HIERARCHIES IN QUANTUM GRAVITY: LARGE NUMBERS, SMALL NUMBERS, AND AXIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

John Eldon Stout

August 2017

HIERARCHIES IN QUANTUM GRAVITY: LARGE NUMBERS, SMALL NUMBERS, AND AXIONS

John Eldon Stout, Ph.D.

Cornell University 2017

Our knowledge of the physical world is mediated by relatively simple, effective descriptions of complex processes. By their very nature, these effective theories obscure any phenomena outside their finite range of validity, discarding information crucial to understanding the full, quantum gravitational theory. However, we may gain enormous insight into the full theory by understanding how effective theories with extreme characteristics—for example, those which realize large-field inflation or have disparate hierarchies of scales—can be naturally realized in consistent theories of quantum gravity. The work in this dissertation focuses on understanding the quantum gravitational constraints on these "extreme" theories in well-controlled corners of string theory.

Axion monodromy provides one mechanism for realizing large-field inflation in quantum gravity. These models spontaneously break an axion's discrete shift symmetry and, assuming that the corrections induced by this breaking remain small throughout the excursion, create a long, quasi-flat direction in field space. This weakly-broken shift symmetry has been used to construct a dynamical solution to the Higgs hierarchy problem, dubbed the "relaxion." We study this relaxion mechanism and show that—without major modifications—it can not be naturally embedded within string theory. In particular, we find corrections to the relaxion potential—due to the ten-dimensional backreaction of monodromy charge—that

conflict with naive notions of technical naturalness and render the mechanism ineffective.

The super-Planckian field displacements necessary for large-field inflation may also be realized via the collective motion of many aligned axions. However, it is not clear that string theory provides the structures necessary for this to occur. We search for these structures by explicitly constructing the leading order potential for C_4 axions and computing the maximum possible field displacement in all compactifications of type IIB string theory on toric Calabi-Yau hypersurfaces with $h^{1,1} \leq 4$ in the Kreuzer-Skarke database. While none of these examples can sustain a super-Planckian displacement—the largest possible is $0.3M_{\rm pl}$—we find an alignment mechanism responsible for large displacements in random matrix models at large $h^{1,1} \gg 1$, indicating that large-field inflation may be feasible in compactifications with tens or hundreds of axions.

These results represent a modest step toward a complete understanding of large hierarchies and naturalness in quantum gravity.

BIOGRAPHICAL SKETCH

John Eldon Stout was born long ago, on April 29, 1988 in Greenville, North Carolina and grew up roaming the woods of Chapel Hill, North Carolina. At age 13, he decided that computers were more interesting than trees. A high school truant, he attended Durham Technical Community College for a year before transferring to North Carolina State University to study Electrical and Computer Engineering, with the intent of becoming a computer architect.

While initially interested, John grew bored. An REU with Prof. Michael Escuti at NC State, studying the optical properties carbon nanotube-doped liquid crystal polymers, convinced John of two things: he should not be doing experimental work, and he should be doing physics. John promptly appended a major in Physics, and after an REU at the University of Maryland with Ed Ott, Michelle Girvan, and Tom Antonsen on synchronization in networks of coupled oscillators, he was hooked.

Two years after the switch to physics, John found himself at Cornell. Unsure of whether he was interested in condensed matter or numerical relativity, John worked on the BCS-BEC crossover for a summer with André LeClair before realizing his interests actually lie in quantum gravity and string theory, completing his descent into the abstract. He then joined Liam McAllister's research group, resulting in this dissertation. John will continue his research as a joint postdoctoral fellow at the University of Amsterdam and the University of Utrecht in the Netherlands.

To my family,

thank you.

ACKNOWLEDGEMENTS

Jim Sethna once told me that we measure time logarithmically, that its passing seems to accelerate as one gets older. Sitting here—at the end of my stint at Cornell—and reflecting on my time as a graduate student, this has never felt truer. Of course, it's terrifying to think of how I'll feel in 90 or 100 years! Still, it comes as a complete shock to me that an entire six years have passed since I first arrived in Ithaca and, even more so, that I am now leaving. I will truly cherish this time, even if it is logarithmically smushed, largely due to the host of wonderful people I have had the great fortune of crossing light cones with.

I'd like to first thank Liam McAllister, who has been an exceptional mentor and advisor. Only in hindsight do I fully appreciate how much luck is involved in finding an advisor as a young graduate student. Liam's passion for physics, clarity of thought, and kindness have left an indelible mark on me as both a person and a physicist, and I feel extremely fortunate to have worked with and learned from him these past few years.

I must thank "Coach" Paul McGuirk, one of the most kind-hearted people I know. Paul provided invaluable encouragement when I was first starting out as a theorist—which coincided with a very difficult time in my life—and was incredibly generous with his time, always willing to do anything he could to help get a young researcher off the ground. It is hard to overstate the impact Paul has had on me as a researcher and as a person, and I strive to one day meet the (very high) bar he set as a postdoc.

Similarly, I owe a deep debt of gratitude to Cody Long, without whom this thesis would have a very different title and graduate school would have been a very different experience. It was Cody who, when I was unsure of what to pursue, sparked my interest in string theory and nudged me to join Liam's group. And it is

Cody who has been an excellent collaborator and—more importantly—a treasured friend, whose influence on this work is immeasurable. Thank you.

The other members of the theory group, too, have had a profound impact on me, and I'd like to thank Jack and Nic especially, my wonderful office mates Nima, Thomas, and Amir, Jeff, Math, Gowri, Sal, Simon, Mario, Wee Hao, Mike, Flip, and our wise postdocs Gautier, Gui, Marco, Yonit, Sachin, Erik, Sandipan, and Javi.

I'd particularly like to thank Tom Hartman for our numerous conversations about quantum field theory and quantum gravity—from which I learned immensely—and his guidance over the years.

I'd like to thank my longtime roommates Matt and Corky, and my shorttime roommate Andy, for keeping it weird. I can't imagine living in Ithaca without you fellas, and I'm glad I won't have to.

Thanks to Sam—my partner-in-soup—for our many lunches, our many more irreverent conversations, and for being a fantastic friend.

I'd like to thank Ethan and Brian who, along with Cody and Andy, formed a wonderful whiskey tasting group that was always a highlight of the month for much of my time at Cornell.

There are a number of people whose friendship have defined my time at Cornell. Thank you Colin, Kayla, Jorge, Ismail, Sam, Khev, Brian, Lauren, Neal, Tomas, Ben, Greg, Joel, Nicole, YJ, and Yariv. Furthermore, I'd like to thank old friends far away from Ithaca—Zach, Logan, and Matt.

I had the great pleasure of teaching and interacting with many bright undergraduates at Cornell, and I would like to thank Chad, Cari, Ven, Laura, Steve, Bryce, Brian, Ben, Artem and many others for making TAing so enjoyable and rewarding.

Penultimately, I'd like to thank my mom, dad, and sister for always doing everything in their power to help me succeed—even once they stopped understanding exactly what I was succeeding at—and for their constant love and encouragement.

Finally, I would like to thank Gimme! Coffee's cold brew, without which none of this work would have been accomplished and—physiologically—I will miss most of all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Each of the chapters contained within this thesis are focused on a common theme: understanding large and small numbers in quantum gravity. The goal of this introduction is to provide some background—and perspective—on why such numbers are interesting and what we stand to gain by understanding them.

1.1 The Unreasonable Effectiveness of Dimensional Analysis

We begin by introducing what is, perhaps, the most powerful tool in physics. Dimensional analysis is a technique for arriving at the correct approximate answer to a physical question without much computational effort (or understanding) on the physicist's part. It is best illustrated with a simple, well-used example.

Consider the rigid pendulum in a gravitational field, pictured in Figure 1.1, consisting of a massless rod of length ℓ fixed to a pivot about one end, with a bead of mass m at the other. A particularly simple parameterization of this system's dynamical degree of freedom is to keep track of the angle the rod makes with the vertical, $\theta(t)$. The equation of motion for this system is then

$$m\,\ddot{\theta}(t) + mg\ell \sin\theta(t) = 0, \tag{1.1}$$

and our childhood experiences tell us that the mass should oscillate about $\theta = 0$ on a time scale dictated by the physical parameters in the system: m, g, and ℓ. The enterprising undergraduate would quickly expand the equation of motion in small fluctuations to show that the period of oscillation is

$$T_0 \approx 2\pi\sqrt{\frac{\ell}{g}}, \tag{1.2}$$

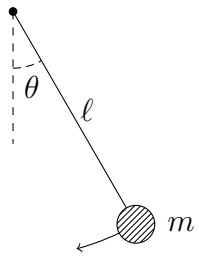

Figure 1.1: The simple pendulum in a gravitational field.

while the overeager first year graduate student might solve (1.1) in terms of Jacobi's amplitude. However, the practiced physicist is much lazier (read: efficient) and may easily guess the correct answer with minimal work.

The essence of the idea is rewrite the equation of motion (1.1) in a dimensionless form. Since θ is already dimensionless, this consists of parameterizing the angle θ in terms of a *natural time scale*, which we will call ξ, such that (1.1) reads

$$\theta''(\xi) + \sin\theta(\xi) = 0, \tag{1.3}$$

with $\xi = \sqrt{g/\ell}\, t$. We are then left with an equation featuring only dimensionless constants, with no very large or very small numbers. Importantly, (1.3) locks two quantities, θ'' and $\sin\theta$, to be of the same magnitude—indeed, in this example they must be exact opposites—so nothing in the equation can become overly large or small, lest it violate (1.3). This is a general expectation in physics: most well-behaved equations of motion have a very difficult time generating extremely large or small numbers. Why? One answer—though certainly not the only one—is that the systems we typically study are those that actually permit study, i.e. those systems which are stable enough for physicists to prod repeatedly! By definition, stable systems cannot generate absurdly large numbers when fed reasonable initial data—divergent physical quantities are a sign that the system under study is unstable. So, by restricting to stable systems, we should expect that the answer to any reasonable

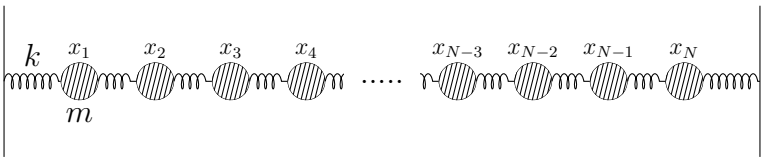

Figure 1.2: N coupled harmonic oscillators of mass m, connected by springs with equal spring constant k.

question we may ask will have a reasonable $\mathcal{O}(1)$ answer, when expressed in the natural scales of the system.

We may apply this logic to our pendulum. We expect that the time scale of oscillation is $\mathcal{O}(1)$ in ξ units, and thus $T_0 \approx \sqrt{\ell/g}$. Without actually solving anything, we may arrive at an accurate idea of how a system will behave simply be performing estimates based on dimensional analysis. It is important to note that dimensional analysis only provides an *expectation* of an answer, and has no way in general of predicting the same 2π factor in (1.2) that the precocious undergraduate would derive. However, when this expectation is not met—say, if the period of oscillation were measured to be $\sim 10^{-12}\sqrt{\ell/g}$—then we should interpret this as a signal flare, a warning that we do not yet fully understand the relevant physics of the system. Indeed, any large deviation from the natural expectation provided by dimensional analysis *demands* to be explained, and this explanation typically provides deeper insight into the nature of the system.

To illustrate this point, consider the system of N coupled harmonic oscillators pictured in Figure 1.2. Each bead has mass m and is coupled to its neighbor with a spring with spring constant k. The system is described by N equations of motion,

$$m\ddot{x}_i = kx_{i-1} - 2kx_i + kx_{i+1}, \qquad i = 1, \ldots, N \quad (1.4)$$

with $x_0 = x_{N+1} = 0$ representing the coordinates of the two fixed walls at either end of the system. This system admits N normal modes of oscillation, each which

oscillate at a frequency ω_ℓ. By dimensional analysis, we expect each of these normal modes to oscillate at a frequency roughly

$$\omega_\ell = \mathcal{O}(1) \times \sqrt{\frac{k}{m}}. \tag{1.5}$$

However, a standard textbook analysis [2] shows that the normal modes oscillate with frequency

$$\omega_\ell = 2\sqrt{\frac{k}{m}} \sin\left(\frac{\ell\pi}{2N+2}\right), \tag{1.6}$$

and so the slowest mode ($\ell = 1$) approximately has frequency

$$\omega_1 \approx \frac{\pi}{N+1}\sqrt{\frac{k}{m}}. \tag{1.7}$$

For arbitrarily large N, this is arbitrarily smaller than our expectation from dimensional analysis, so we should understand what exactly is going on.

Two things have happened here. First, following arguments similar to those we applied to (1.3), we could conclude that the equations of motion (1.4) cannot generate an extremely large or small quantity either. However, this argument relied on large dimensionless quantities being absent in (1.3) and there *is* a large dimensionless parameter in (1.4): the number of beads N. So, we should modify our dimensional analysis expectation to include this dimensionless parameter, that is

$$\omega_\ell \approx f(N) \times \sqrt{\frac{k}{m}}, \tag{1.8}$$

which is validated by the exact answer (1.6). This simple point is crucial to the story described in Chapter 2, where a similar dimensionless parameter provides large corrections to the expectations one derives from a more naive dimensional analysis.

Second, taking the limit $N \to \infty$ restores a symmetry! Adding more and more beads to the system pushes the walls further and further apart, and the system

becomes translationally invariant as $N \to \infty$. Symmetries are very powerful, in that they provide an explanation for why a particular physical quantity vanishes. Slightly broken symmetries, i.e. when N is very large but still finite, retain much of this power and provide an explanation for why a quantity is small. As we will see in the next section, symmetries are particularly useful in quantum mechanical theories, where everything allowed to happen must.

1.2 The Reasonable Effectiveness of Effective Field Theory

A quantum field theory can be defined through the generating functional of its correlation functions, which is typically represented as a path integral over field configurations, weighted by the Feynman measure. For example, for a scalar field φ_Λ in d dimensions, this generating functional takes the form

$$\mathcal{Z}[J] = \int \mathcal{D}\varphi_\Lambda \, \exp\left(\frac{i}{\hbar} S_\Lambda[\varphi_\Lambda] + \frac{i}{\hbar}\int \mathrm{d}^d x \, J(x)\varphi_\Lambda(x)\right). \tag{1.9}$$

We include the subscript Λ to remind ourselves that the path integral can be ill-defined and typically must be regulated in some way, and we will assume that this regulator is a hard momentum cutoff, such that the measure $\mathcal{D}\varphi_\Lambda$ only includes field configurations with momenta $k < |\Lambda|$. The actual physics of this field is dictated by the classical action $S_\Lambda[\varphi]$. By taking functional derivatives of $\mathcal{Z}[J]$ one can generate arbitrary φ_Λ correlation functions which are guaranteed to be consistent with the axioms of quantum mechanics and respect the symmetries of $S_\Lambda[\varphi]$. Thus, the theorist may use the path integral (1.9) as an extremely useful prescription of specifying the quantum mechanical theory of a scalar field φ_Λ with particular, imposed properties.

Crucially, the action $S_\Lambda[\varphi]$ is defined with explicit reference to the cutoff Λ.

We may write this action as the integral over a general local Lagrangian density, for example,[1]

$$S_\Lambda[\varphi_\Lambda] = \int \mathrm{d}^4x \left(-\frac{1}{2}\left(\partial\varphi_\Lambda\right)^2 - \frac{1}{2}\Lambda^2 c_{2,1}(\Lambda)\varphi_\Lambda^2 - \frac{1}{3!}\Lambda c_{3,1}(\Lambda)\varphi_\Lambda^3 - \frac{1}{4!}c_{4,1}(\Lambda)\varphi_\Lambda^4 \right.$$
$$\left. - \sum_{k\geq 6,i} \Lambda^{4-k} c_{k,i}(\Lambda)\mathcal{O}_{k,i}(\varphi_\Lambda, \partial_\mu\varphi_\Lambda, \dots) \right), \tag{1.10}$$

where $\mathcal{O}_{k,i}$ is the i'th operator composed of mass dimension k formed from φ_Λ and its derivatives, and the $c_{k,i}(\Lambda)$ are dimensionless parameters which specify which theory we intend to study.

Wilson [5, 6] realized that as long as one only asks questions below an energy scale Λ', one could integrate out the field configurations with momenta $\Lambda' \leq |k| \leq \Lambda$ to generate an *effective field theory* at the scale Λ'. By writing $\varphi_\Lambda = \varphi_{\Lambda'} + \hat{\varphi}$, we have

$$\begin{aligned}\mathcal{Z}[J_{\Lambda'}] &= \int \mathcal{D}\varphi_{\Lambda'} \left[\int \mathcal{D}\hat{\varphi} \exp\left(iS_\Lambda[\varphi_{\Lambda'} + \hat{\varphi}]\right) \right] \exp\left(i \int \mathrm{d}^4x\, J_{\Lambda'}(x)\varphi_{\Lambda'}(x) \right) \\ &= \int \mathcal{D}\varphi_{\Lambda'} \exp\left(iS_{\Lambda'}[\varphi_{\Lambda'}] + i\int \mathrm{d}^4x\, J_{\Lambda'}(x)\varphi_{\Lambda'}(x) \right),\end{aligned} \tag{1.11}$$

so that

$$\exp\left(iS_{\Lambda'}[\varphi_{\Lambda'}]\right) = \int \mathcal{D}\hat{\varphi}\, \exp\left(iS_\Lambda[\varphi_\Lambda + \hat{\varphi}]\right). \tag{1.12}$$

Because (1.10) was the most general action we could write down for this field, we may also write the effective action as

$$S_{\Lambda'}[\varphi_{\Lambda'}] = \int \mathrm{d}^4x \left(-\frac{1}{2}\left(\partial\varphi_{\Lambda'}\right)^2 - \frac{1}{2}\Lambda'^2 c_{2,1}(\Lambda')\varphi_{\Lambda'}^2 - \frac{1}{3!}\Lambda' c_{3,1}(\Lambda')\varphi_{\Lambda'}^3 - \frac{1}{4!}c_{4,1}(\Lambda')\varphi_{\Lambda'}^4 \right.$$
$$\left. - \sum_{k\geq 6,i} \Lambda'^{4-k} c_{k,i}(\Lambda')\mathcal{O}_{k,i}(\varphi_{\Lambda'}, \partial_\mu\varphi_{\Lambda'}, \dots) \right). \tag{1.13}$$

[1] In general, we should also include wavefunction renormalization factors $Z(\Lambda)$. The interested reader may find a full technical discussion in any modern text on Quantum Field Theory [3, 4]. We will take $\hbar = c = 1$ in the rest of the discussion.

