial^2 \sim M_{Pl}^2 E^2 \gg M_{Pl}^2 m^2$. The only way this will happen is if we can choose the discrete wavefunctions to have all the same overlap integrals as the continuous wavefunctions:

$$o_{i\cdots k} = \frac{1}{R} \int_0^R dz \chi_i \cdots \chi_k = \frac{1}{N} \sum_n c_{in} \cdots c_{kn} \tag{6.43}$$

For a general KK theory this is not possible. But if we compactify on a circle, $\chi_n(z) = e^{2\pi i \frac{nz}{R}}$, then $c_{ab} = e^{2\pi i \frac{ab}{N}}$ satisfies (6.43). Of course, any KK theory has a theory space representa-

tion. but in general the theory space will not simplify the structure of the interactions. For example, it is not possible to satisfy (6.43) on an interval, and so there will be strong interactions simply from the sites' Lagrangians. While amplitudes involving these interactions must cancel other amplitudes at the end of the day, we get this cancellation for free if we choose to the circle where (6.43) holds.

Circle Lagrangian

The truncated KK theory on a circle has a degeneracy at each mass level. We will need the quadratic, cubic. and quartic parts of the KK Lagrangian which come from (6.31):

$$
\mathcal{L} \supset M_{\text{Pl}}^2 \mathcal{L}_K (G_n G_{-n}) - M_{\text{Pl}}^2 m_1^2 \Big[\delta(n + m) mn([G_n G_m] - [G_n][G_m])
$$
$$
+ mn\delta(m + n + p)\{\frac{1}{2}[G_p][G_m G_n]\} - 2[G_m G_p G_n] + [G_n][G_p G_m] + [G_m][G_p G_n]\}
$$
$$
+ mn\delta(m+n+p+q)\{-\frac{1}{4}[G_p G_q][G_n G_m] - [G_n G_p][G_m G_q] + [G_n G_p G_m G_q] + 2[G_n G_p G_q G_m]\} \Big]
$$

$$(6.44)$$

The delta functions enforce KK number conservation. which is equivalent to momentum conservation in the fifth dimension. There are other cubic and quartic terms. but they are irrelevant for the scattering processes we consider. and so we omit them for clarity. $\mathcal{L}_K$ is the standard kinetic Lagrangian for a spin two field. as shown in (6.3). and has the same form for each KK mode.

At this point, we could Fourier transform (6.44) to the site basis. This is not hard. but the explicit form of the site Lagrangian is actually unnecessary for the calculations we are interested in. We know that in unitary gauge, the Lagrangian will have only one general coordinate invariance, under which all the g_j transform as tensors. The $N - 1$ broken GC symmetries correspond to the massive gravitons, whose longitudinal modes produce the strongest interactions in (6.44). To study these interactions. we introduce Goldstone bosons on each site in theory space. That is. we replace each site metric:

$$
g_{\mu\nu}^j \to \tilde{g}_{\mu\nu}^j = \frac{\partial y_j^\alpha}{\partial x^\mu} \frac{\partial y_j^\beta}{\partial x^\nu} g_{\alpha\beta}^j(y_j)
$$

$$(6.45)$$

This restores N copies of general coordinate invariance to the Lagrangian.

This brings up another subtle issue. The transformation properties of the y_α are not uniquely determined, and different choices will actually result in different interactions among the Goldstones. Of course the scale of the strong interactions is the same for all choices, but the calculations of individual amplitudes may be very complicated if we do not introduce Goldstones in a judicious way. For now, we will make it so that the $\bar{g}^\alpha_{\mu\nu}$ are invariant under all N general coordinate transformations. This ensures that the interactions among the Goldstones will respect a global translation invariance around the circle. The drawback of this choice is that the diagonal general coordinate invariance is now obscure and the Lagrangian seems to depend explicitly on N sets of Goldstone bosons. Nevertheless, the Lagrangian must be independent of some non-linear combination of these Goldstone fields.

The y^α are then expanded in terms of Goldstone bosons as usual $y^\alpha_j = x^\alpha + a^\alpha_j$ (and, as usual, we use lower-case for the site basis). The Goldstones are introduced in the site basis, but immediately Fourier transformed: $a^j = e^{2\pi i \frac{jn}{N}} A^n$. Since in the KK basis the masses are diagonal, the bilinear couplings of the Goldstones will be diagonal as well. The quadratic terms in the KK basis are:

$$M_{\text{Pl}}^2(nm_1)^2 \left\{ [(G_n + A_n + A_n^T)(G_{-n} + A_{-n} + A_{-n}^T)] - [G_n + 2A_n][G_{-n} + 2A_{-n}] \right\} \quad (6.46)$$

As with the single massive graviton case, the longitudinal modes of A^n will pick up kinetic terms from mixing with G^n. Thus, the fields have canonical normalization as for a single graviton of mass $m_n = nm_1$.

The interactions we are interested in control the scattering of Goldstones. So after introducing the Goldstones in the site basis, we project out their interactions as in Section 6.3 by setting $g^j_{\mu\nu} = \eta_{\mu\nu}$. In the matrix notation (cf. (6.21) and (6.17)) this corresponds to:

$$g^j \rightsquigarrow a^j + a^{jT} + a^{jT}a^j \quad (6.47)$$

Because we have introduced the Goldstones on the sites, and they are therefore summed over, this implies

$$G^n \rightarrow A^n + A^{nT} + A^{mT}A^{n-m} \quad (6.48)$$

Thus, KK number will be preserved in Goldstone scattering.

We have glossed over another subtle point. Introducing Goldstone bosons on the sites is very different from assigning a general coordinate invariance to each KK mode $G_{\mu\nu}^n$. That would amount to replacing $G^n \leadsto A^n + A^{nT} + A^{nT}A^n$ which leads to interactions which *violate* KK number conservation. We are free to do this, and the physics will be exactly the same since both Lagrangians are the same in unitary gauge, but it will be much harder to calculate. Not only will we have introduced KK number violating interactions, but we will also have introduced interactions into the R_{4D} part of the Lagrangian, which as we noted before, can be canceled for free.

Radion and Graviphoton

The next step is to decide which of the massless fields have strong enough couplings to contribute to scattering of the massive KK gravitons modes. Since the radion, r, and the graviphoton, V_μ, are massless, they get normalized with M_{Pl}. Indeed, V's kinetic term is already present in $\mathcal{L}$ and the radion picks up a kinetic term from mixing with the massless zero-mode graviton. Heuristically,

$$
\begin{aligned}
\mathcal{L} &= M_{Pl}^2 \sqrt{G_0}\{R(G_0)e^r + (V_{\mu,\nu} - V_{\nu,\mu})^2\} + \cdots \\
&= M_{Pl}^2 G^0 \Box G^0 + M_{Pl}^2 r \Box G^0 + M_{Pl}^2 (V_{\mu,\nu} - V_{\nu,\mu})^2 + \cdots
\end{aligned}
\tag{6.49}
$$

So canonical normalization is:

$$
V_\mu^c = M_{Pl} V_\mu \quad \text{and} \quad r^c = M_{Pl} r
\tag{6.50}
$$

Independent of the detailed tensor structure, the couplings to the massive fields have the form:

$$
\mathcal{L} = \cdots + \frac{\partial^2}{M_{Pl}} r^c A_n^c A_n^c + \frac{\partial^3}{M_{Pl} m_n} V^c A_n^c A_n^c + \frac{\partial^4}{M_{Pl} m_n^2} r^c \Phi_n^c \Phi_n^c + \frac{\partial^5}{M_{Pl} m_n^3} V^c \Phi_n^c \Phi_n^c + \cdots
\tag{6.51}
$$

We will study below the strongest contributions to tree-level scattering processes with each type of external lines. We can see from (6.51) that exchange of r is a weaker process than exchange of the Goldstone vector A_μ, while V_μ and A_μ exchange are the same strength. So, to first order, we can ignore r but must include V_μ. We will therefore need the interactions

which are linear in V_μ and do not involve r which were suppressed from the expression (6.32):

$$\mathcal{L}_V = g^{\alpha\beta} g^{\mu\nu} g^{\gamma\delta} V_3 (2g_{\alpha\mu,z} g_{\nu\gamma,\delta} - g_{\alpha\mu,z} g_{\gamma\delta,\nu} - g_{\gamma\delta,z} g_{\alpha\mu,\nu} + \frac{1}{2} g_{\gamma\delta,z} g_{\mu\nu,\alpha} - \frac{3}{2} g_{\gamma\mu,z} g_{\delta\nu,\alpha} + g_{\gamma\mu,z} g_{\nu\alpha,\delta})$$

$$+ g^{\mu\nu} g^{\alpha\beta} (V_{3,\alpha} g_{\mu\nu,z} - V_{3,\mu} g_{\alpha\nu,z}) + g^{\mu\nu} V_3 (2g_{\mu\nu,\alpha,z} - 2g_{\alpha\mu,\nu,z}) + \cdots \quad (6.52)$$

All of the couplings that we will need below come from the following terms in the KK decomposition of (6.52):

$$\mathcal{L} \supset -i M_{\text{Pl}}^2 m_1 n \delta(n+m) V_3 (2G_{\mu\nu}^m G_{\mu\nu,3}^n + G_{\mu\nu,3}^m G_{\mu\nu}^n - 2G_{\mu\nu}^m G_{\mu3,\nu}^n) \quad (6.53)$$

Now let us come back to one of the subtle issues mentioned above. In addition to the radion and the KK gauge boson. we must contend with the zero mode Goldstone boson. A_μ^0. It does not have a kinetic term. and does not pick one up by mixing. But it does have interactions. with the Lagrangian in its current form. The only reason we have this mode at all, is because when we introduced the Goldstone bosons in (6.47) we included one set for each of the N sites. even though the Lagrangian has only $N-1$ broken coordinate invariances. The interactions of A_μ^0 exist because the preserved symmetry is not the one under which all the fields on the sites transform nicely: it is the one where all the KK modes transform nicely. Suppose we had chosen the transformations of the Goldstone bosons so that each $\tilde{g}_{\mu\nu}^j$ were covariant under changes of a single coordinate y^s and invariant under all the others. Then, $\tilde{g}_{\mu\nu}^s = g_{\mu\nu}^s$ and $\tilde{g}_{\mu\nu}^j = y_{j,\mu}^\alpha y_{j,\nu}^\beta g_{\alpha\beta}^j$ for $j \neq s$. Then there would be only $N-1$ sets of Goldstones and we would not have the peculiar A_μ^0 field. The Lagrangian would be the same, but with $a_\mu^s = 0$. In terms of the KK modes of the Goldstones, this implies