The effect of integrating out the degrees of freedom between the momenta $\Lambda' \leq |k| \leq \Lambda$ can then be incorporated by changing the dimensionless coupling constants, $c_{k,i}(\Lambda')$, which are called the *Wilson parameters*. Since the Wilson parameters determine[2] how the theory behaves at the scale Λ, then we may similarly understand how the theory behaves at the scale Λ' by simply keeping track of how the Wilson parameters change and forgetting about the physics between $\Lambda' \leq |k| \leq \Lambda$. Said differently, the high-energy degrees of freedom have *decoupled*, but have modified the coupling constants $c_{k,i}(\Lambda)$ in the process.

Now, by integrating out an infinitesimal shell of momentum, $\Lambda - \epsilon \leq |k| \leq \Lambda$, we may describe the flow $c_{k,i}(\Lambda) \to c_{k,i}(\Lambda - \epsilon)$ using the *renormalization group equations*,

$$\frac{\mathrm{d}c_{k,i}(\Lambda)}{\mathrm{d}\log\Lambda} = \beta_{k,i}\Big(\{c_{l,j}(\Lambda)\}\Big). \tag{1.14}$$

Like (1.3) for the rigid pendulum, (1.14) may be interpreted as the dimensionless equations of motion for a dynamical system, and following the arguments of the previous section we would expect that under any reasonable flow and for generic initial data the Wilson coefficients cannot become extremely large or small, i.e.

$$c_{k,i}(\Lambda) \sim \mathcal{O}(1).^{3} \tag{1.15}$$

Since the Wilson coefficients $c_{k,i}$ only appear in the action (1.10) dressed with appropriate powers of the cutoff Λ^{4-k}, we expect that its contribution to a physical observable appears in powers of $(\Lambda/E)^{4-k}$. Thus, *relevant* ($k < 4$) and *marginal* ($k = 4$) operators have a much greater impact on low energy observables than

[2]Much of the wording here intrinsically assumes that the theory is weakly-coupled at both Λ and Λ'.

[3]This logic implicitly assumes that the theory remains weakly-coupled and that the degrees of freedom being integrated out have a characteristic mass scale roughly $\sim \Lambda$, so no large or small ratios $(E/\Lambda)^{\ell}$ can drastically alter this expectation.

irrelevant ($k \geq 4$) operators.

This is the wonderful power of effective field theory: at low energies, one only needs to get a small number of (aptly-titled) relevant parameters right. The irrelevant operators in the action (1.10) have a negligible effect on low energy physics! This *universality* provides physics with enormous predictive power: many systems with drastically different microphysics can have the same low energy behavior.

Unfortunately, the high-energy theorist is interested in the inverse problem: can one determine the original microscopic theory from the effective theory at low energies? Here, universality works against us. Inherent experimental error prevents the low-energy physicist from determining the irrelevant Wilson coefficients to any accuracy, forcing the high-energy theorist to solve the inverse problem with incomplete information. Said differently, because universality predicts that a large class of microphysical systems have the *same* low-energy effective description, the high-energy theorist cannot solve the inverse problem without additional tools or constraints.

1.3 Naturalness and Effective Field Theory

As demonstrated in §1.1, we can use general expectations about the nature and behavior of a physical system to arrive at, or at least point towards, a solution to this inverse problem. We rely on the assumption that the system under study is "typical," and we argued in the previous section that typical effective field theories have $\mathcal{O}(1)$ Wilson coefficients. That is, effective theories which satisfy (1.15) are *natural* or generic. We may use this presumption of *naturalness* as an additional constraint in our attempts to divine the true microphysical theory from its low-

energy effective description.

A classic example of utility of naturalness is the prediction of the positron. Let us model [7, 8] the electron as a smooth charge distribution with characteristic size $r_{\rm e}$. We would expect that the mass of the electron is at least of the order of the energy that it carries in the electromagnetic field it produces. Using $r_{\rm e}$ as a cutoff, we expect that

$$m_{\rm e} \sim \frac{1}{8\pi} \int {\rm d}^3x \, |\mathbf{E}|^2 \sim \frac{\alpha}{r_{\rm e}}, \tag{1.16}$$

where $\alpha = e^2/4\pi \sim 1/137$ is the fine-structure constant. However, modern day experiment has shown that $r_{\rm e} < 10^{-17}$ cm, implying that the electron mass should naturally be of the order $m_{\rm e} \sim 100$ GeV, not $m_{\rm e} \approx 0.511$ MeV.

Absent an explanation, we would argue that the measured electron mass is *unnatural* and that there must be a delicate fine tuning between the bare mass of the electron and the correction due to quantum effects,

$$\underset{\sim 1 \text{ MeV}}{m_{\rm e}} = \underset{\sim 10^5 \text{ MeV}}{m_{\rm e}^{(0)}} + \underset{\sim 10^5 \text{ MeV}}{\alpha / r_{\rm e}}, \tag{1.17}$$

for the electron mass to be ≈ 0.511 MeV rather than ~ 100 GeV, roughly a cancellation to one part in 10^5. While such a fine tuning is allowed, it should be interpreted as a signal that there may be additional physics—i.e. new degrees of freedom which we do not yet thoroughly understand—that render the small parameter $m_{\rm e}$ natural. We expect that these new degrees of freedom show up at the scale dictated by naturalness, in this case at an energy scale $\Lambda \sim 1$ MeV.

Indeed, the introduction of the positron (with mass $m_{\rm p} = m_{\rm e}$) alters the form of the quantum corrections,

$$m_{\rm e} = m_{\rm e}^{(0)} \left(1 - \frac{6\alpha}{4\pi} \log \left(m_{\rm e} r_{\rm e} \right) \right), \tag{1.18}$$

and renders $m_{\mathrm{e}} \approx 0.511\,\mathrm{MeV}$ natural. Much like the $N \to \infty$ limit restored translational invariance in our coupled oscillators example, the introduction of the positron restores a chiral symmetry to the theory of quantum electrodynamics. While this symmetry is broken by the presence of the electron and positron mass, it still provides an *explanation* for why this mass is much smaller than the naive expectation $m_{\mathrm{e}} \sim \alpha r_{\mathrm{e}}^{-1}$. A small parameter is said to be *technically natural* if a symmetry is restored when that parameter is tuned to vanish.

Naturalness has proved to be a very effective tool for physicists and has been used to "predict" a variety of new particles and phenomena. Indeed, this expectation that our physical theories are not *fine-tuned* has become so strong that apparent violations of naturalness represent major problems—deemed *hierarchy problems*—in theoretical physics. Solutions to the *Higgs hierarchy problem*, the *strong CP problem*, and the *cosmological constant problem* attempt to explain why the Higgs mass, QCD θ-angle, and cosmological constant are 10^{17}, 10^{13}, and 10^{120} times smaller than our expectations, respectively. Such solutions provide microphysical mechanisms which explain why these quantities are so different from their natural values, and would provide deep insights into the structure of our world at high energies.

1.4 Naturalness and Quantum Gravity

In the previous sections, we saw two examples—the system of N coupled oscillators and the electron mass in quantum electrodynamics—where the unnatural smallness of a physical quantity could be explained by a global or spacetime symmetry. This is a common theme in solutions to hierarchy problems: one typically declares

victory[4] if a small parameter can be tied to an approximate symmetry of the theory. However, it is widely believed folklore [9, 10, 11] that continuous global symmetries cannot exist in a consistent quantum gravitational theory. Instead, the only exact continuous symmetries that can exist in quantum gravity are those which are gauged: quantum gravity permits only constraints, not symmetries.

As with most deep facts about quantum gravity, we can argue against the existence of global symmetries by considering the properties of evaporating black holes. Following [12], let us consider a quantum gravitational theory with heavy matter charged under an exact U(1) global symmetry. We may form a black hole of mass M with a large global charge Q and, because black holes have no hair, an observer sitting outside the black hole cannot discern that it is charged. Thus, the black hole will evaporate thermally, losing most of its mass to light, uncharged quanta until the black hole becomes hot enough to emit the heavier charged matter. Because the U(1) symmetry is exact, the charge cannot simply vanish. So, unless the theory contains matter with arbitrarily high charge-to-mass ratios, the black hole will not be able to get rid of all of its charge and will not be able to evaporate entirely. The end state of this evaporation process thus leads to a highly-charged stable object called a remnant. These remnants contain huge amounts of entropy, and necessarily [13] drive the gravitational coupling to 0, leaving us with a non-gravitational quantum theory. So, the existence of continuous global symmetries in quantum gravitational theories lead to unacceptable pathologies, and we expect that any such symmetries are necessarily broken by gravitational effects. This is most clearly demonstrated in [9], which shows that, in a simple system with a global symmetry, wormholes can eat global charge by carrying it away from our universe.

[4] "Victory" is alternatively written as "technically natural."

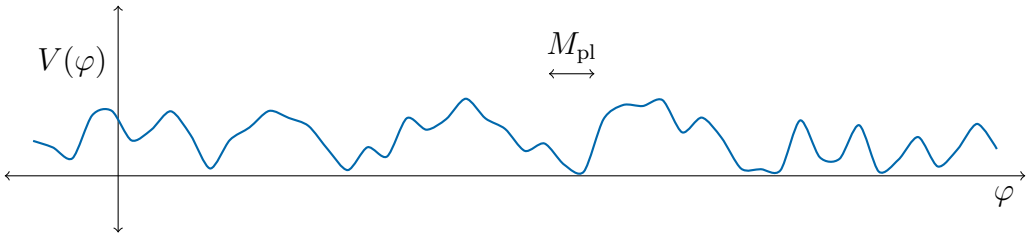

Figure 1.3: A typical potential (1.19) with $c_k \sim \mathcal{O}(1)$ Wilson coefficients.

That global symmetries are absent in theories of quantum gravity strips us of a particularly powerful tool for explaining unnaturally large or small quantities in these theories, especially if the large hierarchy we are interested in is intrinsically tied to gravitational physics as, for example, it is in the cosmological constant problem and in large field inflation.

Much of this thesis is devoted, in one way or another, to realizing large field inflation in quantum gravity and, in particular, string theory. Simple models of inflation involve a single scalar field φ—the inflaton—undergoing slow, controlled evolution in a relatively flat potential. In large field inflation, the inflaton executes a path in field space with super-Planckian arc-length. From our discussion above and in §1.2, we expect that the inflaton potential is naturally of the form

$$V(\varphi) = M_{\rm pl}^4 \sum_{k=1}^{\infty} c_k \left(\frac{\varphi}{M_{\rm pl}}\right)^k \tag{1.19}$$

with the Wilson coefficients $c_k \sim \mathcal{O}(1)$, and schematically looks like Figure 1.3. Requiring that this potential be flat over super-Planckian $\Delta\varphi \gg M_{\rm pl}$—that is, requiring that it look like Figure 1.4—amounts to ensuring that all Wilson coefficients are unnaturally small, $c_k \ll 1$! Large field inflation apparently requires a function's worth of fine-tuning.

Excitingly, large field inflation is tied to a detectably bright CMB B-mode signal which will be definitively discovered or ruled out within the next decade [14]. Thus,

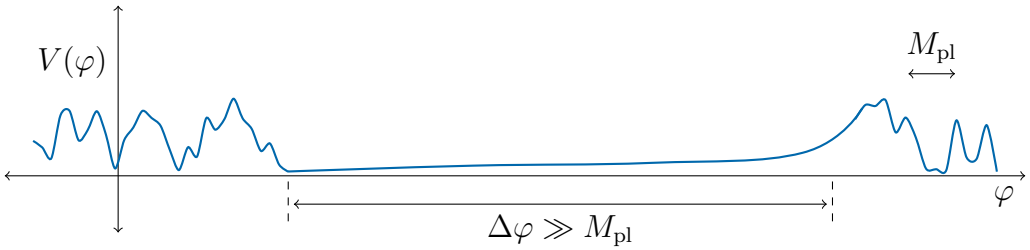

Figure 1.4: A potential (1.19) which admits large field inflation and requires a delicate fine-tuning of the Wilson coefficients c_i.

while understanding the role of naturalness in quantum gravity—and whether or not there exist robust mechanisms for generation of large hierarchies—promises to yield deep insights into the structure of the theory, it may also provide the pathway to the first experimental tests of quantum gravity.

1.5 Outline of Thesis

As discussed in the previous section, theories of quantum gravity do not allow continuous global symmetries. This is very unfortunate from a computational perspective, as symmetries can drastically reduce the complexity of a given problem. Computational control is a luxury when studying quantum gravity, so one must be certain to work under a "lamp post," i.e. in a particularly well-controlled regime or sector of the theory. Most of this thesis is thus concerned with the physics of *axions* in string theory—specifically axions in Type IIB string theory compactified on a Calabi Yau orientifold.

Discussed in detail in §2.A, axions arise from p-form gauge fields dimensionally reduced along non-trivial p-cycles in an internal manifold. The p-form gauge symmetry protects the axion from receiving potentially dangerous quantum cor-

rections. Indeed, the axion a enjoys a continuous shift symmetry $a \mapsto a + \text{const.}$ at the perturbative level. This continuous symmetry is broken to a discrete shift symmetry $a \mapsto a + f$ by nonperturbative effects, with f the axion decay constant. The discrete shift symmetry is an example of a 0-form gauge symmetry. Axionic theories are thus highly constrained and may serve as a well-controlled arena to understand these large hierarchies in quantum gravity.

The work in Chapter 2 focuses on the relaxion mechanism in string theory. The relaxion is a beautiful attempt to provide a dynamical resolution to the Higgs hierarchy problem, in which the Higgs mass dynamically relaxes to its measured, unnaturally small value. The mechanism relies on *axion monodromy*, the spontaneous breaking an axion's discrete shift symmetry, and ties the final value of the Higgs mass to this (seemingly) technically natural small quantity. However, we argue that this scenario is inconsistent with many general properties of quantum gravity, and illustrate these inconsistencies in a well-controlled string compactification, allowing us to make fairly general statements about the role of technical naturalness in axion monodromy.

Theories with multiple axions also present promising candidates for realizing large field ranges in string theory. While general reasoning prevents each axion from having a super-Planckian decay constant (and thus super-Planckian field range), multiple axions may undergo a collective motion which realizes an effective super-Planckian displacement. It is thus important to understand if there are general structures in string theory (and, more generally, in quantum gravity) that prevent such a collective motion. In Chapter 3, we initiate a systematic study of these theories in a well-controlled corner of string theory. Specifically, we compute the largest possible field range for every Calabi Yau threefold with $h^{1,1} \leq 4$ in the

Kreuzer-Skarke database, with the aim of understanding the extent to which the field range can be enhanced by this collective motion.

As we saw in §1.2, tension with naturalness can be resolved by the introduction of new degrees of freedom. One potential resolution to the Higgs hierarchy problem is to realize the Standard Model fields as composite—instead of fundamental—degrees of freedom. At low energies, the Standard Model degrees of freedom behave like particles, while at higher energies they act like a bound state of more fundamental, necessarily strongly coupled, degrees of freedom. AdS/CFT provides a powerful tool to investigate strongly-coupled dynamics, and Chapter 4 utilizes this tool to study a toy version of these composite models realized in string theory. Specifically, we determine the spectrum of chiral and non-chiral bifundamental mesons on stacks of intersecting D7-branes in $AdS_5 \times S^5$.

CHAPTER 2

RUNAWAY RELAXION MONODROMY

Abstract[1]

We examine the relaxion mechanism in string theory. An essential feature is that an axion winds over $N \gg 1$ fundamental periods. In string theory realizations via axion monodromy, this winding number corresponds to a physical charge carried by branes or fluxes. We show that this monodromy charge backreacts on the compact space, ruining the structure of the relaxion action. In particular, the barriers generated by strong gauge dynamics have height $\propto e^{-N}$, so the relaxion does not stop when the Higgs acquires a vev. Backreaction of monodromy charge can therefore spoil the relaxion mechanism. We comment on the limitations of technical naturalness arguments in this context.

[1]This chapter is based on L. McAllister, P. Schwaller, G. Servant, J. Stout and A. Westphal, "Runaway Relaxion Monodromy," 1610.05320.

We thank Nima Afkhami-Jeddi, Tom Hartman, Nemanja Kaloper, David E. Kaplan, Eric Kuflik, Cody Long, Miguel Montero, Surjeet Rajendran, Michael Stillman, Amir Tajdini, Irene Valenzuela, and Timo Weigand for valuable discussions.

2.1 Introduction

Why is the Higgs mass so small? Graham, Kaplan, and Rajendran (GKR) have proposed a novel solution to the electroweak hierarchy problem, the *relaxion mechanism*, in which the evolution of an axion field ϕ drives the Higgs mass m_h to relax dynamically to a value much smaller than the cutoff, $|m_h^2| \ll M^2$ [16]. Achieving a large hierarchy in this way requires very small dimensionless couplings, as well as field excursions $\Delta\phi \gg M$, but GKR argued that the requisite couplings are technically natural.

In this work, we study the impact of ultraviolet completion on the relaxion mechanism. The large field excursions required by the mechanism, while technically natural in effective field theory, turn out to be *source terms* in string theory! Winding an axion ϕ over $N \gg 1$ fundamental periods leads to the accumulation of N units of monodromy charge, providing a large source term in ten dimensions. This changes the shape of the compactification and alters the couplings of the effective theory, eliminating the barrier that is needed to stop the relaxion once the Higgs acquires a vev.

The root of the problem is that new states linked to the monodromy charge, which are too massive in the initial configuration to be visible, are eventually drawn below the cutoff M. These new light states induce changes in the couplings of the effective theory. In particular, the gauge coupling $g_{\rm YM}$ of the gauge theory that generates the stopping potential receives a correction $\delta g_{\rm YM}^{-2} \sim N$. This leads to an exponential suppression of the stopping potential, with barrier heights $\sim e^{-N}$, and therefore to a runaway relaxion. This problem persists even in the limit in which the relaxion shift symmetry appears to be restored.

Although we work in string theory, and quantum gravity completion is the central question, our results do not hinge on super-Planckian displacements $\Delta\phi \gg M_{\rm pl}$, which are famously challenging in quantum gravity. The problems that we expose occur even for $\Delta\phi \ll M_{\rm pl}$. The core issue is indeed one of large displacements, but here large means compared to the natural scale (or periodicity) of the effective theory. When ϕ is an axion with decay constant f, the backreaction of monodromy charge is significant for $\Delta\phi \gg f$.

Our analysis does not amount to a complaint that the effective theories given in [16] contain small dimensionless parameters. Constructing a solution of string theory that yields an effective field theory containing small numbers plausibly requires fine-tuning, e.g. of the discrete data of a compactification. Quantifying this obvious issue is not our aim. The backreaction phenomenon that we identify is a much more severe problem: even granting fine-tuned data that gives rise to an apparently-suitable relaxion Lagrangian in the probe approximation that omits the monodromy charge as a source in ten dimensions, the full Lagrangian beyond the probe approximation is not of the form given in [16], and does not allow for relaxation of a hierarchy.

Our goal is to identify the challenges that confront the relaxion mechanism in string theory. Though we analyze a specific realization in type IIB string theory, we find a set of surprising, plausibly general, qualitative lessons about the nature of hierarchies and technical naturalness in low energy effective field theories descending from string theory.

The remainder of §2.1 is a microcosm of the paper. We begin with a review of the relaxion mechanism and then provide an overview of our results, leaving detailed analysis for the main text. The casual reader need only read §2.1.

Overview of the Relaxion

The simplest model of electroweak scale relaxation involves adding to the Standard Model a single axion ϕ, the relaxion, with the potential[2]

$$V(\phi, h) = \underbrace{\Big(M^2 - gM\,(\phi_{\rm init} - \phi)\Big)}_{\text{Ⓐ}}|h|^2 + \underbrace{gM^3\phi}_{\text{Ⓑ}} + \underbrace{V_{\rm stop}(\phi, v)}_{\text{Ⓒ}}. \tag{2.1}$$

Here h is the Higgs field and v is its vacuum expectation value, $v^2 \equiv \langle |h|^2 \rangle$, M is the cutoff of the effective field theory, and g is a dimensionless parameter that controls the explicit (albeit weak) complete breaking of the relaxion's perturbatively exact continuous shift symmetry $\phi \mapsto \phi + \text{const}$. The coupling Ⓐ promotes the Higgs mass m_h^2 to a dynamical variable, so that evolution of ϕ scans over a range of Higgs masses, while Ⓑ is a potential that forces ϕ to smaller values, $\phi_{\rm final} \ll \phi_{\rm init}$. Finally, Ⓒ is a non-perturbatively generated, oscillatory "stopping potential" $V_{\rm stop}(\phi, v) = V_{\rm stop}(\phi + f, v)$, whose height grows with the Higgs vev v. For now, we take this potential to be

$$V_{\rm stop}(\phi, v) = \Lambda_c^3\, v \cos\left(\frac{2\pi\phi}{f}\right) \tag{2.2}$$

with Λ_c the confinement scale of a gauge theory G to which ϕ has an axionic coupling, though we will consider more general potentials in §2.2. This potential is generated by strong gauge dynamics and disappears when the theory is in a phase with unbroken chiral symmetry, i.e. in a phase with massless quarks. Thus, the stopping potential vanishes unless the Higgs has developed a vev.

The mechanism is illustrated in Figure 2.1. The relaxion starts at a large value $\phi_{\rm init}$, where $m_{\rm h}^2 \sim M^2$, and begins to slowly roll down the linear potential Ⓑ. For

[2] We follow the same notation as [16], except that we take the coupling g to be dimensionless, $g_{\rm GKR} = g \times M$, and shift the origin of the relaxion ϕ field space.

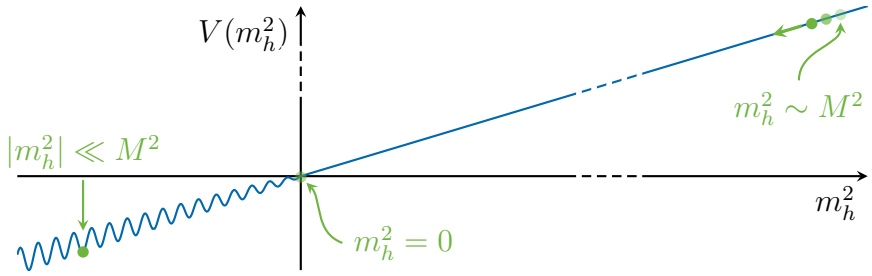

Figure 2.1: Schematic plot of the relaxion potential (2.1).

generic initial conditions, the relaxion will roll a distance

$$\Delta\phi \sim M/g \tag{2.3}$$

in field space before the Higgs becomes massless, Ⓐ $= m_h^2 = 0$. The Higgs then develops a vev and the stopping potential is generated. The relaxion continues to roll, halting once the stopping potential grows strong enough to counterbalance the linear potential—roughly when

$$\frac{v}{M} \sim g \left(\frac{M}{\Lambda_c}\right)^3 \frac{f}{M}. \tag{2.4}$$

The hierarchy between the Higgs vev and the cutoff of the theory is thus controlled by the shift-symmetry breaking parameter g. In effective field theory, it is technically natural for g to be arbitrarily small. However, we will see that there are obstacles to such a structure in string theory.

Requirements for Relaxation

We now summarize the necessary ingredients for a successful relaxation of the electroweak scale.

1. The Higgs mass must be made dynamical by introducing an axion[3] field ϕ

[3]As is clear from the name, it is important the relaxion ϕ be an axion: the axionic shift

with a coupling to the Higgs of the form

$$\mathcal{L}_h \supset \mathcal{G}(\phi)|h|^2 \tag{2.5}$$

where $\mathcal{G}(\phi)$ is some polynomial in ϕ. Evolution in ϕ scans over Higgs masses.

2. The dynamics of ϕ must be attractive, with the late-time (when $m_h^2 \sim 0$) behavior of ϕ being independent of the initial conditions.

3. ϕ must stop when the Higgs mass is approximately its observed, unnatural value.

For the evolution of the relaxion to be both attractive and dominated by classical dynamics, some friction is necessary. Therefore, the relaxion scenario has been assumed to take place during inflation (for an alternative source of friction from particle production see [17]). In this paper we will not discuss the underlying model of inflation (e.g. see [18, 19, 20]), nor its possible realization in string theory; these issues are the subject of an extensive literature (see for example [8]). We assume inflation to be operative, and concentrate instead on the relaxion potential and examine how it may arise in string theory constructions.

Typically, the stopping potential is generated by non-perturbative effects and is f-periodic. This ensures that only Ⓐ and Ⓑ explicitly break the discrete shift symmetry $\phi \mapsto \phi + f$, and protects against possibly disastrous corrections. The height of the stopping potential must depend on the Higgs vev. Furthermore, we require the minima of (2.1) to scan through Higgs masses finely enough so that a small overshoot does not dramatically increase the final electroweak scale; since

symmetry protects the potential against undesirable corrections. One could envision a more general relaxation scenario involving a field ϕ that is not an axion, but it would then be necessary to explain how the structures in (2.1) could be technically natural.

the stopping potential minima are spaced roughly $\Delta\phi \sim f$ apart, this translates into the requirement that $\mathcal{G}'(\phi) f \ll v^2$.

While this appears to be a beautiful solution to the Higgs hierarchy problem, there is some cause for concern: g must be an exceptionally small number in order to generate a sizable hierarchy. The simplest model of [16] requires $g \sim 10^{-28}$; see §2.2 for the requirements in variants of the model. It is reasonable to ask whether the associated large number $1/g$ infects any other terms in the effective action.

Note that although we used the same g in Ⓐ and Ⓑ, these two terms could in principle be different. Let us temporarily distinguish them and denote the coupling in Ⓐ as g_h. If $g \ll g_h$ at tree level, Higgs loops will drive the coupling in Ⓑ to be of order g_h so that the two couplings in Ⓐ and Ⓑ are not very different. If, on the other hand, we take $g \gg g_h$ at tree level, this hierarchy is stable but the required field excursion in (2.3) increases to $M/g_h \gg M/g$. So, models with $g_h \sim g$ undergo the smallest field excursion, and for this reason we only consider one g coupling in Eq. (2.1).

Although all the phenomena that we will uncover in this work can be encoded in an effective field theory, appropriately extended to include the effects of states that enter the spectrum as the relaxion makes its long excursion, these effects are not easily seen without the perspective of an ultraviolet theory. This is to say that the technical naturalness reasoning of [16] amounts to a set of premises about the field content and interactions of an effective theory, together with conclusions that indeed follow from those premises. In this work we question these premises, asking whether string theory imposes restrictions or refinements on the possible effective theories. We first critically examine technical naturalness arguments in this context and then turn to a string theory embedding of the relaxion.

Technical Naturalness and Large Displacements

Technical naturalness is often used as a panacea in model building: one begins with a symmetry that protects against potentially disastrous quantum corrections and then weakly breaks it, confident that all corrections induced by this breaking are necessarily small. If g is a dimensionless parameter measuring the weak symmetry breaking, and the symmetry is restored for $g \to 0$, corrections in the effective theory are proportional to positive powers of g, and so are well-controlled for $g \ll 1$. This logic must be used with care in the presence of field excursions $\Delta\phi$ that are large compared to the effective theory's cutoff M. The essential problem is that $\Delta\phi/M$ provides a new large parameter and corrections can depend both on g and on $\Delta\phi/M$.

As a toy example, consider a four-dimensional effective theory for a scalar field ϕ with Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial\phi)^2 - M^4 \sum_{i=1}^{\infty} c_i \left(\frac{\phi}{M}\right)^{d_i} g^{e_i}\,, \tag{2.6}$$

where M is a physical ultraviolet cutoff (the scale of some new physics), $g \ll 1$ is a dimensionless parameter, the c_i are dimensionless Wilson coefficients, and the d_i and e_i are non-negative numbers. As long as

$$e_i \neq 0 \qquad \forall i, \tag{2.7}$$

all quantum corrections are proportional to powers of g, and the continuous shift symmetry $\phi \mapsto \phi + \text{const.}$ is restored in the limit $g \to 0$. However, we stress that (2.7) must be checked for every term in (2.6), as any $e_i = 0$ term, no matter how irrelevant, could potentially provide disastrous corrections.

At large displacements $\phi \gg M$, the condition (2.7) is far from sufficient to ensure that quantum corrections are under control at small but finite g. The theory

contains a new large parameter, ϕ/M, and corrections proportional to $g^{e_i}(\phi/M)^{d_i}$ are not necessarily small for $g \ll 1$ and $\phi/M \gg 1$. Ensuring that the corrections to the classical equations of motion are small requires knowledge about the *entire sequence* $\{c_i, d_i, e_i\}$, and so the full Lagrangian (2.6).

In systems allowing axion monodromy, there is an additional subtlety: the limit $g \to 0$ is not smooth,[4] because the field space *discontinuously* changes from a helix (for $g \neq 0$) to a circle (for $g = 0$). Standard technical naturalness arguments that rely on the $g \to 0$ limit can therefore become problematic.

Now suppose one obtains an effective theory from the top down, beginning in a vacuum of quantum gravity and integrating out Planck-scale degrees of freedom, for example by performing dimensional reduction in a string compactification with stabilized moduli. Then the low-energy theory in four dimensions could still take the form (2.6), but with two important caveats. First, the exponents d_i, e_i are dictated by the vacuum configuration of the underlying theory, and the condition (2.7) must be established rather than assumed. Second, in configurations with ≤ 4 supercharges in four dimensions, in practice one never obtains complete information about the infinite sum in (2.6): some terms can be computed in different approximations, but other terms remain incalculable, although they are in principle determined by the underlying vacuum.

Because we do not have the ability to compute every term in (2.6) in *any* halfway-realistic solution of string theory, it is difficult to prove that (2.7) is possible in quantum gravity. As a result, there is a disjunction between bottom-up reasoning based on technical naturalness, and top-down reasoning based on ob-

[4] This observation led the authors of [21] to argue that the $g \to 0$ limit is not technically natural.

taining effective theories from quantum gravity: the former strictly requires the condition (2.7), which appears not to be provable in quantum gravity.

In our view, the difficulty in establishing (2.7) in any particular solution of string theory is not just that the computation is challenging; it is that plausible general reasoning about black hole thermodynamics in quantum gravity suggests that (2.7) is in fact false. Exact continuous global internal symmetries are thought by many to be impossible in quantum gravity and have not appeared in string theory to date. We therefore expect quantum gravity to dramatically affect the $g \to 0$ limit. Although our results will turn out to be compatible with this general expectation, we do *not* rely on bottom-up reasoning about quantum gravity at any point in our analysis. In particular, we do not assume any form of the Weak Gravity Conjecture (WGC).[5]

We will argue that axion monodromy in string theory is very generally characterized by the existence of one or more terms in the effective action (2.6) with $e_i = 0$, and the theory is poorly-controlled in the limit $g \to 0$, $\phi/M \to \infty$. The physical origin of these problematic terms is *backreaction by monodromy charge*, as we now explain.

New States from Monodromy

In a viable relaxion theory, we must find that every shift symmetry breaking term in the relaxion Lagrangian is proportional to a power of g, the parameter that controls the weak breaking in (2.1). However, as we will explain qualitatively now

[5]For work applying the WGC to the relaxion, see e.g. [22, 23].

and quantitatively in §2.4, the monodromy charge

$$N \equiv \frac{\Delta\phi}{f}\,, \tag{2.8}$$

leads to corrections that are *not* dressed by powers of g, so that (2.7) does not hold. We will begin with an example and then draw more general lessons.

Suppose (cf. the detailed discussion in §2.3.3) that the relaxion is associated to a two-cycle wrapped by an NS5-brane. Further, suppose that the stopping potential arises from the dynamics of a strongly-coupled non-Abelian gauge theory, with group G, living on a stack of D7-branes wrapping a four-cycle Σ_4. The height of the stopping potential depends on the coupling $g_{\rm YM}$ of this D7-brane gauge theory:

$$|V_{\rm stop}| \propto \Lambda_c^3 \propto \exp\left(-\frac{8\pi^2}{g_{\rm YM}^2\, c_G}\right)\,. \tag{2.9}$$

Here the constant c_G is determined by the type of non-perturbative effects that generate $V_{\rm stop}$, and may be set to unity for our purposes. The gauge coupling function of G is proportional to the warped volume of Σ_4, cf. (2.71):

$$\frac{1}{g_{\rm YM}^2} = \frac{{\rm vol}_{\rm W}(\Sigma_4)}{2\pi\ell_s^4}\,. \tag{2.10}$$

When the system is wound up over N cycles, N units of monodromy charge—which in this scenario is D3-brane charge—accumulate on the NS5-brane. This charge is a source in the ten-dimensional Einstein equations, and so leads to changes in the metric of the internal space and the warp factor. The backreaction thus alters the warped volume ${\rm vol}_{\rm W}(\Sigma_4)$. Then, through (2.10), the gauge coupling function—and hence the height of the barriers—depend on N. In §2.4.2, we will show that

$$\delta\left(\frac{8\pi^2}{g_{\rm YM}^2}\right) \sim N, \tag{2.11}$$

without dependence on g.

The correction (2.11) can be understood in a dual description as resulting from new light states associated to the source of monodromy. The one-loop $\overline{\text{MS}}$ β-function in a Yang-Mills theory with n_{F} fermions, n_{S} complex scalars, and coupling constant g_{YM} can be written

$$\frac{\mathrm{d}}{\mathrm{d}\log\mu}\left(\frac{8\pi^2}{g_{\text{YM}}^2}\right) = \frac{11}{3}T(\text{Ad}) - \frac{2}{3}\sum_{i=1}^{n_{\text{F}}} T(\text{R}_i) - \frac{1}{3}\sum_{a=1}^{n_{\text{S}}} T(\text{R}_a), \tag{2.12}$$

where $T(\text{R}_i)$ is the index of representation R_i and Ad denotes the adjoint representation. The introduction of N light states will typically lead to a change

$$\delta\left(\frac{8\pi^2}{g_{\text{YM}}^2}\right) = \gamma_{\text{br}} N\,, \tag{2.13}$$

with γ_{br} a constant independent of N.