$$A_\mu^0 = -e^{2\pi i \frac{sn}{N}} A_\mu^n \quad (6.54)$$

For example. if we we take $s = 0$ then $A_\mu^0 \rightarrow -A_\mu^1 - A_\mu^2 - \cdots$. Now. there will be many KK number violating vertices. In particular, all the heavy KK modes would contribute to scattering of the low modes. But when we sum over all diagrams. KK number should not be violated. It would be nice if we could just use the residual general coordinate invariance

to set $A_\mu^0 = 0$. but it not clear that this is consistent. There is no simple way of introducing Goldstones so that A_μ^0 does not appear at all. As it turns out. if we calculate the scattering from terms generated by the substitution (6.54) into the cubic and quartic vertices involving the A_μ^0, everything vanishes. This implies that we are probably free to just set $A_\mu^0 = 0$.

Goldstone Interactions

From now on. we use the notation $1 \equiv A_{\mu,\nu}^1$. where A_μ^1 is the first KK mode of the vector of Goldstone bosons introduced. Scattering of the the two modes in the first massive level involves exchange of the two modes on the second massive level. So we need $N \geq 5$ to see a non-trivial cancellation. Therefore. we take $N = 5$. Note that taking $N > 5$ will not change the relevant vertices. as these get no contribution from higher KK modes. So we consider the lightest five modes $0, 1, 2, -1 \equiv \bar{1}$ and $-2 \equiv \bar{2}$. and restrict to diagrams involving only external 1 and $\bar{1}$. We will also immediately go to canonical normalization $A_\mu^n \rightarrow M_{\text{Pl}} m_n A_\mu^n$ and $\Phi^n \rightarrow M_{\text{Pl}} m_n^2 \Phi^n$. Now. we present the relevant interactions among these fields.

We start with the cubic vertices. After a lengthy calculation we can isolate the following terms relevant for diagrams with external 1 and $\bar{1}$:

$$\mathcal{L}_C \supset 8[11^T\bar{2}] - 4[1^T1\bar{2}] + 12[11\bar{2}] + 8[1^T1^T\bar{2}] - 14[11][\bar{2}] + 2[1^T1][\bar{2}]$$

$$-16[1][1\bar{2}] - 8[1][1^T\bar{2}] + 12[\bar{2}][1][1] + c.c. \tag{6.55}$$

It is interesting to note that we can perform a field redefinition to remove all the cubic terms:

$$2_\mu \rightarrow 2_\mu + \frac{1}{8}1_\jmath 1_{\mu,\jmath} \tag{6.56}$$

Note that this substitution preserves KK number conservation. This is an example of the type of non-linear transformation we were wary of in Section 6.3. But it does not actually make the calculations of this appendix any easier, so we leave (6.55) as it is.

Next, the couplings of the graviphoton can be written as (from (6.53)):

$$\mathcal{L} \supset 4iV_\jmath(1_{\mu,\nu}\bar{1}_{\jmath,\mu,\nu} - \bar{1}_{\mu,\nu}1_{\jmath,\mu,\nu}) \tag{6.57}$$

We have already dropped everything which vanishes when the Goldstones, which always appear on shell in V-exchange diagrams, are massless. The interactions in (6.57) contribute to 11 scattering at the order Λ_3, which is the same order as the quartic vertices and as 2 exchange.

Next, we look at the quartic vertices with 1 and $\bar{1}$:

$$\mathcal{L}_Q \supset -8[\bar{1}\bar{1}11] - 4[1^T\bar{1}\bar{1}1] + 4[\bar{1}\bar{1}1^T1] - 4[\bar{1}^T11\bar{1}] + 4[\bar{1}^T\bar{1}11]$$

$$- 4[\bar{1}^T\bar{1}^T11] + 2[\bar{1}^T1^T1\bar{1}] + 2[1^T\bar{1}^T\bar{1}1] - 4[\bar{1}1\bar{1}1] - 4[\bar{1}^T1^T\bar{1}1] - 4[\bar{1}^T1\bar{1}1] - 4[1^T\bar{1}1\bar{1}]$$

$$+ 6[11][\bar{1}\bar{1}] - [11][\bar{1}\bar{1}^T] - [11^T][\bar{1}\bar{1}] + 4[1\bar{1}][1\bar{1}] + 4[1\bar{1}][1\bar{1}^T]$$

There are other terms involving [1] and $[\bar{1}]$. These will not contribute when the external lines are on shell, so we have not displayed them here.

To calculate the interactions, it is easiest to project out the real and imaginary parts of the KK fields. This diagonalizes the kinetic terms of the physical fields. To account for the normalization as well, we make the substitution $1 = \frac{1}{\sqrt{8}}(A + iB)$ and $2 = \frac{1}{2\sqrt{8}}(C + iD)$.

Then the entire normalized Lagrangian we will need is:

$$\mathcal{L} \supset -\frac{3}{16}[A^4] - \frac{1}{8}[A^T A^3] - \frac{1}{16}[A^T A^T A^2] + \frac{5}{32}[A^2]^2 + \frac{1}{32}[A^2][AA^T]$$

$$-\frac{3}{16}[B^4] - \frac{1}{8}[B^T B^3] - \frac{1}{16}[B^T B^T B^2] + \frac{5}{32}[B^2]^2 + \frac{1}{32}[B^2][BB^T]$$

$$-\frac{1}{8}[ABAB] + \frac{1}{8}[A^T BAB] + \frac{1}{8}[B^T ABA] + \frac{1}{8}[A^T B^T AB]$$

$$-\frac{1}{4}[A^2 B^2] - \frac{3}{8}[A^T ABB] - \frac{3}{8}[B^T BAA] - \frac{3}{16}[A^T B^T BA] - \frac{3}{16}[B^T A^T AB]$$

$$+\frac{1}{8}[B^T AAB] + \frac{1}{8}[A^T BBA] + \frac{1}{8}[A^T A^T BB] - \frac{1}{16}[B^2][A^2] + \frac{3}{8}[AB][AB]$$

$$+\frac{3}{32}[A^2][BB^T] + \frac{3}{32}[B^2][AA^T] - \frac{1}{8}[AB][AB^T]$$

$$+\frac{1}{\sqrt{8}}[AA^T C] + \frac{1}{\sqrt{8}}[A^T A^T C] - \frac{1}{\sqrt{8}}[AA][C]$$

$$-\frac{1}{\sqrt{8}}[BB^T C] - \frac{1}{\sqrt{8}}[B^T B^T C] + \frac{1}{\sqrt{8}}[BB][C]$$

$$+\frac{1}{\sqrt{8}}[AB^T D] + \frac{1}{\sqrt{8}}[BA^T D] + \frac{1}{\sqrt{8}}[A^T B^T D] + \frac{1}{\sqrt{8}}[B^T A^T D] - \frac{1}{\sqrt{2}}[AB][D]$$

$$-\frac{1}{2}[A^T A] + \frac{1}{2}[A^2] - \frac{1}{2}[B^T B] + \frac{1}{2}[B^2] - \frac{1}{2}[C^T C] + \frac{1}{2}[C^2] - \frac{1}{2}[D^T D] + \frac{1}{2}[D^2]$$

$$-\frac{1}{4}(V_{\mu,\nu} - V_{\nu,\mu})^2 + V_3(A_{\mu,\nu} B_{3,\mu,\nu} - B_{\mu,\nu} A_{3,\mu,\nu})$$

$$(6.58)$$

Test the Lagrangian

To test the Lagrangian, we will look at some characteristic processes. Recall that the strength of a vertex for involving a graviton of mass m_g is given by $\Lambda_\lambda = (m_g^{\lambda-1} M_{\rm Pl})^{1/\lambda}$ where λ is given by (6.6). The strongest processes involve scalar exchange. And the vector exchange diagrams have the same strength as the diagrams coming from the quartic vertices as well as the diagrams with graviphoton exchange. If we separate out scalar exchange, which must cancel by itself, then the strength of a process is determined by the external lines. Since this is the case, we do not have even to project out the scalar in the external lines, we can just leave it as the longitudinal polarization of the vector field. We only insist that the polarization of the eternal vector field satisfies $\varepsilon_{\mu,\mu} = 0$.

The masses of the Goldstone bosons are gauge dependent, but we will only be concerned

with the lowest order tree level effects. and so we take all the Goldstones to be massless. The corrections, of order m^2/k^2 will only contribute to higher order processes which we will ignore. In practice. this means dropping terms which contain $[A]$. $[B]$. $A_{\mu,\mu}$. $B_{\mu,\mu}$. $A_{\mu,\nu,\nu}$ or $B_{\mu,\nu,\nu}$. We have already done this for the Lagrangian (6.58) but more simplifications come about for particular process.
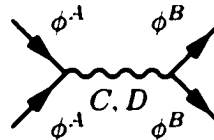
The cubic terms' contribution to C and D exchange can be written as:

$$\mathcal{L} \supset \; - \; \frac{1}{\sqrt{2}} C_\mu (A_{\alpha,3} A_{\mu,\alpha,3} - A_{\alpha,3} A_{3,\alpha,\mu} - B_{\alpha,3} B_{\mu,\alpha,3} + B_{\alpha,3} B_{3,\alpha,\mu}) \qquad (6.59)$$

$$- \; \frac{1}{\sqrt{2}} D_\mu (A_{\nu,\alpha} B_{\mu,\alpha,\nu} + A_{\mu,\alpha,3} B_{3,\alpha} - A_{\alpha,3,\mu} B_{3,\alpha} - A_{\alpha,3} B_{3,\alpha,\mu}) \qquad (6.60)$$

It is easy to see that for scalar exchange $C_\mu \to \phi_{,\mu}^C$ and $D_\mu \to \phi_{,\mu}^D$ the above terms vanish after integrating by parts. This means that all the scalar exchange processes. which are stronger than the corresponding vector exchange processes. vanish.