Where do these new light states come from? The N units of D3-brane charge in the NS5-brane can be viewed as resulting from N actual D3-branes (up to a binding energy that does not affect our argument). So there are N new states in the theory, corresponding to strings stretching from the D7-brane stack, where the gauge theory lives, to the D3-branes. These states transform in the fundamental of G, and so may be described as N species of quarks from the viewpoint of G. Including these species in loops leads to (2.13).

The lesson is that $\mathcal{O}(N)$ new states associated with the source of monodromy—in our examples, fundamental strings stretching from the source of monodromy to the gauge theory D-branes—can give large loop corrections. These states could easily be missed in field theory, but in a string theory configuration with two D-brane gauge theories G_1, G_2, the presence of bifundamentals is hard to avoid. The only question is whether the bifundamentals are so massive that they are physically unimportant. In our setting, we will find (cf. Appendix 2.B) that arbitrarily short—and hence, light—bifundamental strings are present.

The fact that for each unit of monodromy charge there is a new state coming down in mass that contributes to the gauge coupling of the effective theory—even though this state was far above the cutoff in the vacuum at zero winding—is a consequence of the structure of the ultraviolet completion. The new states described above arise from stretched strings, and so obviously have their origin in string theory *per se*, but there are also new states that arise simply from the presence of extra dimensions: these are Kaluza-Klein (KK) states made light by monodromy. Thus, our considerations can be extended to extra-dimensional "partial" ultraviolet completions of four-dimensional field theories, without invoking string theory.

Perhaps the simplest illustration of this phenomenon is the model of [24] (see also [25]), which describes axion monodromy arising from a Stueckelberg massive U(1) gauge field coupled to a massless charged scalar field in a five-dimensional spacetime with the extra dimension compactified on a circle,

$$S_{5\mathrm{D}} = \int \mathrm{d}^4x \int_{\mathrm{S}^1} \mathrm{d}y \, \sqrt{-g} \left(-\frac{1}{4} F_{MN} F^{MN} - \frac{1}{2} m^2 \mathcal{A}_M \mathcal{A}^M - \left(D_M \Phi\right)^\dagger \left(D^M \Phi\right) \right) , \tag{2.14}$$

where $D_m = \partial_M - iqA_M$, $F_{MN} = \partial_{[M} A_{N]} = \partial_{[M} \mathcal{A}_{N]}$, and $\mathcal{A}_M = A_M - ie^{i\theta} \partial_M e^{-i\theta}$ denotes the Stueckelberg covariant U(1) gauge field. Now we perform a KK reduction on the circle, whose circumference we denote by $2\pi R$. We decompose the five-dimensional fields into an infinite series of discrete Fourier (KK) modes on the circle, and focus on the KK modes of the scalar Φ,

$$\Phi(x^\mu, y) = \frac{1}{\sqrt{2\pi R}} \sum_{n \in \mathbb{Z}} \Phi_n(x^\mu) \exp\left(\frac{iny}{R}\right) \tag{2.15}$$

This yields the effective four-dimensional action

$$S_{5\mathrm{D}} \supset \int \mathrm{d}^4x \left(\frac{1}{2} m^2 \phi^2 + \sum_{n \in \mathbb{Z}} \left(\frac{n}{R} - q\phi \right)^2 |\Phi_n|^2 \right) . \tag{2.16}$$

Here, $\phi \sim A_5^{(0)}$ denotes the four-dimensional axion field corresponding to the five-

dimensional gauge field Wilson line around the S^1. The axion ϕ evidently experiences monodromy, acquiring a quadratic potential.

The key observation is that the masses of the KK modes Φ_n,

$$m_n^2 = \left(\frac{n}{R} - q\phi\right)^2 \tag{2.17}$$

depend on the vev of the axion ϕ. As ϕ scans across its field space, one KK mode after another falls below the cutoff R^{-1} in mass and thus enters the spectrum of the low-energy effective theory. In particular, as ϕ moves over N units of its fundamental domain, N KK modes fall below the cutoff R^{-1}, in analogy with the string theory effect discussed above.

We should clarify that in our examples, monodromy affects mass spectra in two very different ways. One effect is *shifting*, in which $\phi \mapsto \phi + f$ leaves the set of masses m in a sector invariant, but permutes the states associated with these masses. For example, in (2.17), changing $\phi \mapsto \phi + (qR)^{-1}$ increases by one unit the Kaluza-Klein charge of the state at each mass level.

The other effect is *compression*, in which a monodromy $\phi \mapsto \phi + f$ changes the mass spectrum. Typically, as the axion winds up and stores more energy, the masses in affected sectors are reduced. A shifting spectrum is compatible with an exact discrete shift symmetry of the theory; the number of states below a fixed cutoff does not change, but the labels of the states change. Compression violates even a discrete shift symmetry, as the number of states below a fixed cutoff depends on ϕ.

With this terminology, we remark that the five-dimensional example above displays only shifting, not compression. This is a consequence of the oversimplified nature of the model. We will show below that axion monodromy also causes com-

pression of the mass spectrum of Kaluza-Klein excitations of an NS5-brane. Thus, stretched string states are not the only states that experience compression, and we expect that compressed spectra can arise in purely extra-dimensional scenarios without string theory.

Why do we not provide a purely four-dimensional field theory toy model showing the effects of shifting and compression, for instance in the case of axion monodromy from a four-form field strength [25, 26, 27, 28]? The issue is that although the core mechanism of axion monodromy arising via the Stueckelberg mechanism can be described in four-dimensional field theory, the results of [25] make it clear that backreaction effects, including those of massive states entering the spectrum, are described by higher-derivative corrections arising from higher powers of the four-form field strength. These corrections must be determined in the ultraviolet completion of gravity, as explicitly noted in [25] as well. That is, the two-derivative, four-dimensional field theory Kaloper-Sorbo model of axion monodromy [26] *is not* a magic wand that suppresses or controls backreaction effects.

Exponential Suppression of the Stopping Potential

We have argued that backreaction by N units of monodromy charge leads to a large correction to the gauge coupling (2.13). Thus,

$$|V_{\rm stop}| \propto \exp\left(-\gamma_{\rm br} N\right), \tag{2.18}$$

where $\gamma_{\rm br}$ is a number that has no parametric dependence on N or on the shift symmetry breaking parameter g. When $\gamma_{\rm br}$ is positive, the immediate and fatal consequence is the exponential suppression of the stopping potential.[6] The stop-

[6]Can one fine-tune $\gamma_{\rm br}$ to avoid the suppression in (2.18)? In §2.4.2 we explain what such a tuning would correspond to in terms of compactification parameters, but it is already clear that

ping potential, including the backreaction effect encoded in (2.18), is far too small to halt the evolution when the Higgs is almost massless, $|m_{\rm h}|^2 \ll M^2$. The result is a runaway relaxion. If instead $\gamma_{\rm br}$ is negative, the story is more involved. But, as we will see in §2.4.2, the result is still exponential suppression of the stopping potential.

Now to make things worse, there are two independent requirements that necessitate placing the source of monodromy in a region with a large background D3-brane charge $N_{\rm D3}$ that obeys $N_{\rm D3} \gg N$. This background charge introduces an additional, larger exponential suppression of the stopping potential.

First, achieving a large hierarchy between the Higgs vev v and the cutoff M necessitates an extremely small g. This parameter controls how strongly the relaxion shift symmetry is broken and is determined by the amount of energy introduced into the configuration per unit of monodromy charge. Since the source of monodromy corresponds to a physical quantized object, the amount of energy introduced by an additional winding is, in a sense, irreducible. However, *warping* the source of monodromy reduces this quantum of energy compared to other scales in the problem. So, an extremely small g—and thus a large hierarchy—may be realized by placing the source of monodromy in a strongly warped region, as in Figure 2.5 on page 54. As we will show in §2.3.2, we may characterize this warping by the amount of effective D3-brane charge $N_{\rm D3}$ needed to create the warped throat, and

$$g \propto N_{\rm D3}^{-1}. \tag{2.19}$$

this cannot be a satisfactory solution. Independent of where and how the stopping potential is generated, one would need to ensure that $\gamma_{\rm br} \lesssim \mathcal{O}(N^{-1})$, which by (2.3) reintroduces a tuning on the order of the hierarchy the mechanism was supposed to naturally explain.

Moreover, we must require a large background D3-brane charge to retain computational control and to ensure stability of the ten-dimensional configuration. The D3-brane charge induced as the relaxion is wound must be a small correction to the charge of the ambient space for the backreaction not to overwhelm the background configuration, and we therefore must require $N_{\rm D3} \gg N$.

Now as in the preceding section, the D3-brane charge $N_{\rm D3}$ of the ambient space has an effect akin to that of $N_{\rm D3}$ actual D3-branes, which would give rise to $N_{\rm D3}$ species of quarks in the fundamental of G. Loops of these quarks yield

$$\delta\left(\frac{8\pi^2}{g_{\rm YM}^2}\right) = \gamma_{\rm bg} N_{\rm D3}\,, \tag{2.20}$$

where $\gamma_{\rm bg}$ is a positive constant.

We conclude that the stopping potential is *exponentially suppressed* by the warping required to achieve a weak shift symmetry breaking $g \ll 1$ and to maintain control over the model. Schematically, we have

$$|V_{\rm stop}| \propto \exp\left(-\gamma_{\rm bg} N_{\rm D3}\right)\,, \tag{2.21}$$

with $N_{\rm D3} \gg N$. Here we remind the reader that $N \gg 1$ is the large number of windings required to substantially ameliorate the hierarchy problem. The suppression (2.21) renders the barriers utterly negligible.[7] Using (2.19), (2.21) can be written as

$$|V_{\rm stop}| \propto \exp\Big(-\mathcal{O}\left(1/g\right)\Big), \tag{2.22}$$

so we see that the generated hierarchy is no longer proportional to g. Instead, suppressing the shift symmetry breaking scale simultaneously suppresses the stopping potential barriers, leading to a relaxion runaway.

[7] This effect holds regardless of the sign of $\gamma_{\rm br}$ in (2.18). But, for $\gamma_{\rm br} < 0$, (2.21) is the only relevant exponential suppression, while for $\gamma_{\rm br} > 0$ (2.18) amounts to an independent suppression $\sim e^{-N}$ which even on its own is sufficient to cause a runaway.

Overview of the paper

The organization of this paper is as follows. In §2.2 we briefly survey relaxion models constructed in effective field theory, and identify the parameter ranges that allow relaxation of a hierarchy. In §2.3 we introduce axion monodromy in string theory, emphasizing the fact that monodromy results from a physical, quantized source. We review the scenario of axion monodromy on NS5-branes, and then explain how the relaxion mechanism could be realized in this setting. In §2.4 we determine the microphysical constraints that arise in such a realization. An executive summary appears in §2.4.1. We discuss generalizations in §2.5, and conclude in §2.6. The appendices contain more technical material. Appendix 2.A provides background on axions in string theory. In Appendix 2.B we prove that the D7-branes responsible for the stopping potential must intersect the NS5-brane. In Appendix 2.C we give the actions for D5-branes and NS5-branes in warped compactifications of type IIB string theory. In Appendix 2.D we analyze the backreaction of D3-brane and anti-D3-brane charge and tension on the metric of the internal space.

2.2 Relaxion Zoology

We now briefly overview a selection of existing relaxion models in field theory. To present a unified synopsis of the genus *Relaxion* in its various speciations, we discuss these models in a consistent notation and, since the number of windings N is severely constrained in string theory, we pay special attention to the field excursions required to generate a large hierarchy.

For generic initial displacements, the relaxion must scan a field range $\Delta\phi \sim M/g$ to reach $m_h^2 = 0$. If we generalize (2.2) and (2.4) by including a more generic dependence on the Higgs vev v as in [1], schematically

$$V_{\rm stop}(v,\phi) = \Lambda^4(v)\cos\left(\frac{2\pi\phi}{f}\right) = \epsilon\,\Lambda_c^4\left(\frac{v}{\Lambda_c}\right)^r\cos\left(\frac{2\pi\phi}{f}\right) \tag{2.23}$$

and

$$\left(\frac{v}{M}\right)^r \sim \frac{g}{\epsilon}\frac{f}{M}\left(\frac{M}{\Lambda_c}\right)^{4-r} \tag{2.24}$$

with ϵ a constant coefficient, this excursion implies a winding charge of

$$N = \frac{\Delta\phi}{f} \sim \frac{1}{\epsilon}\left(\frac{M}{\Lambda_c}\right)^4\left(\frac{\Lambda_c}{v}\right)^r. \tag{2.25}$$

For $\Lambda_c \sim v$, N scales with the fourth power of the ratio of the cutoff scale M to the weak scale, further increasing if $\Lambda_c \ll v$.

2.2.1 Original models

Two explicit constructions were originally proposed in [16], and the dynamics of these models were explained in §2.1. The relaxion in the first model (GKR1) is the QCD axion and the potential barriers are generated by strong chromodynamic forces. The potential barriers in (2.23) then scale as $\Lambda^4(v) \sim \Lambda_{\rm QCD}^3 m_u$, i.e. $r = 1$, $\Lambda_c \sim \Lambda_{\rm QCD}$, and ϵ is the up-quark Yukawa coupling y_u. The main drawback of this model is that it destroys the solution to the strong CP problem. The PQ solution may be restored, as discussed in [16], by introducing additional dynamics at the end of inflation, which removes the slope of the relaxion potential at the end of inflation. However, in this case the cutoff scale cannot be pushed higher than $M \approx 30$ TeV. The hierarchy (2.25) is then multiplied by a factor of the QCD angle $\theta_{\rm QCD}$. In either case, the number of windings obeys $N \geq (M/\Lambda_{\rm QCD})^4$.

Because the generated hierarchy (2.25) grows with the confinement scale Λ_c, the second model (GKR2) introduces a new strongly-interacting gauge sector G, whose axion is the relaxion. The PQ solution to the strong CP problem is then untouched and Λ_c can be much larger than $\Lambda_{\rm QCD}$. However, this does not allow one to make the barrier arbitrarily high. New electroweak scale fermions couple the Higgs sector to this new sector and the barrier height depends quadratically ($r = 2$) on the Higgs vev. But, a constant term ($r = 0$) will also be generated by quantum corrections so the barriers can schematically be written as [1]

$$\Lambda^4(v) \sim \epsilon\, \Lambda_c^4 \left(1 + \left(\frac{v}{\Lambda_c}\right)^2\right). \tag{2.26}$$

So, the barrier will not depend strongly enough on the Higgs vev v for the relaxion mechanism to work unless $\Lambda_c \lesssim v$. Still, GKR2 can generate a much larger hierarchy $M \lesssim 10^8$ GeV than GKR1, with a similar parametric scaling of the number of windings $N \geq (M/v)^4$. Unfortunately, GKR2 requires that new electroweak scale fermions be put in by hand, and this coincidence of scales must be explained.

2.2.2 CHAIN

A solution to this coincidence problem was suggested in [1], and involves taking the barrier height to depend on an extra scalar field. The relaxion mechanism is then able to explain the near-criticality of the Higgs without a coincidence of scales. Instead, there is only one scale in the problem, the cutoff M, which also sets the barrier height $\Lambda_c \sim M$. The extra scalar σ, which need not be an axion, controls the height of the stopping potential,

$$\Lambda^4(h, \phi, \sigma) = \epsilon M^4 \left(\beta + c_\phi \frac{g\phi}{M} - c_\sigma \frac{g_\sigma \sigma}{M} + \frac{h^2}{M^2}\right). \tag{2.27}$$

	GKR 1	GKR 2	CHAIN with $f \sim M$
f	$f_{PQ} \sim 10^{10} - 10^{12}$ GeV	$\gtrsim M_{\rm GUT} \sim 10^{16}$ GeV	$\gtrsim M$
g	$\left(\frac{\Lambda_{\rm QCD}}{M}\right)^4 \frac{M}{f_{\rm PQ}} \theta_{\rm QCD} \lesssim 10^{-36}$	$\left(\frac{\Lambda_{\rm EW}}{M}\right)^4 \frac{M}{M_{\rm GUT}} \sim 10^{-30}$–$10^{-20} \lesssim$	$v^4/M^4 \sim 10^{-26}$–10^{-6}
$M_{\rm max}$	3×10^3 GeV	10^8 GeV	10^9 GeV
m_ϕ	$\frac{\Lambda_{\rm QCD}^2}{f_{\rm PQ}} \lesssim 10^{-11}$ GeV	$\frac{\Lambda_{\rm EW}^2}{M_{\rm GUT}} \lesssim 10^{-12}$ GeV	$\sqrt{gM^4/v^2} \lesssim v$
$\Delta\phi/f$	$\theta_{\rm QCD}^{-1}\left(\frac{M}{\Lambda_{\rm QCD}}\right)^4 \gtrsim 10^{30}$	$(M/\Lambda_{\rm EW})^4 \sim 10^8 - 10^{24}$	$g^{-1} \sim 10^6 - 10^{26}$

Table 2.1: Summary of parameter values in the three non-supersymmetric relaxion models discussed in §2.2.

The initial conditions are very different from both GKR1 and GKR2. At first, the barriers are large and the relaxion is stuck in one of its minima. As the second field σ evolves, its vev will eventually cancel this barrier and allow the relaxion to roll. In contrast with GKR2, there are no constraints on the decay constant f from reheating.

Given that now the barriers are allowed to be high, $\Lambda_c \gg v$, one might hope that the required number of windings for a given cutoff scale is substantially reduced. A more careful analysis, however, reveals that this is not the case. Instead, because classical evolution must dominate over quantum fluctuations, we require that $\epsilon \lesssim v^2/M^2$, while imposing that the Higgs barrier in (2.27) is solely responsible for stopping ϕ requires that $v^2 \sim gMf/\epsilon$. Together, these imply that

$$N \sim \frac{M}{gf} \gtrsim \left(\frac{M}{v}\right)^4, \tag{2.28}$$

and so the CHAIN model also requires a large number of windings to resolve the hierarchy problem.

A comparison of the three models is shown in Figure 2.2 and Table 2.1, while further phenomenological constraints are discussed in [29].

2.2.3 Supersymmetric models

Inflation limits the achievable cutoff scale to $M \sim 10^9$ GeV. The energy stored in the relaxion must not dominate over the energy driving inflation,

$$M^4 < H^2 M_{\rm pl}^2 \,, \tag{2.29}$$

where H is the Hubble rate during inflation, and barriers cannot form unless $H < \Lambda_c$. This immediately implies a bound on the cutoff $M \lesssim \sqrt{v M_{\rm pl}} \sim 10^9$ GeV for GKR2. While this argument does not directly apply to the CHAIN model, there one finds the same bound $M \lesssim 10^9$ GeV. Since we must also explain the remaining hierarchy between the cutoff M and the Planck scale $M_{\rm pl}$, a natural candidate solution is that supersymmetry is restored above M and the relaxion is embedded within a supersymmetric model.

A supersymmetric version of GKR1 was presented in [30], on which the following discussion is based. The relaxion becomes part of a chiral superfield S:

$$S = \frac{s + ia}{\sqrt{2}} + \sqrt{2}\theta\tilde{a} + \theta^2 F + \dots \,, \tag{2.30}$$

which contains the (dimensionless) relaxion $a = \phi/f$, a srelaxion field s, and the relaxino $\tilde{a}$. The Peccei-Quinn symmetry acts as $S \mapsto S + i\alpha$. The linear term Ⓑ in the relaxion potential (2.1) descends from the superpotential term

$$W \supset \frac{1}{2} m f^2 S^2 \,. \tag{2.31}$$

Small $m \ll f$ is technically natural since m breaks the PQ symmetry, which is non-linearly realized via the term

$$W \supset \mu_0 \, e^{-qS} H_u H_d. \tag{2.32}$$

Apart from S, the model contains only SM particles and their superpartners (including the usual second Higgs doublet). The effective potential for s and a is then

$$V = \frac{1}{2} m^2 f^2 \left(s^2 + a^2\right) \kappa(s) \,, \tag{2.33}$$

with $\kappa(s)$ a function of s. As in all relaxion models, a starts out at a field value far away from its minimum at $a = 0$, and so breaks supersymmetry, with $F \propto ma$. As a evolves towards its minimum, it scans the SUSY breaking scale, and therefore the soft masses of the gauginos and the scalar superpartners. In particular, the determinant of the Higgs mass matrix was shown to scale as a^4 for $a \gg \mu_0/m$, far away from any electroweak symmetry breaking minima. As a approaches the critical value $a_* = \mu_0/m$, electroweak symmetry is broken, the Higgs(es) obtain a vacuum expectation value, and barriers appear that halt the evolution of a.

For a suitable choice of parameters, the model explains the hierarchy between the electroweak symmetry breaking scale v and the mass scale of the superpartners $\mu_0 \gg v$, thus solving the supersymmetric little hierarchy problem. According to [30], the number of windings scales as

$$N \sim \Delta a \sim \frac{f^2 \mu_0^2}{\Lambda_{\text{QCD}}^4} \,, \tag{2.34}$$

where $f \sim 10^9$–10^{12} GeV is the QCD axion decay constant and μ_0 plays the role of the UV cutoff. For $\mu_0 = 10^5$ GeV, a field excursion of $\Delta a \sim 10^{30}$ is required. Without further modifications, this model is phenomenologically unacceptable since it predicts $\theta_{\text{QCD}} \sim \mathcal{O}(1)$. A variation with a non-QCD axion similar to GKR2 is briefly discussed in [30]. In this case, a larger range of decay constants is allowed—for $f = \mu_0$ and $\Lambda \sim v$ one obtains the same scaling as in GKR2, namely $N \sim \mu_0^4/v^4$.

A supersymmetrization of the CHAIN model was proposed in [31]. The philosophy is similar to the above discussion—now both the relaxion and the addi-

tional scalar σ are promoted to chiral superfields. The barriers for the relaxion are generated by a new SU(N_g) gauge theory, with confinement scale Λ_g, which communicates with the Higgs sector via a set of vector-like leptons. The required field excursion in this model is

$$N \equiv \frac{\Delta\phi}{f} \gtrsim \frac{m_{\rm SUSY}}{|m_{\rm S}|}\,, \tag{2.35}$$

where $m_{\rm SUSY} \sim \mu_0$ is the supersymmetry-breaking scale, and $m_{\rm S}$ is the relaxion mass coming from a term similar to (2.31). For $\Lambda_g \sim f \sim m_{\rm SUSY}$, a supersymmetry-breaking scale $\sim 10^9$ GeV may be generated through a field excursion of $N \sim 10^{27}$. So while the field excursion seems to grow more moderate as a function of the cutoff, it is still as large as in the earlier models for the largest possible cutoff (e.g. M^4/v^4 is of order 10^{26}). Instead, even the best case scenarios with $m_{\rm SUSY} \sim 10^4$ GeV require a large number of windings $N \gtrsim 10^8$.

2.2.4 Summary

The models presented above do not represent a complete classification of genus *Relaxion*. In particular, we are not considering models that rely on the alignment of multiple axions or use friction from particle production to halt the evolution of ϕ. We discuss both of these further in §2.5. However, the models that we examine represent a large cross-section of *Relaxion* and share a common trait: the required field excursion scales parametrically with the hierarchy generated, and so the associated number of windings around the relaxion field space $N \equiv \Delta\phi/f$ is enormous. In what follows, we will argue that N is a *physical charge* in string theory, which backreacts on the ten-dimensional configuration and tragically destroys the structures in (2.1), allowing for a runaway relaxion.

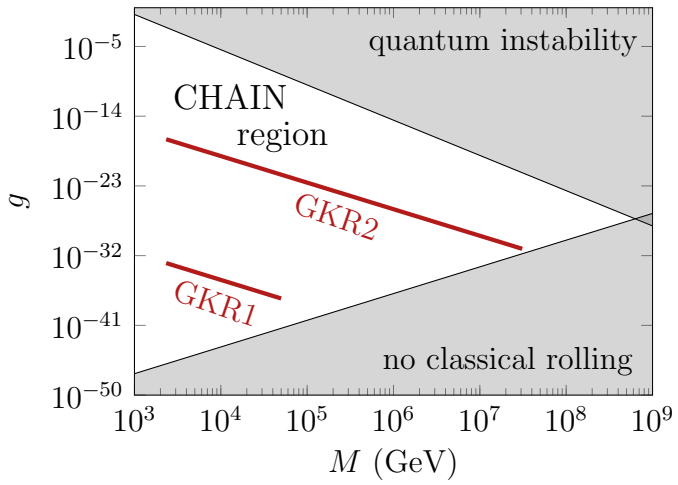

Figure 2.2: Schematic parameter space in the three main non-supersymmetric relaxion models. See [1] for the derivation of the constraints on the parameter space.

2.3 Relaxion Monodromy

2.3.1 Axion monodromy in string theory

Axions are commonplace in string compactifications,[8] and arise when a p-form gauge potential—either the NS–NS two-form B_2 or an R–R p-form C_p—is dimensionally reduced along a non-trivial cycle Σ_p in the compactification manifold X_6. The ten-dimensional supergravity action is invariant under the gauge symmetry $B_2 \mapsto B_2 + \mathrm{d}\Lambda_1$ and $C_p \mapsto C_p + \mathrm{d}\Lambda_{p-1}$ which, upon reduction to four dimensions, ensures that the axion enjoys a perturbatively exact shift symmetry. For an axion a associated with a non-trivial cycle Σ_p, the shift symmetry $a \mapsto a + \text{const.}$ may be broken to a discrete shift symmetry by non-perturbative effects, or completely broken by a brane wrapping Σ_p. In the latter case, the explicit breaking is proportional to the brane's tension. For example, if one wraps an NS5-brane along a

[8] A detailed treatment of the material that follows can be found in [8], §5.4.2. An overview is given in Appendix 2.A, and Table 2.2 gives a simple dictionary.

Relaxion Quantity	String Theory Origin
Axion ϕ	NS-NS or R-R p-form gauge field, dimensionally reduced along non-trivial p-cycle, §2.3.2
Discrete shift symmetry $\phi \mapsto \phi + f$	Ten-dimensional NS-NS or R-R gauge symmetry, exact in absence of brane or flux, §2.3.2
Physical source of monodromy explicitly breaks $\phi \mapsto \phi + f$	Wrapped brane or flux along axion p-cycle, §2.3.2
Shift symmetry-breaking scale $gM^3 f$	Warped brane tension, §2.3.2
Winding number $N \equiv \Delta\phi/f$	Quantized monodromy charge, §2.3.2
Axion decay constant f	Set by internal six-dimensional geometry, §2.A
Stopping potential barrier height $\Lambda(v)$	Set by warped volume of a four-cycle, §2.3.3

Table 2.2: A quick string theory-relaxion dictionary. This is a summary table, with more extended explanations given throughout this chapter and in its appendices.

two-cycle Σ_2 in the compactification manifold, the axion field c, defined by

$$c \equiv \frac{1}{\ell_s^2} \int_{\Sigma_2} C_2 \,, \tag{2.36}$$

experiences monodromy. The four-dimensional action for the dimensionless field c takes the form [32]

$$\mathcal{L} = -\frac{1}{2} f^2 (\partial c)^2 - \varepsilon \mu_0^3 f c \,, \tag{2.37}$$

where f is the axion decay constant,[9] μ_0 is a parameter with dimensions of mass, determined by the geometry of X_6, and ε parameterizes the warp factor at the location of the NS5-brane. (In terms of the warped line element (2.48), we have $\varepsilon = e^{4A_\cup}$.) We will explain the potential (2.37) in more detail in §2.3.2.

[9]The axion decay constant depends on the topology and geometry of the six-dimensional compact manifold X_6: see Appendix 2.A.

Defining the canonically normalized axion $\phi \equiv fc$, we have

$$\mathcal{L} = -\frac{1}{2}(\partial\phi)^2 - \varepsilon\mu_0^3\phi\,. \tag{2.38}$$

Comparing to the relaxion potential (2.1), we have the correspondence

$$gM^3 = \varepsilon\mu_0^3\,. \tag{2.39}$$

So $\varepsilon \ll 1$ corresponds to $g \ll 1$ in the relaxion model. Since the breaking of the shift symmetry $\phi \mapsto \phi + \text{const.}$ is proportional to the warp factor at the location of the fivebrane, strong warping could lead to weak breaking of the symmetry, and hence to the small values of g required for a relaxion model.

The potential (2.38) has the desirable property that the entire potential is proportional to the warp factor, so it appears completely natural to make this potential small. However, a central observation of this paper is that achieving small g through warping, without unintended consequences elsewhere in the action, is challenging.

Requirements for Axion Monodromy

Let us first summarize the core ingredients mentioned above. For a model of axion monodromy in string theory, one requires:

1. An axion field descending from a p-form, and a *source of monodromy*: a brane, flux, or other physical ingredient that causes the configuration space to be a multi-cover of the axion circle, rather than just a single circle.
2. To have a plausible mechanism for making the breaking of the shift symmetry *weak*, the source of monodromy should be in a *warped region*.

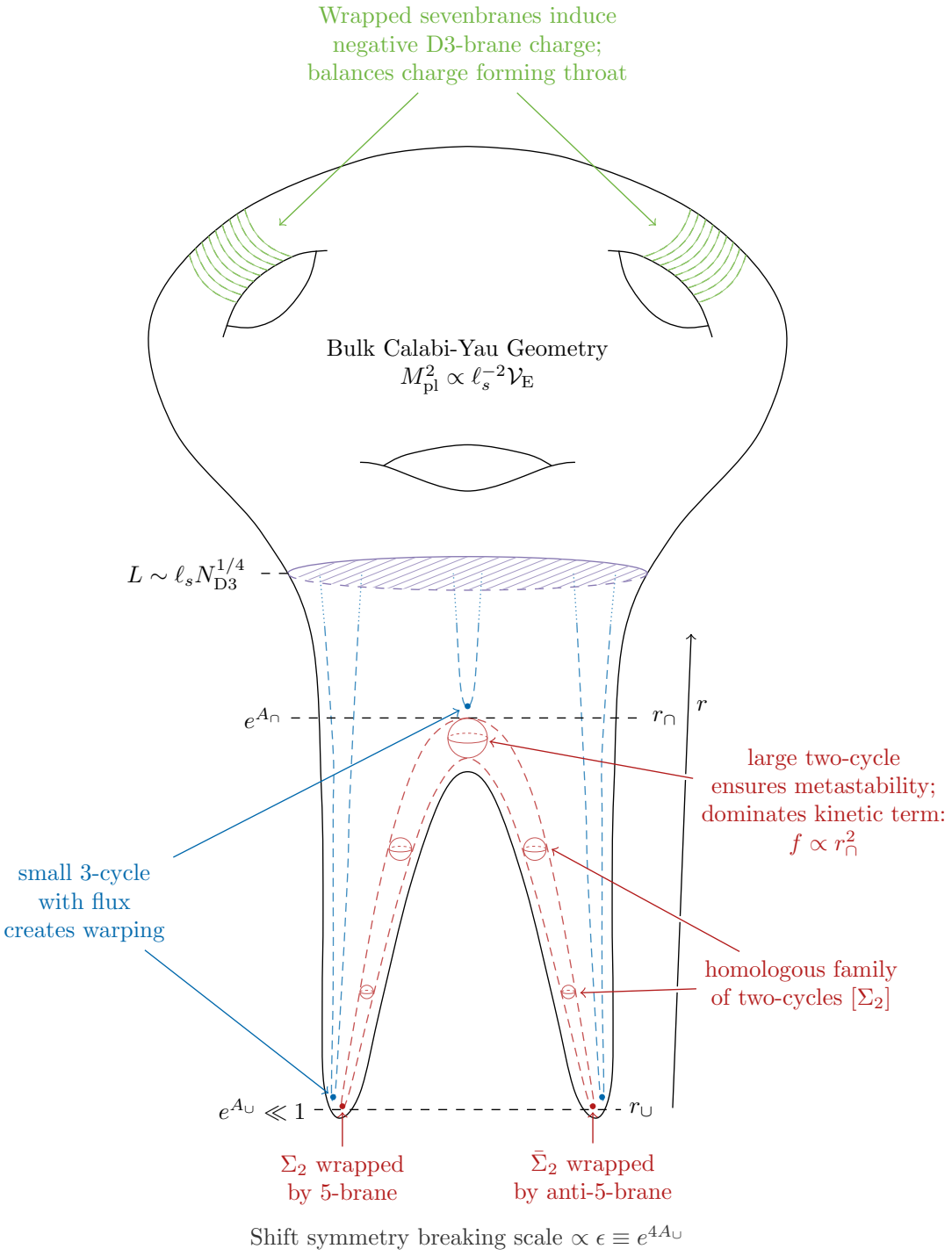

Figure 2.3: Minimal bifurcated warped throat setup for relaxion monodromy with 5-branes.

3. Most of the issues that arise as possible obstacles become visible only in vacua with stabilized moduli: if one ignores the moduli sector, many problems disappear. But, of course, moduli stabilization is needed for a cosmological model. So the axion and the source of monodromy must be situated in a *vacuum with stabilized moduli*.

4. Since the compactification must have finite volume in order to lead to a finite four-dimensional Newton constant, Gauss's law imposes strict constraints on the charges in the compact space X_6, and so we must *satisfy all tadpole conditions*.

There are many mechanisms in the literature that achieve (1), for instance [33]. But there is only one model currently available that achieves (1)-(3) [32, 34]: this is a model with an NS5-brane/anti-NS5-brane pair in a warped throat region of a type IIB flux compactification whose complex structure moduli are stabilized by fluxes, and whose Kähler moduli are stabilized by nonperturbative effects and possibly also by perturbative effects. We will call this model, whose detailed properties we will review in §2.3.2, the *NS5-brane model*.

The central physics of the NS5-brane model is that transporting the dimensionless axion over a period induces one unit of D3-brane charge on the NS5-brane, and one unit of anti-D3-brane charge on the anti-NS5-brane. That is, "winding up" the axion by one cycle develops a D3-brane dipole in the compact space; the axis of the dipole is the line from the NS5-brane to the anti-NS5-brane. The entire dipole is in the infrared region of the warped throat where the fivebrane pair lives. See Figure 2.3.