Next, consider the process $\phi^A \phi^A \to \phi^B \phi^B$. This contributes at the scale Λ_4. If we make the substitution $A_\mu \to \phi_{,\mu}^A$ and $B_\mu \to \phi_{,\mu}^B$ into (6.59) and (6.60) we see that there is no contribution from vector C and D exchange:
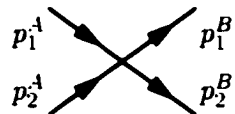

$$= 0 \qquad (6.61)$$

The quartic terms are:

$$\mathcal{L} \supset -[\Phi^A \Phi^A \Phi^B \Phi^B] + \frac{1}{4}[\Phi^A \Phi^B \Phi^A \Phi^B] + \frac{1}{4}[\Phi^A \Phi^B][\Phi^A \Phi^B] + \frac{1}{8}[\Phi^A \Phi^A][\Phi^B \Phi^B] \qquad (6.62)$$

This does not vanish by integration by parts. If we define the Mandelstam variables $s = p_1^A \cdot p_2^A$, $t = p_1^A \cdot p_1^B$ and $u = p_1^A \cdot p_2^B$ then this contributes:

$$= -2s^2(t^2 + u^2) + t^2 u^2 + \frac{1}{2}(t^4 + u^4) + \frac{1}{2}s^4 \neq 0 \qquad (6.63)$$

However, there is also a contribution from the graviphoton:

$$\mathcal{L} \supset V_3(\phi_{,\mu,\nu}^A \phi_{,3,\mu,\nu}^B - \phi_{,\mu,\nu}^B \phi_{,3,\mu,\nu}^A) \qquad (6.64)$$

This contributes through the t- and u- channels:

$$\text{(diagram)} = -\left\{ \frac{t^4(2s+2u)}{-2t} + \frac{u^4(2s+2t)}{-2u} \right\} \tag{6.65}$$

If we apply the relation $s - t - u = 0$ the quartic and exchange contributions exactly cancel.

Note that the C and D fields do not contribute at either Λ_5 (scalar scattering through scalar exchange) or at Λ_4 (scalar scattering through vector exchange). So even if we truncated the theory at the level at the first massive mode. the strong coupling scale would already be $(m_1^2 M_{\text{Pl}})^{1/3}$. However, to see the cancellation of the remaining diagrams. we need to include the effects of the heavier fields.

It is straightforward to work through remaining scattering processes. The computations are more involved. but at tree-level all the amplitudes involving the Lagrangian (6.58) are exactly zero.

Chapter 7

Conclusions

It this dissertation, I have presented a number of applications of effective field theory to the study of quantum gravity. I hope to have shown that even though gravity presents a non-renormalizable quantum field theory, we can learn a tremendous amount about its subtle and non-trivial structure just from a self-consistent low energy effective field theory approach. In curved backgrounds. it was shown in Chapter 2 that self-consistency forces the cutoff to vary with position. Among other things. this confirms the hypothesis of the holographic principle in a very straightforward way. The curved regulator was applied to the calculation of one loop beta functions in Chapter 3 to show that that high scale perturbative unification is compatible with the Randall Sundrum model. In Chapter 4 we saw how theories with massive gravitons can be studied with effective field theory. A formalism was introduced for projecting out the Goldstone bosons associated with broken general coordinate invariance. This formalism was shown to make a number of peculiarities associated with massive gravity, such as the Fierz-Pauli tensor structure and the van Dam-Veltman-Zakharov discontinuity, completely transparent. It was also applied, in Chapter 5, to the construction of discrete gravitational extra dimensions. There, the UV/IR correspondence, observed in Chapter 2, resurfaced. We found that the high energy scale where the discrete theory breaks down is determined in part by the low energy, infrared scale, of the continuum space it is trying to emulate.

134

The search for a consistent theory which contains both quantum mechanics and gravity is one of the most pressing tasks in particle physics. Undoubtedly, entirely new ideas will be needed to understand quantum gravity completely. But it is important, and far more practical, to approach the problem by studying the ideas we already have. My work as a graduate student, as presented in this dissertation, is an attempt to unravel the intricacies of quantum gravity from the ground up. I hope that this work, and more importantly, this approach, will provide an important contribution to the quest for a fundamental understanding of our physical universe.

References

[1] O. Aharony, M. Berkooz, D. Kutasov, and N. Seiberg. Linear dilatons, ns5-branes and holography. *JHEP*, 10:004, 1998.

[2] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali. New dimensions at a millimeter to a fermi and superstrings at a tev. *Phys. Lett.*, B436:257–263, 1998.

[3] C. Aragone and S. Deser. Constraints on gravitationally coupled tensor fields. *Nuovo Cim. A*, 3, 1971.

[4] N. Arkani-Hamed, A. G. Cohen, and H. Georgi. Accelerated unification. 2001.

[5] N. Arkani-Hamed, A. G. Cohen, and H. Georgi. (de)constructing dimensions. *Phys. Rev. Lett.*, 86:4757–4761, 2001.

[6] N. Arkani-Hamed, A. G. Cohen, and H. Georgi. Electroweak symmetry breaking from dimensional deconstruction. *Phys. Lett.*, B513:232–240, 2001.

[7] N. Arkani-Hamed, A. G. Cohen, and H. Georgi. Twisted supersymmetry and the topology of theory space. *JHEP*, 07:020, 2002.