The key point is that the Lagrangian (2.38) arising from the NS5-brane DBI action, which is intended to be the relaxion Lagrangian, holds in the so-called *probe*

approximation. That is, the potential in (2.38) follows from including the tension of the D3-branes and anti-D3-branes as a contribution to the four-dimensional vacuum energy, i.e. as a source in the four-dimensional Einstein equations, but *not* including this tension as a source in the ten-dimensional Einstein equations. The effects of a particular source on the ten-dimensional field configuration are termed the *backreaction* of that source, and so the probe approximation consists of neglecting the backreaction of D3-branes and anti-D3-branes.[10]

An immediate question is whether applying the probe approximation is consistent; in other words, can the backreaction of D3-branes be neglected? In the context of axion monodromy inflation in string theory, this question has been addressed, with the outcome that backreaction can be suppressed to some degree, by a variety of mechanisms, but nevertheless remains as a leading constraint on model-building [34]. However, the sources of backreaction are the D3-brane charge and tension, both proportional to the number of windings N of the axion. In the present context of relaxion monodromy, N needs to be extremely large, and so the problem of backreaction is much more severe than in the corresponding inflationary models. The constraints examined in [34] must therefore be revisited under this more severe test.

In this work, we will carefully examine the consequences of D3-brane backreaction for the NS5-brane model of relaxion monodromy in string theory. The first step is to explain how to compute the backreaction in this scenario.

[10]For brevity we will often speak of "D3-branes," "D3-brane backreaction," etc., with the understanding that both D3-branes and anti-D3-branes are included.

2.3.2 Fivebrane axion monodromy

Our analysis will rely on detailed properties of the action for NS5-branes wrapping curves in a warped region of a type IIB flux compactification, so we now give some essential background. We will begin by discussing D5-branes, to facilitate comparison with the string theory literature, even though our eventual interest will be NS5-branes.

The action of a D5-brane is the sum of a Dirac-Born-Infeld term related to the worldvolume $\mathcal{W}$ of the brane,

$$S_{\rm DBI} = -g_s T_5 \int_{\mathcal{W}} \mathrm{d}^6\sigma\, e^{-\Phi} \sqrt{-\det(G_{ab} + \mathcal{F}_{ab})}\,, \tag{2.40}$$

and a Chern-Simons term encoding the coupling of the D5-brane to the Ramond-Ramond p-form potentials C_0, C_2, C_4, and C_6,

$$S_{\rm CS} = \mu_5 \int_{\mathcal{W}} \sum_p C_p \wedge e^{\mathcal{F}}\,, \tag{2.41}$$

with $\mathcal{F} = B + 2\pi\alpha' F$. Here g_s is the string coupling, T_5 is the D5-brane tension, G_{ab} is the metric induced on the D5-brane, μ_5 is the D5-brane charge, and $\mathcal{F}$ is the gauge-invariant two-form field strength on the D5-brane. The integral in (2.41) picks out the six-forms C_6, $C_4 \wedge \mathcal{F}$, $C_2 \wedge \mathcal{F} \wedge \mathcal{F}$, and $C_0 \wedge \mathcal{F} \wedge \mathcal{F} \wedge \mathcal{F}$.

Now suppose that $\mathcal{W} = \mathcal{M}^{3,1} \times \Sigma_2$, with Σ_2 a two-cycle in the internal six-manifold X_6. If the field strength $\mathcal{F}$ obeys[11]

$$\frac{1}{\ell_s^2} \int_{\Sigma_2} \mathcal{F} = N \in \mathbb{Z}\,, \tag{2.42}$$

then the Chern-Simons coupling becomes

$$\mu_5 \int_{\mathcal{W}} C_4 \wedge \mathcal{F} \to N\mu_3 \int_{\mathcal{M}^{3,1}} C_4. \tag{2.43}$$

[11] We define the string length to be $\ell_s \equiv 2\pi\sqrt{\alpha'}$.

The interaction (2.43) is precisely N times the Chern-Simons coupling of a single D3-brane to the Ramond-Ramond four-form potential C_4, under which the D3-brane is electrically (and also magnetically) charged. The coupling (2.43) should be understood as a generalization of the worldline coupling

$$\mathcal{L}_{\text{int}} = -\frac{e}{c}\int A_\mu \, \mathrm{d}x^\mu \tag{2.44}$$

in electromagnetism. In particular, (2.43) shows that a D5-brane wrapping Σ_2, with N units of $\mathcal{F}$ flux on Σ_2, carries N units of D3-brane charge. Equivalently, the D5-brane can be said to contain N D3-branes dissolved in the D5-brane. This fact, while well-known, will be crucial for our considerations.

The Σ_2-wrapping D5-brane can fluctuate in the space orthogonal to $\mathcal{M}^{3,1}\times\Sigma_2$. We denote these corresponding canonically-normalized fluctuations as X^i. Defining the dimensionless field

$$b \equiv \frac{1}{\ell_s^2}\int_{\Sigma_2} B_2 \,, \tag{2.45}$$

we may expand the DBI action (2.40) to second order in these fluctuations,[12]

$$\begin{aligned} S_{\text{DBI}} = -\frac{T_5}{2}\int \mathrm{dvol}_4 \, \mathrm{d}^2 z \times \\ \Bigg(\underset{①}{\sqrt{4\tilde{g}_2 + \ell_s^4 b^2}} + \underset{②}{\partial_\mu X^i \partial^\mu X^i} + \underset{③}{\frac{4\tilde{g}_2}{4\tilde{g}_2 + \ell_s^4 b^2}\partial_a X^i \partial^a X^i} + \dots \Bigg), \end{aligned} \tag{2.46}$$

where $\tilde{g}_2$ is the determinant of the metric on Σ_2. Upon integrating over Σ_2 and denoting its volume as ℓ^2, ① yields a four-dimensional potential for b

$$V(b) = \frac{\varepsilon}{(2\pi)^3 \alpha'^2}\sqrt{\left(\frac{\ell}{\ell_s}\right)^4 + \frac{b^2}{4}}, \tag{2.47}$$

[12] As in Appendix 2.C, we denote $\mathcal{M}^{3,1}$ indices with μ,ν, etc.; Σ_2 indices with a,b, etc.; directions orthogonal to $\mathcal{M}^{3,1}\times\Sigma_2$ with indices i,j, etc.; and we parameterize Σ_2 using the coordinates y and z, with $\mathrm{d}y\wedge\mathrm{d}z = \mathrm{d}^2 z$; see Table 2.3.

In the absence of a wrapped D5-brane, b would enjoy an approximate continuous shift symmetry, $b \mapsto b + \text{const.}$, that is broken to a residual exact discrete shift symmetry, $b \mapsto b + 1$, by instanton effects.[13] However, the potential (2.47) induced by the D5-brane *completely breaks* this symmetry. In fact, the D5-brane introduces a monodromy, in that upon traversing the axion circle, from $b \mapsto b + 1$, the potential energy is increased, rather than being periodic. For large b, the potential (2.47) becomes linear, as claimed in (2.38) for the related case of an NS5-brane.

The strength of this symmetry breaking is proportional to ε, the *warp factor* $\varepsilon = e^{4A_\cup}$ at the location of the fivebrane. In a warped compactification, the ten-dimensional metric takes the form

$$\mathrm{d}s_{10}^2 = e^{2A(y)} g_{\mu\nu} \mathrm{d}x^\mu \, \mathrm{d}x^\nu + e^{-2A(y)} \tilde{g}_{mn} \, \mathrm{d}y^m \, \mathrm{d}y^n . \tag{2.48}$$

By placing the fivebranes in a warped throat, the energy of this shift symmetry breaking can be gravitationally redshifted to an energy much smaller than the natural scale of breaking due to unwarped fivebranes.

The monodromy is closely related to the induced D3-brane charge (2.43). Starting from an initial configuration with $b = b_0$ and moving to $b = b_0 + N$ (for $N > 0$) corresponds to shifting

$$\mathcal{F} \mapsto \mathcal{F} + N\omega_2 \, , \tag{2.49}$$

with ω_2 a two-form obeying $\ell_s^{-2} \int_{\Sigma_2} \omega_2 = 1$. This is an increase, of N units, of the gauge-invariant field strength $\mathcal{F}$. This change is manifest in the potential (2.47), which increases linearly. The change is also visible in the D3-brane charge carried by the D5-brane, which increases by N units. We refer to this process as "winding up the axion N times."

[13] Moduli-stabilizing effects further break this symmetry, as explained in [32].

A justifiable complaint at this stage is that in a compact space, the total D3-brane charge should be fixed: in fact it must vanish by Gauss's law. So winding up the axion would appear to be forbidden. However, to cancel the D5-brane tadpole, we may suppose that in addition to the D5-brane wrapping Σ_2, there is an anti-D5-brane also wrapping Σ_2. The anti-D5-brane Chern-Simons coupling differs from the D5-brane Chern-Simons coupling (2.41) by an overall minus sign. Thus, winding up the axion N times induces N units of D3-brane charge on the D5-brane, as well as $-N$ units of D3-brane charge on the anti-D5-brane, so that no net D3-brane charge is produced, and if Gauss's law is obeyed in the initial configuration, it is also obeyed after winding.

A coincident D5-brane and anti-D5 brane will quickly annihilate. However, if a D5-brane wraps Σ_2, and an anti-D5-brane wraps a two-cycle $\overline{\Sigma}_2$ that is *homologous* to Σ_2, but is not coincident with Σ_2, then the D5-brane/anti-D5-brane configuration can be metastable and cosmologically long-lived [35]. Because the induced D3-brane charges are determined by the homology classes of Σ_2 and $\overline{\Sigma}_2$, if $[\Sigma_2] - [\overline{\Sigma}_2]$ is trivial in homology then no net D3-brane charge is induced, just as in the case of a strictly coincident D5-brane/anti-D5-brane pair, and Gauss's law does not preclude winding up the axion.

Let us summarize the physics of B_2 monodromy from a wrapped D5-brane. The D5-brane is a source of monodromy and gives rise to the non-periodic potential (2.47). The order parameter measuring the distance from the origin in the b field space is the number of windings, $N \in \mathbb{Z}$, which also counts the D3-brane charge induced on the D5-brane. This is the *monodromy charge* in the fivebrane model. Winding up corresponds to moving away from the origin in field space and storing energy in the form of the D3-branes dissolved in the D5-brane, and anti-D3-branes

dissolved in the anti-D5-brane: that is, the energy is stored in the monodromy charge.

The potential (2.47) is that of a *probe* D5-brane, in the same sense that (2.44) includes the potential energy of an electron in a background electromagnetic field. However, just as (2.44) also encodes the fact that electrons source electromagnetic fields, the couplings (2.40) and (2.41) encode the effects that a D5-brane has on the background fields. To determine this *backreaction* of the D5-brane on the bulk field, including the metric and the p-form fields, we simply include the couplings (2.40) and (2.41) when varying the ten-dimensional action with respect to these fields φ. Schematically,

$$0 = \frac{\delta}{\delta\varphi} S_{\text{10d,bulk}} + \frac{\delta}{\delta\varphi} S_{\text{DBI}} + \frac{\delta}{\delta\varphi} S_{\text{CS}} \,. \tag{2.50}$$

Any D5-brane serves as a source for the ten-dimensional metric (it has tension), and as a source for C_6. But a D5-brane with

$$\frac{1}{\ell_s^2} \int_{\Sigma_2} \mathcal{F} = N \neq 0 \tag{2.51}$$

also serves as a source for C_4; this is just to say that such a D5-brane carries D3-brane charge. The DBI action (2.40) may be interpreted as the product of the brane tension and its "effective volume," which grows with N. This growth has two principal effects. The mass of the five-brane is also, schematically, the product of its tension and this effective volume, and thus as N grows the charged D5-brane will more strongly source the ten-dimensional metric. Furthermore, there are Kaluza-Klein excitations arising from the dimensional reduction of (2.40) whose masses *decrease* as this effective volume grows; indeed, the dimensional reduction of ② and ③ in (2.46)—which correspond to the transverse fluctuations of the five-brane—yield Kaluza-Klein modes with masses m_{bKK} that are smaller than the

naive estimate $m_{\rm KK} \propto \ell^{-1}$ by a factor of (see Appendix 2.C.2)

$$\frac{m_{\rm bKK}}{m_{\rm KK}} \sim \frac{\ell^2}{\sqrt{\ell^4 + \ell_s^4 b^2}}\,. \tag{2.52}$$

Axions descending from B_2 generically suffer an η problem [32], meaning that in expansion around a vacuum with stabilized moduli, the actual potential for the axion, taking into account all couplings to moduli, is very different from the potential (2.47) that arises from the probe D5-brane action alone. This problem can be ameliorated by considering an axion descending from the Ramond-Ramond two-form C_2 and exchanging the D5-branes in the above discussion for NS5-branes. The analogous potential is then given by

$$V(c) = \frac{\varepsilon}{(2\pi)^3 g_s \alpha'^2}\sqrt{\left(\frac{\ell}{\ell_s}\right)^4 + \frac{g_s^2 c^2}{4}}. \tag{2.53}$$

As we will argue in §2.3.3, for a construction of a relaxion model via fivebrane axion monodromy in string theory one needs an extremely large winding $N \gg 1$. There is a correspondingly large induced D3-brane charge, the effect of which must be included in the ten-dimensional field equations. Backreaction of this charge and its effect on the five-brane cannot be neglected: the potential for the axion is no longer simply given by (2.47), and new light modes appear.

2.3.3 Fivebrane relaxion monodromy

To understand string theoretic constraints on the relaxion mechanism, we require an embedding of the four-dimensional potential

$$V(\phi, h) = \underbrace{\Big(M^2 - g_h M\,(\phi_{\rm init} - \phi)\Big)}_{\text{Ⓐ}}|h|^2 + \underbrace{gM^3\phi}_{\text{Ⓑ}} + \underbrace{V_{\rm stop}(\phi, h)}_{\text{Ⓒ}}\,, \tag{2.54}$$

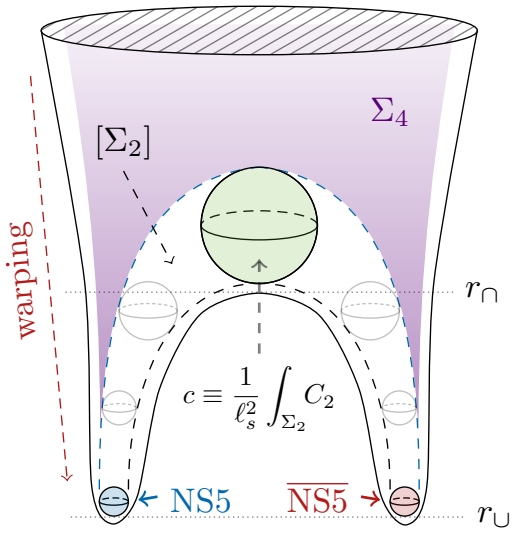

Figure 2.4: Ten-dimensional realization of Ⓑ and Ⓒ of (2.54). Ⓒ is generated by strong gauge dynamics on seven-branes wrapping a divisor Σ_4, which must necessarily intersect the minimum volume representative $[\Sigma_2]$ wrapped by the NS5-/anti-NS5-brane.

or of something functionally equivalent, in a well-controlled compactification of string theory. As noted in the introduction, the ratio g_h/g need not be $\mathcal{O}(1)$, and so in (2.54) we distinguish between the two.

In §2.3.2, we focused on realizing Ⓑ as the potential energy of an NS5-/anti-NS5-brane pair wrapping the minimum volume representatives of the homology class $[\Sigma_2]$ associated with the axion $c = \ell_s^{-2}\int_{\Sigma_2} C_2$, where $\phi \equiv fc$. Ⓑ provides a potential that is self-similar (ignoring backreaction effects) over a very large distance $\Delta\phi \gg f$ in field space. Hubble friction eventually dominates and the late-time dynamics are independent of the initial conditions for ϕ.

Crucially, the small parameter g is controlled by the warp factor at the position $r_\cup$ of the five-branes. Specifically,

$$gM^3 f \equiv \frac{2\pi}{\ell_s^4} e^{4A_\cup}, \tag{2.55}$$

where $e^{4A_\cup}$ is the warp factor at $r_\cup$, the location of the fivebranes and the bottom of

the "tooth" in Figure 2.4. We may think of the "roots" of the tooth as Klebanov-Strassler or similar warped throat geometries. Away from the tip, the warp factor is roughly $e^{4A} \sim r^4/L^4$, with L the characteristic size of the warped throat. A simple way to describe this warping is by the number of D3-branes it would take to form a similarly sized warped throat,

$$L^4 \sim g_s N_{\rm D3} \ell_s^4. \tag{2.56}$$

As explained in §2.A, the axion decay constant f is determined by the radial position of the arch of the "tooth,"

$$f^2 \sim g_s \frac{r_\cap^2}{\ell_s^4}, \tag{2.57}$$

so the shift symmetry breaking scale is given by

$$gM^3 \sim \frac{2\pi}{g_s^{3/2} \ell_s^3 N_{\rm D3}} \left(\frac{r_\cup}{r_\cap}\right) \left(\frac{r_\cup}{\ell_s}\right)^3. \tag{2.58}$$

The cutoff scale M depends on how the Higgs is realized and does not necessarily depend on the total D3-brane charge $N_{\rm D3}$. However, regardless of where the Higgs is located in the internal space, the smallness of g is necessarily tied to a large $N_{\rm D3}$. For example, if the Higgs sector is realized somewhere in the bulk geometry, then $M \propto N_{\rm D3}^0$ and $g \propto N_{\rm D3}^{-1}$, as in (2.58). If instead the Higgs sector is realized at the top of the warped throat at $r \sim L$ in Figure 2.4, then we may take $M^3 \sim L^{-3}$ and so

$$g \sim \frac{1}{(g_s N_{\rm D3})^{1/4}} \left(\frac{r_\cup}{r_\cap}\right) \left(\frac{r_\cup}{\ell_s}\right)^3. \tag{2.59}$$

We will not consider a Higgs realized deep within the warped throat, as this would lead to an exponential suppression of M, corresponding to a supersymmetric resolution of the hierarchy.

We will be agnostic about the detailed origin of the Higgs coupling Ⓐ. While its specific form would be relevant in a complete model, it is not needed to expose

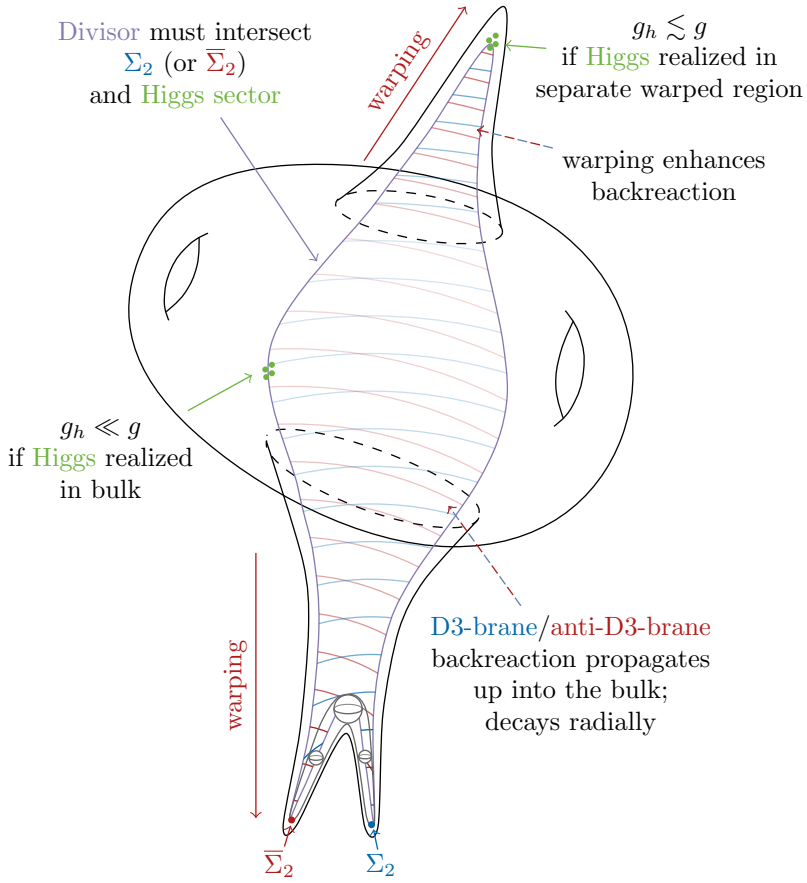

Figure 2.5: Schematic structure of the extra dimensions, showing a string theory setup that realizes the main relaxion features and couplings. The central region is the "bulk" of the extra dimensions, which does not experience position-dependent warping. The coupling g_h depends on where in the bulk Calabi-Yau the Higgs sector is realized.

and quantify the issues that concern us here, which mainly deal with the interplay between the linear potential and the stopping potential. In the spirit of this agnosticism, we instead focus on the hierarchy generated between the string scale M_s and the electroweak scale v.

Even so, a concrete picture of one possibility may be helpful. The Higgs could arise from open strings stretching between stacks of D3-branes or D7-branes. The Higgs mass is then proportional to the distance between the $\mathrm{U}(1)_\mathrm{Y}$ brane and the $\mathrm{SU}(2)_\mathrm{W}$ stack. The coupling Ⓐ is generated by backreaction of the monodromy

charge on the internal geometry, which changes distance between these branes and thus the Higgs mass, as in Figure 2.5. Because this backreaction decays as it propagates throughout the six-dimensional space, there may be an appreciable hierarchy between g_h and g which depends on where the Higgs sector is realized in the internal geometry. We may mitigate this hierarchy somewhat by placing the Higgs in another warped throat—the backreaction will then be blue-shifted, leading to an increased coupling g_h—though, as explained above, placing the Higgs in a warped region will naturally suppress M.

Finally, for generic initial conditions, the relaxion traverses a distance $\Delta\phi \sim M/g_h$ in field space. This is associated with the dissipation of

$$N \sim \frac{\Delta\phi}{f} \sim \frac{M^4}{g_h M^3 f} \sim g_s N_{\mathrm{D3}} \left(\frac{g}{g_h}\right) \left(\frac{M}{M_s}\right)^4 \left(\frac{\ell_s}{r_\cup}\right)^4 \tag{2.60}$$

units of monodromy charge.

Barriers from D7-branes

We will be more specific about how Ⓒ is realized. Perturbatively in g_s, the axion has a continuous shift symmetry $\phi \mapsto \phi + \text{const.}$ which is broken by non-perturbative effects (in g_s) to the discrete shift symmetry $\phi \mapsto \phi + f$. If Ⓐ and Ⓑ are to be the only terms that break this discrete shift symmetry, as is implicitly assumed in the relaxion construction, then V_{stop} must be generated non-perturbatively in g_s. As noted above, we take the Higgs and relaxion sectors to be separated in the internal geometry, and so in order for the stopping potential to depend on both of these sectors, it must be generated by physics on one or more extended objects—by either Euclidean Dp-branes or strong gauge dynamics on a stack of Dp-branes.

For simplicity, we will assume that $V_{\rm stop}$ is generated by the strong dynamics of a gauge theory, with group G, realized on a stack of D7-branes wrapping a holomorphic four-cycle Σ_4, as illustrated in Figure 2.4. The D7-branes couple to the C_2 axion through the Chern-Simons action

$$S_{\rm CS} \supset \mu_7 \int_{\mathcal{W}} \mathcal{F} \wedge C_2 \wedge \mathcal{F} \wedge \mathcal{F} \,. \tag{2.61}$$

A key observation is that the D7-branes must enter the warped throat region (see Appendix 2.B for a proof). The coupling (2.61) leads to a potential of the schematic form

$$V(\phi, v) = \Lambda_c^3 \, v \cos\left(\frac{2\pi\phi}{f}\right), \tag{2.62}$$

but can, in general, involve a more complicated polynomial of the Higgs vev v and a general f-periodic function in ϕ. The confinement scale Λ_c is naturally related to the string scale and the D7-brane gauge coupling $g_{\rm YM}$,

$$\Lambda_c^3 \propto \ell_s^{-3} \exp\left(-\frac{8\pi^2}{g_{\rm YM}^2 \, c_G}\right), \tag{2.63}$$

where c_G is a constant determined by the particular effects that generate (2.62). For example, c_G is simply the dual Coxeter number of G if the stopping potential is realized through gaugino condensation. In known examples, c_G is at most $\mathcal{O}(10^2)$, and we will take $c_G = 1$ henceforth. The generated hierarchy between the string and electroweak scales is then

$$\frac{M_s}{v} \propto g_s N_{\rm D3} \exp\left(-\frac{8\pi^2}{g_{\rm YM}^2}\right) \left(\frac{\ell_s}{r_\cup}\right)^4. \tag{2.64}$$

Since $r_\cup \gtrsim \ell_s$, the hierarchy is controlled by the warp factor and at first sight appears to be proportional to $N_{\rm D3}$. Thus, an arbitrarily large hierarchy could apparently be realized via substantially warping the source of monodromy. However, as we will discuss in §2.4, this is too naive.

2.4 Microphysical Constraints

Relaxation of a hierarchy by the relaxion mechanism occurs only in theories that meet several stringent requirements. Arguably the most challenging requirements from the viewpoint of ultraviolet completion in string theory are both the *large displacement* $\Delta\phi \sim M/g$, and the comparatively *short stopping length*. That is, the relaxion must evolve slowly over a large distance, gradually reducing the Higgs mass, but then rapidly come to rest after the Higgs acquires a vev. These disparate distance scales in field space correspond to very different energy scales in the potential: the final Higgs vev v is determined by the ratio of the shift symmetry breaking scale gM^3f to the stopping potential scale Λ_c^3, cf. (2.4). In field theory, one can obtain a controllably large hierarchy by taking g to be extremely small while holding Λ_c fixed.

This limit is problematic in string theory. As we will show in §2.4.2, the scale Λ_c depends on g, and is exponentially suppressed as $g \to 0$. This dramatically limits the hierarchy that can be generated.

At the same time, the large field excursion on its own implies that the initial configuration carries a very large monodromy charge. This gradually dissipating monodromy charge will serve as a changing source for the ten-dimensional equations of motion. For $N \gg 1$, this backreaction has profound effects on the compactification geometry and so on the four-dimensional relaxion potential (2.54).

It is tempting to argue that all of the corrections that result from backreaction must ultimately originate in the breaking of the axionic shift symmetry, and so must involve powers of the shift-symmetry breaking parameter g. This is *not* correct. In the NS5-brane model, the breaking parameter g is small because the

DBI action of an NS5-brane is proportional to the warp factor at the NS5-brane location, cf. (2.55). Backreaction effects sourced directly through the DBI action are indeed proportional to powers of g. However, the NS5-brane Chern-Simons action is not warped, and could not be: it is topological, and counts the (integer) D3-brane charge induced on the NS5-brane, i.e. the monodromy charge N.

Thus, backreaction effects sourced by the Chern-Simons action are proportional to N, without factors of g. For example, the integral of the R-R field strength F_5 over a Gaussian surface—say, an S^5—surrounding the NS5-brane is simply given by N, even as $g \to 0$. One consequence, as we shall see, is that the monodromy charge provides a large correction to the stopping potential.[14]

In this section we provide an array of calculations that reveal the concrete obstacles to achieving a large displacement and a short stopping length in the NS5-brane model.

2.4.1 Overview of microphysical constraints

We first preview a number of constraints on relaxion monodromy constructions, which originate from microphysical limitations on string compactifications that provide the desiderata listed in §2.3.1. Each of these constraints will be detailed in turn in §§2.4.2-2.4.2.

§2.4.2 **Universal effect on the geometry.** The shape of the warped throat region is dramatically altered by backreaction, leading to large changes in the

[14]The backreaction sourced by this topological term does not need to propagate far to be "detected," i.e. to influence a significant term in the four-dimensional Lagrangian: see Appendix 2.B.

effective action.

§2.4.2 **Tadpole constraints**. To accommodate $N \gg 1$ units of monodromy charge without the loss of perturbative control, we must construct a background throat with $N_{\rm D3} \gg N \gg 1$. Gauss's law—i.e., the D3-brane charge tadpole—then implies that there must be a source that is equivalent to $-N_{\rm D3}$ D3-branes. To avoid the instabilities created by a large number of actual anti-D3-branes, this source must be supersymmetric, and arise from the topology of an elliptically-fibered fourfold: the D3-brane charge is then $-\chi/24$, where χ is the Euler number of the fourfold. The largest known Euler number of an elliptically-fibered fourfold is 1,820,448. So in this setting, N will have to be much smaller than 75,852.

§2.4.2 **Barrier suppression from warping**. The D7-branes that generate the stopping potential must wrap a four-cycle Σ_4 that intersects the minimum volume two-cycle Σ_2. The D7-brane gauge coupling function, which depends on the *warped* four-volume of Σ_4, is then directly suppressed by the same warping responsible for the miniscule monodromy energy scale (2.55). A weakly broken shift symmetry therefore leads to extremely small barriers.

§2.4.2 **Barrier suppression from backreaction**. The induced D3-brane charge and tension backreact on the warped four-volume of Σ_4 and therefore perturb the D7-brane gauge coupling function. This perturbation introduces an exponential dependence of the gauge coupling on the monodromy charge N, with *no* powers of $g \sim N_{\rm D3}^{-1}$. This contradicts naive applications of technical naturalness: the dangerous term that arises is not negligible in the limit $g \to 0$ where the shift symmetry breaking is weak.

§2.4.2 **Effects on the moduli potential**. The sources responsible for Kähler moduli stabilization are exponentially sensitive to perturbations of the warp

factor. So the moduli potential depends on the relaxion field, i.e. there are new terms in the relaxion potential not captured by (2.54). This was extensively studied in [34].

§2.4.2 **Effects on the axion decay constant**. Large backreaction will also affect the axion decay constant, which depends on the volume of the cycle the axion threads as well as on the overall volume of the internal manifold.

§2.4.2 **Classical annihilation of the dipole**. The compactification detailed in §2.3.2 is metastable. The NS5-brane and anti-NS5-brane attract one another because of the induced D3-brane charge that each carries, but the fivebranes must stretch over a large-volume representative of $[\Sigma_2]$ in order to meet one another. This costs energy, because the fivebranes have tension. For modest windings N, the tension energy can be much larger than the Coulomb energy from the D3-branes and anti-D3-branes, and the system is controllably metastable. However, for $N \gg 1$, the Coulomb energy can overpower the tension energy, and the fivebrane/anti-fivebrane dipole can classically annihilate.

§2.4.2 **Constraints from anti-D3-brane annihilation**. An anti-D3-brane at the tip of a large Klebanov-Strassler throat is a metastable and cosmologically long-lived configuration. However, the barrier that ensures metastability depends on the number of anti-D3-branes in the throat. For some number $N_{\rm KPV}$ of anti-D3-branes—and thus for windings $N \geq N_{\rm KPV}$—the barrier disappears and the anti-D3-branes can classically annihilate against the flux of the throat. Thus, the accumulation of anti-D3-branes on the anti-NS5-brane creates a risk of instability.

§2.4.2 **Tunneling via light brane KK modes**. The accumulation of D3-branes in the NS5-brane pair leads to a reduction in the tension of the NS5-branes,

and correspondingly a reduction in the mass of Kaluza-Klein excitations of the NS5-branes. This Kaluza-Klein spectrum has spacing proportional to m_0/N when the axion is wound up by N cycles, with m_0 associated to the IR scale of the warped throat. These light brane KK modes provide another pathway for classical annihilation of the dipole. If the throat is put at some temperature, say from a source of supersymmetry breaking elsewhere in the internal space, thermal fluctuations of the light brane KK modes could enable the NS5-branes to reach up towards one another, allowing for a quantum mechanical tunneling event.

2.4.2 Consequences of D3-brane backreaction

D3-branes and anti-D3-branes source warping, and so the D3-brane dipole that develops when the axion is wound up leads to a change in the local warp factor. The warped throat region is itself produced by some number $N_{\rm D3}$ of D3-branes that have dissolved into flux, and when the number of windings N becomes comparable to $N_{\rm D3}$, the D3-brane dipole is a large correction to the background in which it is sitting. The probe approximation is not valid for such a configuration, and the backreaction of the D3-brane dipole affects many couplings in the four-dimensional theory.

Universal effect on geometry

The backreaction of the tension and charge of N induced D3-branes will be a small perturbation to the overall configuration as long as the ratio $g_s\ell_s^4 N/L^4$ is small, where L is the radius of the warped throat: see Appendix 2.D. Using $N_{\rm D3}$ to denote

the effective D3-brane charge of the warped throat (2.56) we must require that

$$N \ll N_{\rm D3}. \tag{2.65}$$

To intuitively motivate (2.65), we may replace the N D3-branes with an AdS_5 warped throat with radius

$$R_N^4 \sim g_s \ell_s^4 N \tag{2.66}$$

via a geometric transition. The perturbed geometry will be drastically different unless size of this extra throat is much smaller than the original warped throat, $R_N^4 \ll L^4$. So, we require that $N \ll N_{\rm D3}$ in order to maintain perturbative control.

The volume of the warped throat is necessarily bounded by the total volume of the internal space, $L^6 \lesssim \ell_s^6 \mathcal{V}_{\rm E}$.[15] which determines the four-dimensional Planck mass via $M_{\rm pl}^2 \ell_s^2 = 4\pi \mathcal{V}_{\rm E}$. From (2.65) we find the constraint

$$N \ll \frac{1}{g_s} \frac{M_{\rm pl}}{M_{\rm KK}}, \tag{2.67}$$

where $M_{\rm KK} = 1/(\ell_s \mathcal{V}_E^{1/6})$. This imposes a constraint on the number of windings for reasonable hierarchies between the compactification and Planck scales, and for reasonable values of g_s.

Tadpole constraint

The higher-dimensional equations of motion must be satisfied in a consistent string compactification. In particular, the higher-dimensional analog of Gauss's law for the five-form flux F_5 becomes a powerful constraint on the ten-dimensional configuration. In a non-compact manifold, flux lines are allowed to extend to infinity and Gauss's law places no constraint on the amount of charge allowed in a given configuration. However, in a compact manifold a flux line must end on a charge and

[15] We denote the total volume of the internal space X_6, measured in Einstein frame, as $\ell_s^6 \mathcal{V}_{\rm E}$,

Gauss's law provides a *tadpole constraint*: the total amount of D3-brane charge in the compactification must vanish. As discussed above, the warped throats pictured in Figure 2.3 are supported by a total of N_{D3} units of D3-brane charge. The tadpole constraint requires that this charge be canceled elsewhere in the Calabi-Yau geometry.

This cancellation could occur by including anti-D3-branes elsewhere in the internal space, or by forming another, oppositely charged, warped throat elsewhere with a large amount of negative D3-brane charge. In both cases, the D3-branes supporting the relaxion's warped throat and these additional anti-D3-branes will attract and the entire model will generically be unstable.