[8] N. Arkani-Hamed, A. G. Cohen, D. B. Kaplan, A. Karch, and L. Motl. Deconstructing (2,0) and little string theories. *JHEP*, 01:083, 2003.

[9] N. Arkani-Hamed and S. Dimopoulos. New origin for approximate symmetries from distant breaking in extra dimensions. *Phys. Rev.*, D65:052003, 2002.

136

[10] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali. Phenomenology, astrophysics and cosmology of theories with sub-millimeter dimensions and tev scale quantum gravity. *Phys. Rev.*, D59:086004, 1999.

[11] N. Arkani-Hamed, S. Dimopoulos, and J. March-Russell. Logarithmic unification from symmetries enhanced in the sub-millimeter infrared. 1999.

[12] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz. Effective field theory for massive gravitons and gravity in theory space. 2002.

[13] N. Arkani-Hamed, M. Porrati, and L. Randall. Holography and phenomenology. *JHEP*, 08:017, 2001.

[14] N. Arkani-Hamed and M. D. Schwartz. Discrete gravitational dimensions. 2003.

[15] V. Balasubramanian and P. Kraus. Spacetime and the holographic renormalization group. *Phys. Rev. Lett.*, 83:3605–3608, 1999.

[16] M. Bander. Gravity in dynamically generated dimensions. *Phys. Rev.*, D64:105021, 2001.

[17] J. M. Bardeen, B. Carter, and S. W. Hawking. The four laws of black hole mechanics. *Commun. Math. Phys.*, 31:161–170, 1973.

[18] J. D. Bekenstein. Black holes and the second law. *Nuovo Cim. Lett.*, 4:737–740, 1972.

[19] J. D. Bekenstein. Black holes and entropy. *Phys. Rev.*, D7:2333–2346, 1973.

[20] J. D. Bekenstein. Generalized second law of thermodynamics in black hole physics. *Phys. Rev.*, D9:3292–3300, 1974.

[21] J. D. Bekenstein. Statistical black hole thermodynamics. *Phys. Rev.*, D12:3077–3085, 1975.

[22] D. Bigatti and L. Susskind. Tasi lectures on the holographic principle. 1999.

[23] D. G. Boulware. Quantum field theory in schwarzschild and rindler spaces. *Phys. Rev.*, D11:1404, 1975.

[24] D. G. Boulware and S. Deser. Inconsistency of finite range gravitation. *Phys. Lett.*, B40:227-229, 1972.

[25] R. Bousso. A covariant entropy conjecture. *JHEP*, 07:004, 1999.

[26] I. L. Buchbinder, D. M. Gitman, V. A. Krykhtin, and V. D. Pershin. Equations of motion for massive spin 2 field coupled to gravity. *Nucl. Phys.*, B584:615-640, 2000.

[27] J. Callan, Curtis G., S. R. Coleman, J. Wess, and B. Zumino. Structure of phenomenological lagrangians. 2. *Phys. Rev.*, 177:2247-2250, 1969.

[28] A. H. Chamseddine, A. Salam, and J. Strathdee. Strong gravity and supersymmetry. *Nucl. Phys.*, B136:248-258, 1978.

[29] S. Chang, J. Hisano, H. Nakano, N. Okada, and M. Yamaguchi. Bulk standard model in the randall-sundrum background. *Phys. Rev.*, D62:084025, 2000.

[30] H.-C. Cheng. Doublet-triplet splitting and fermion masses with extra dimensions. *Phys. Rev.*, D60:075015, 1999.

[31] H.-C. Cheng, C. T. Hill, S. Pokorski, and J. Wang. The standard model in the latticized bulk. *Phys. Rev.*, D64:065007, 2001.

[32] R. S. Chivukula and H.-J. He. Unitarity of deconstructed five-dimensional yang-mills theory. *Phys. Lett.*, B532:121-128, 2002.

[33] S. R. Coleman, J. Wess, and B. Zumino. Structure of phenomenological lagrangians. 1. *Phys. Rev.*, 177:2239-2247, 1969.

[34] C. Csaki, M. L. Graesser, and G. D. Kribs. Radion dynamics and electroweak physics. *Phys. Rev.*, D63:065002, 2001.

[35] T. Damour and I. I. Kogan. Effective lagrangians and universality classes of nonlinear bigravity. *Phys. Rev.*, D66:104024, 2002.

[36] T. Damour, I. I. Kogan, and A. Papazoglou. Non-linear bigravity and cosmic acceleration. *Phys. Rev.*, D66:104025, 2002.

[37] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo. Bulk gauge fields in the randall-sundrum model. *Phys. Lett.*, B473:43–49, 2000.

[38] C. Deffayet, G. R. Dvali, G. Gabadadze, and A. I. Vainshtein. Nonperturbative continuity in graviton mass versus perturbative discontinuity. *Phys. Rev.*, D65:044026, 2002.

[39] J.-G. Demers, R. Lafrance, and R. C. Myers. Black hole entropy without brick walls. *Phys. Rev.*, D52:2245–2253, 1995.

[40] S. Deser, R. Jackiw, and G. 't Hooft. Three-dimensional einstein gravity: Dynamics of flat space. *Ann. Phys.*, 152:220, 1984.

[41] K. R. Dienes, E. Dudas, and T. Gherghetta. Grand unification at intermediate mass scales through extra dimensions. *Nucl. Phys.*, B537:47-108, 1999.

[42] S. Dimopoulos, S. Kachru, N. Kaloper, A. E. Lawrence, and E. Silverstein. Small numbers from tunneling between brane throats. *Phys. Rev.*, D64:121702, 2001.

[43] L. Dolan and M. J. Duff. Kac-moody symmetries of kaluza-klein theories. *Phys. Rev. Lett.*, 52:14, 1984.

[44] M. J. Duff and J. T. Liu. Complementarity of the maldacena and randall-sundrum pictures. *Phys. Rev. Lett.*, 85:2052–2055, 2000.

[45] M. J. Duff, C. N. Pope, and K. S. Stelle. Consistent interacting massive spin-2 requires an infinity of states. *Phys. Lett.*, B223:386, 1989.

[46] G. R. Dvali, G. Gabadadze, and M. Porrati. 4d gravity on a brane in 5d minkowski space. *Phys. Lett.*, B485:208-214, 2000.

[47] M. Fierz and W. Pauli. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. Roy. Soc. Lond.*, A173:211–232, 1939.

[48] J. Garriga, O. Pujolas, and T. Tanaka. Radion effective potential in the brane-world. *Nucl. Phys.*, B605:192–214, 2001.

[49] H. Georgi and S. L. Glashow. Unity of all elementary particle forces. *Phys. Rev. Lett.*, 32:438–441, 1974.

[50] T. Gherghetta and A. Pomarol. Bulk fields and supersymmetry in a slice of ads. *Nucl. Phys.*, B586:141–162, 2000.

[51] S. B. Giddings and E. Katz. Effective theories and black hole production in warped compactifications. *J. Math. Phys.*, 42:3082–3102, 2001.

[52] S. B. Giddings, E. Katz, and L. Randall. Linearized gravity in brane backgrounds. *JHEP*, 03:023, 2000.

[53] W. D. Goldberger and I. Z. Rothstein. Quantum stabilization of compactified ads(5). *Phys. Lett.*, B491:339–344, 2000.

[54] W. D. Goldberger and M. B. Wise. Modulus stabilization with bulk fields. *Phys. Rev. Lett.*, 83:4922–4925, 1999.

[55] M. B. Green and C. B. Thorn. Continuing between closed and open strings. *Nucl. Phys.*, B367:462–484, 1991.

[56] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov. Gauge theory correlators from non-critical string theory. *Phys. Lett.*, B428:105–114, 1998.

[57] L. J. Hall and Y. Nomura. Gauge unification in higher dimensions. *Phys. Rev.*, D64:055003, 2001.

[58] S. W. Hawking. Gravitational radiation from colliding black holes. *Phys. Rev. Lett.*, 26:1344–1346, 1971.

[59] S. W. Hawking. Particle creation by black holes. *Commun. Math. Phys.*, 43:199–220, 1975.

[60] S. W. Hawking and D. N. Page. Thermodynamics of black holes in anti-de sitter space. *Commun. Math. Phys.*, 87:577, 1983.

[61] C. T. Hill, S. Pokorski, and J. Wang. Gauge invariant effective lagrangian for kaluza-klein modes. *Phys. Rev.*, D64:105005, 2001.

[62] S. J. Huber and Q. Shafi. Higgs mechanism and bulk gauge boson masses in the randall- sundrum model. *Phys. Rev.*, D63:045010, 2001.

[63] T. Jacobson and R. C. Myers. Black hole entropy and higher curvature interactions. *Phys. Rev. Lett.*, 70:3684–3687, 1993.

[64] V. Jejjala, R. G. Leigh, and D. Minic. The cosmological constant and the deconstruction of gravity. *Phys. Lett.*, B556:71–79, 2003.

[65] N. Kan and K. Shiraishi. Multi-graviton theory, a latticized dimension, and the cosmological constant. 2002.

[66] A. Karch, E. Katz, and L. Randall. Absence of a vvdz discontinuity in ads(ads). *JHEP*, 12:016, 2001.

[67] A. Karch and L. Randall. Locally localized gravity. *JHEP*, 05:008, 2001.

[68] Y. Kawamura. Triplet-doublet splitting, proton stability and extra dimension. *Prog. Theor. Phys.*, 105:999–1006, 2001.

[69] I. I. Kogan. Multi(scale)gravity: A telescope for the micro-world. 2001.

[70] I. I. Kogan, S. Mouslopoulos, and A. Papazoglou. The m $-¿$ 0 limit for massive graviton in ds(4) and ads(4): How to circumvent the van dam-veltman-zakharov discontinuity. *Phys. Lett.*, B503:173–180, 2001.

[71] I. I. Kogan, S. Mouslopoulos, A. Papazoglou, and L. Pilo. Radion in multibrane world. *Nucl. Phys.*, B625:179–197, 2002.

[72] P. Langacker. Physics implications of precision electroweak experiments. *J. Phys.*, G29:35–48, 2003.

[73] A. Lue. Cosmic strings in a brane world theory with metastable gravitons. *Phys. Rev.*, D66:043509, 2002.

[74] J. M. Maldacena. The large n limit of superconformal field theories and supergravity. *Adv. Theor. Math. Phys.*, 2:231–252, 1998.

[75] C. R. Nappi and L. Witten. Interacting lagrangian for massive spin two field. *Phys. Rev.*, D40:1095, 1989.

[76] D. N. Page. Black hole information. 1993.

[77] A. W. Peet and J. Polchinski. Uv/ir relations in ads dynamics. *Phys. Rev.*, D59:065011, 1999.

[78] M. E. Peskin and D. V. Schroeder. An introduction to quantum field theory. 1995. Reading, USA: Addison-Wesley (1995) 842 p.

[79] J. Polchinski. String theory and black hole complementarity. 1995.

[80] J. Polchinski, L. Susskind, and N. Toumbas. Negative energy, superluminosity and holography. *Phys. Rev.*, D60:084006, 1999.

[81] A. Pomarol. Gauge bosons in a five-dimensional theory with localized gravity. *Phys. Lett.*, B486:153–157, 2000.

[82] A. Pomarol. Grand unified theories without the desert. *Phys. Rev. Lett.*, 85:4004–4007, 2000.

[83] M. Porrati. No van dam-veltman-zakharov discontinuity in ads space. *Phys. Lett.*, B498:92–96, 2001.

[84] M. Porrati. Fully covariant van dam-veltman-zakharov discontinuity, and absence thereof. *Phys. Lett.*, B534:209–215, 2002.

[85] L. Randall and C. Csaki. The doublet-triplet splitting problem and higgses as pseudogoldstone bosons. 1995.

[86] L. Randall, V. Sanz, and M. D. Schwartz. Entropy-area relations in field theory. *JHEP*, 06:008, 2002.

[87] L. Randall and M. D. Schwartz. Quantum field theory and unification in ads5. *JHEP*, 11:003, 2001.

[88] L. Randall and M. D. Schwartz. Unification and the hierarchy from ads5. *Phys. Rev. Lett.*, 88:081801, 2002.

[89] L. Randall and R. Sundrum. A large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999.

[90] R. Rattazzi and A. Zaffaroni. Comments on the holographic picture of the randall-sundrum model. *JHEP*, 04:021, 2001.

[91] M. D. Schwartz. The emergence of localized gravity. *Phys. Lett.*, B502:223–228, 2001.

[92] M. D. Schwartz. Constructing gravitational dimensions. 2003.

[93] R. Sekhar Chivukula, D. A. Dicus, and H.-J. He. Unitarity of compactified five dimensional yang-mills theory. *Phys. Lett.*, B525:175–182, 2002.

[94] W. Siegel. Hidden gravity in open string field theory. *Phys. Rev.*, D49:4144–4153, 1994.

[95] A. Strominger and C. Vafa. Microscopic origin of the bekenstein-hawking entropy. *Phys. Lett.*, B379:99–104, 1996.

[96] A. Sugamoto. 4d gauge theory and gravity generated from 3d ones at high energy. *Prog. Theor. Phys.*, 107:793–804, 2002.

[97] L. Susskind. The world as a hologram. *J. Math. Phys.*, 36:6377-6396, 1995.

[98] L. Susskind, L. Thorlacius, and J. Uglum. The stretched horizon and black hole complementarity. *Phys. Rev.*, D48:3743-3761, 1993.

[99] L. Susskind and N. Toumbas. Wilson loops as precursors. *Phys. Rev.*, D61:044001, 2000.

[100] L. Susskind and J. Uglum. Black hole entropy in canonical quantum gravity and superstring theory. *Phys. Rev.*, D50:2700-2711, 1994.

[101] L. Susskind and E. Witten. The holographic bound in anti-de sitter space. 1998.

[102] G. 't Hooft. On the quantum structure of a black hole. *Nucl. Phys.*, B256:727, 1985.

[103] G. 't Hooft. Dimensional reduction in quantum gravity. 1993.

[104] G. 't Hooft. The holographic principle: Opening lecture. 1999.

[105] S. Thomas. Holographic vacuum energy. 2000.

[106] D. J. Toms. Quantized bulk fields in the randall-sundrum compactification model. 2000.

[107] A. I. Vainshtein. To the problem of nonvanishing gravitation mass. *Phys. Lett.*, B39:393-394, 1972.

[108] H. van Dam and M. J. G. Veltman. Massive and massless yang-mills and gravitational fields. *Nucl. Phys.*, B22:397-411, 1970.

[109] P. Van Nieuwenhuizen. On ghost-free tensor lagrangians and linearized gravitation. *Nucl. Phys.*, B60:478-492, 1973.

[110] H. Verlinde. Holography and compactification. *Nucl. Phys.*, B580:264-274, 2000.

[111] M. Visser. Dirty black holes: Entropy as a surface term. *Phys. Rev.*, D48:5697-5705, 1993.

[112] R. M. Wald. Black hole entropy in noether charge. *Phys. Rev.*. D48:3427-3431. 1993.

[113] N. Weiner. Unification without unification. 2001.

[114] E. Witten. (2+1)-dimensional gravity as an exactly soluble system. *Nucl. Phys.*. B311:46, 1988.

[115] E. Witten. Anti-de sitter space and holography. *Adv. Theor. Math. Phys.*. 2:253-291. 1998.

[116] V. I. Zakharov. *JETP Letters (Sov. Phys.)*, 12. 1970.