Fortunately, there exist well-known sources of supersymmetric negative D3-brane charge, and thus one may satisfy the tadpole constraint while maintaining stability. Seven-branes wrapping non-trivial cycles in the internal space provide curvature-induced negative D3-brane charge. F-theory compactified on elliptically-fibered Calabi-Yau fourfolds provides a framework for analyzing type IIB compactifications at arbitrary coupling, and the negative charge is related to the fourfold's Euler number $\chi(\mathrm{CY}_4)$ via

$$N_{\mathrm{D3}}^{\mathrm{CY}_4} = -\frac{\chi(\mathrm{CY}_4)}{24}. \tag{2.68}$$

The largest known Euler number of an elliptic-fibered Calabi-Yau fourfold is $\chi(\mathrm{CY}_4) = 1{,}820{,}448$ [36], which imposes the constraint

$$N_{\mathrm{D3}} \leq 75{,}852. \tag{2.69}$$

Requiring $N \ll N_{\mathrm{D3}}$ to maintain control over the configuration, we then have the constraint

$$N \ll 75{,}852. \tag{2.70}$$

The bound (2.69) on the Euler number of known fourfolds thus translates to a strong upper limit on the number of windings, and so constrains the maximum possible field excursion undergone by the relaxion.

The bound (2.70) applies only in the present case in which the monodromy charge is D3-brane charge. However, in alternative axion monodromy scenarios, it would still be necessary to arrange that the background solution at zero winding carries a large background monodromy charge analogous to N_{D3}. In such a setting we expect topological upper bounds analogous to (2.69) on the amount of monodromy charge that can be included without creating rapid instabilities.

Suppression from warping

Relaxation of a large hierarchy requires that the shift symmetry is very weakly broken, with $g \ll 1$. In the ten-dimensional model of §2.3.3, the breaking is made small by placing the source of monodromy—NS5-branes wrapping the minimum-volume two-cycles Σ_2 and $\overline{\Sigma}_2$—in a heavily warped region. However, we prove in Appendix 2.B that supersymmetric D7-branes can generate a relaxion stopping potential only if the four-cycle Σ_4 they wrap intersects Σ_2 or $\overline{\Sigma}_2$. So the D7-brane stack responsible for the stopping potential necessarily descends into the warped region. As we will now see, elementary locality arguments show that the small parameter associated with this warping, g, in Ⓐ and Ⓑ of (2.54) then generically infects the stopping potential Ⓒ realized on the D7-brane stack, leading to an *exponential suppression* of the stopping potential barriers.

The gauge coupling g_{YM} on a spacetime-filling D7-brane wrapping a four-cycle

Σ_4 is proportional to the *warped* four-volume of Σ_4 in string units,

$$\frac{1}{g_{\rm YM}^2} = \frac{1}{2\pi\ell_s^4}\int_{\Sigma_4} \mathrm{d}^4\xi\,\sqrt{\tilde{g}_4}\,e^{-4A}, \tag{2.71}$$

with $\tilde{g}_4$ the induced, unwarped metric on Σ_4. Defining a reference warp factor profile $\exp(4\bar{A}) = r^4/L^4$, cf. (2.56), we may express (2.71) as

$$g_{\rm YM}^{-2} = \alpha^{-1} g_s N_{\rm D3}, \tag{2.72}$$

where

$$\alpha^{-1} \propto \int_{\Sigma_4} \mathrm{d}^4\xi\,\sqrt{\tilde{g}_4}\,r^{-4}e^{-4(A-\bar{A})} \tag{2.73}$$

is a dimensionless coefficient capturing the geometry of the embedding of Σ_4 in the warped throat.

We may estimate α as follows. We have shown that Σ_4 must reach down the warped throat to intersect Σ_2 at $r_\cup$. Assuming that Σ_4 roughly factorizes into a radial part and an angular part with volume $\breve{v}$, that it extends up into the bulk geometry as in Figure 2.5, and that the integral is dominated in the region where $A \sim \bar{A}$, we find

$$\alpha^{-1} \gtrsim \breve{v}\,\log\left(\frac{L}{r_\cup}\right) \sim \breve{v}\,\log\left(\frac{g_s N_{\rm D3}\ell_s^4}{r_\cup}\right). \tag{2.74}$$

Importantly, α^{-1} is not naturally $\mathcal{O}(N_{\rm D3}^{-1})$, and in fact grows with the size of the throat, $L^4 \propto N_{\rm D3}$. So, $g_{\rm YM}^{-2} \sim \mathcal{O}(N_{\rm D3})$ unless the angular volume $\breve{v}$ is finely tuned to be exceptionally small, to one part in g^{-1}, which is of order the desired hierarchy. In other words, fine-tuning the angular volume $\breve{v}$ to eliminate the effects of this warping amounts to constructing the entire hierarchy by this fine-tuning. This suppression therefore renders the relaxation mechanism ineffectual.

From (2.63) and (2.72), the stopping potential is *exponentially suppressed* in $N_{\rm D3}$,

$$\Lambda_c^3 \propto \ell_s^{-3}\exp\left(-\gamma_{\rm bg}N_{\rm D3}\right), \tag{2.75}$$

with $\gamma_{\rm bg} \sim 8\pi^2/(g_s \alpha c_G)$. The hierarchy generated including this suppression is then

$$\frac{M_s}{v} \sim g_s \gamma_{\rm bg}^{-1} \left(\gamma_{\rm bg} N_{\rm D3} e^{-\gamma_{\rm bg} N_{\rm D3}}\right) \left(\frac{\ell_s}{r_\cup}\right)^4, \tag{2.76}$$

and since $xe^{-x} \leq e^{-1}$, the maximum resolvable hierarchy is simply

$$\frac{M_s}{v} \sim g_s \gamma_{\rm bg}^{-1} \sim \alpha\, c_G \tag{2.77}$$

which is, crucially, not $\mathcal{O}(g^{-1})$ unless α is severely fine tuned.

Generically, the warping responsible for the suppression of the shift symmetry breaking energy scale also suppresses the scale of the stopping potential. This suppression drives a runaway relaxion, and precludes the dynamical generation of a large hierarchy in the absence of an acute fine tuning.

We expect this suppression to be very general. We argued in §2.3.3 that the stopping potential must be generated by non-perturbative effects on a $(p+1)$-dimensional extended object, and Lorentz invariance requires this extended object to either fill spacetime and wrap an internal cycle ($p > 3$) or else be instantonic. For a Dp-brane wrapping a p-cycle Σ_p, the gauge coupling is given by

$$\frac{1}{g^2_{\rm YM,p}} = \frac{1}{2\pi \ell_s^{p+1}} \int_{\Sigma_p} {\rm d}^{p-3}\xi \, \sqrt{\tilde{g}_{p-3}}\, e^{(7-p)\Phi/4 - (p-3)A}. \tag{2.78}$$

Similarly, for a Euclidean Dp-brane wrapping the same cycle, the action is

$$S_{\rm EDp} = \frac{2\pi}{\ell_s^{p+1}} \int_{\Sigma_p} {\rm d}^{p+1}\xi \, \sqrt{\tilde{g}_{p+1}}\, e^{-(p-3)\Phi/4 - (p+1)A}. \tag{2.79}$$

Both depend on powers of e^{-A} and thus *positive* powers of $N_{\rm D3}$. So, any potential barrier generated by these effects will suffer from the same exponential suppression, albeit with different powers of $N_{\rm D3}$.

Suppression from backreaction

The backreaction of D3-brane charge is out of control unless the induced D3-brane charge N is a small fraction of the total D3-brane charge forming the throat, $N/N_{\rm D3} \ll 1$, so that we may perform a perturbative expansion of the ten-dimensional field configuration in this ratio. We should thus expect corrections to (2.54) to involve powers of $N/N_{\rm D3}$, which is consistent with the expectation that, because the monodromy charge is related to the shift symmetry breaking, any corrections due to backreaction will come dressed with powers of g. Crucially, however, it is *fractional* corrections to the field configurations—i.e. $\delta\varphi/\varphi$ for some field φ—that involve powers of $N/N_{\rm D3}$. If some quantity—say, a D7-brane gauge coupling function—also scaled with $N_{\rm D3} \propto g^{-1}$, the the absolute (additive) correction correction to this quantity is *not* necessarily small when $N/N_{\rm D3} \ll 1$.

Indeed, the monodromy charge induces a perturbation to (2.71),

$$\delta\left(\frac{8\pi^2}{g_{\rm YM}^2}\right) \sim g_s N_{\rm D3} \int_{\Sigma_4} \mathrm{d}^4\xi\, \sqrt{\tilde{g}_4}\, e^{-4(A-A_0)}\, r^{-4} \underbrace{\left(\frac{1}{2}\frac{\delta\tilde{g}_4}{\tilde{g}_4} - \frac{\delta e^{4A}}{e^{4A}}\right)}_{\mathcal{O}(N/N_{\rm D3})} \propto g_s N \equiv \frac{\gamma_{\rm br}\phi}{f}. \tag{2.80}$$

As discussed in detail in Appendix 2.D, the fractional perturbations are $\mathcal{O}(N/N_{\rm D3})$ and thus the entire perturbation to the gauge coupling is $\mathcal{O}(g_s N)$. We have again grouped specific geometric details into a coefficient $\gamma_{\rm br}$.

In the introduction, we gave an interpretation of this backreaction in terms of new light states entering the spectrum of the theory upon a monodromy $\phi \mapsto \phi + f$. Open/closed-string duality dictates that the supergravity (closed-string channel) correction (2.80) must match the one-loop correction to the gauge coupling $g_{\rm YM}$ calculated in the open-string channel. In the open-string picture of the configuration pictured in Figure 2.4, we are interested in the one-loop correction to the

SU(N_c) gauge theory living on the D7-brane stack wrapping Σ_4, in the presence of N D3-branes dissolved in the NS5-brane on Σ_2 and N anti-D3-branes dissolved in the anti-NS5-brane on $\bar{\Sigma}_2$. Crucially, the N D3-branes introduce N light 3-7 strings transforming in the fundamental of SU(N_c), which provide a contribution to the one-loop β-function (2.12).

Accounting for this backreaction changes the structure of the potential (2.54). In particular, from (2.80) the monodromy charge induces further relaxion-dependence of the height of the stopping potential barriers,

$$\Lambda^4(v) \to \Lambda^4(\phi) e^{-\gamma_{\rm br}\phi/f}. \tag{2.81}$$

A priori, it is not obvious that $\gamma_{\rm br}$ is either always positive or always negative, so we will consider $\gamma_{\rm br} > 0$ and $\gamma_{\rm br} < 0$ separately. Assuming that the Higgs quartic coupling takes the form

$$\mathcal{L}_h \supset -\frac{\lambda}{2}|h|^4, \tag{2.82}$$

v is given by

$$v(\phi) = \sqrt{\frac{gM}{\lambda}\,(\phi_{\rm h} - \phi)}, \tag{2.83}$$

where $\phi_{\rm h} = \phi_{\rm init} - M/g$ is generically $\mathcal{O}(M/g)$, and thus the corresponding induced monodromy charge when the Higgs develops a vev is very large, $N_{\rm h} \equiv \phi_{\rm h}/f \gg 1$. Ignoring the backreaction effect (2.80) and assuming that $f \ll \lambda v^2/gM$, the relaxion will stop rolling when

$$\frac{\Lambda_c^3}{f}\sqrt{\frac{gM}{\lambda}\,(\phi_{\rm h} - \phi)} \sim gM^3 \tag{2.84}$$

and it will be stabilized at

$$\phi_h - \phi \sim \frac{\lambda}{gM}\left(\frac{gM^3 f}{\Lambda_c^3}\right)^2. \tag{2.85}$$

If we now include the backreaction (2.80), (2.84) becomes

$$\frac{2\gamma_{\rm br}}{f}\,(\phi_{\rm h} - \phi)\, e^{2\gamma_{\rm br}(\phi_{\rm h}-\phi)/f} \sim \frac{2\lambda}{gMf}\frac{\gamma_{\rm br}}{(1+\gamma_{\rm br})^2}\left(\frac{gM^3 f}{\Lambda_c^3}\right)^2 e^{2\gamma_{\rm br}\phi_{\rm h}/f}. \tag{2.86}$$

Because $\phi_{\rm h}/f \gg 1$, the asymptotic behavior of solutions to (2.86) is determined solely by the sign of $\gamma_{\rm br}$. For $\gamma_{\rm br} > 0$, the stopping potential barriers are exponentially suppressed by the backreaction and (2.86) predicts that the relaxion stops at

$$\phi \sim -\frac{f}{2\gamma_{\rm br}} \log\left(\frac{2\lambda\gamma_{\rm br}}{(1+\gamma_{\rm br})^2}\frac{gM^3 f}{\Lambda_c^4}\frac{M^2}{\Lambda_c^2}\right) < 0. \tag{2.87}$$

However, the linear potential $gM^3\phi$ in (2.54) is only an approximation for a potential of the form (2.53) and cannot be used for arbitrarily small values of ϕ/f. From (2.87) we see that this approximation breaks down. We should therefore understand (2.87) as an indication that the relaxion stops roughly when it has dissipated all of its charge, near $\phi = 0$. The electroweak scale is then fixed at

$$v \sim \frac{gM}{\lambda}\phi_{\rm h} \sim \frac{M^2}{\lambda}, \tag{2.88}$$

leaving the hierarchy unresolved.

For $\beta < 0$, the barriers are exponentially *enhanced*, and the relaxion stops at

$$\phi \sim \phi_{\rm h} - \frac{\lambda}{gM}\left(\frac{gM^3 f}{\Lambda_c^3}\right)^2 \frac{1}{(1+\gamma_{\rm br})^2} e^{2\gamma_{\rm br}\phi_{\rm h}/f}, \tag{2.89}$$

and the electroweak scale

$$v \sim \frac{gM^3 f}{\Lambda_c^3}\frac{1}{|1+\gamma_{\rm br}|} e^{-|\gamma_{\rm br}|\phi_{\rm h}/f} \tag{2.90}$$

is suppressed by the backreaction.

Can one use this barrier enhancement to save the relaxion from the exponential suppression discussed in §2.4.2? Unfortunately, this backreaction enhancement is not enough to overcome the suppression from warping. We may combine (2.90) with (2.75) to find

$$\frac{M_s}{v} \sim \frac{g_s|1+\beta|}{(\gamma_{\rm bg} - |\gamma_{\rm br}|N_{\rm h}/N_{\rm D3})}\left(\frac{\ell_s}{r_\cup}\right)^4 \left[(\gamma_{\rm bg}N_{\rm D3} - |\gamma_{\rm br}|N_{\rm h})\, e^{-\gamma_{\rm bg}N_{\rm D3}+|\gamma_{\rm br}|N_{\rm h}}\right]. \tag{2.91}$$

Since we require that $N_{\rm h}/N_{\rm D3} \ll 1$ for control and we expect the geometric factors to be on the same order $\gamma_{\rm bg} \sim |\gamma_{\rm br}|$, (2.91) implies that the necessary fine-tuning is still of the same order as the hierarchy one wishes to generate.

Effects on the moduli potential

In §2.4.2 we considered the backreaction of D3-brane charge on the gauge coupling of the D7-branes that generate the stopping potential. As shown in Appendix 2.B, this particular D7-brane stack must enter the strongly warped region, and so the backreaction does not need to propagate far to impact them. The result is a very large change in the gauge coupling of the D7-brane worldvolume theory, leading to exponential suppression of the stopping potential.

Let us now ask about the impact of backreaction on the moduli potential. In the NS5-brane scenario, the Kähler moduli of the compactification are stabilized by nonperturbative effects on a collection of four-cycles, either Euclidean D3-branes or gaugino condensation on D7-branes. The moduli potential also involves exponentials of the warped volumes of these cycles. Backreaction of D3-brane charge will change the warped volumes of these cycles, and so the moduli potential will typically be a rapidly varying function of the relaxion ϕ.

The argument of Appendix 2.B does not imply that the four-cycles supporting the Kähler moduli potential enter the warped region, so in contrast to §2.4.2, the backreaction has to propagate across the internal geometry to influence the moduli potential. It is tempting to argue that backreaction has a negligible effect on a sufficiently distant four-cycle. This is *not* correct. We will give a heuristic explanation here, and refer the reader to [8] for a complete quantitative treatment.

To understand whether backreaction of D3-brane charge can decouple from D7-branes on a particular four-cycle Σ_4, we work in the open string picture, where the effect of backreaction is translated into the open string one-loop threshold correction to the D7-brane gauge coupling. On very general grounds, this effect gives non-negligible contributions to the relaxion potential unless the masses M_{3-7} of the stretched open strings obey

$$M_{3-7} \gtrsim M_{\rm pl}\,, \tag{2.92}$$

for then the non-renormalizable operators coupling the relaxion to the moduli are suppressed by more than the Planck mass. In a compact space, the diameter of the space determines an upper bound on the mass M_{3-7}, and one finds that at weak coupling and large volume, $M_{3-7} \ll M_{\rm pl}$ [8] (cf. also [37]). This is easily checked in simple geometries, but holds more generally.[16]

The upshot is that the four-cycles supporting the moduli potential cannot be taken far enough away from the source of monodromy to avoid significant backreaction: the moduli potential depends strongly on ϕ. One consequence is that the relaxion potential is *not* simply given by the probe DBI action (2.53), but instead has important contributions from couplings to moduli. This is an incarnation of the eta problem, which hinders the construction of natural models of inflation.

If all the other obstacles enumerated here could be overcome in some manner, leaving only the problem of relaxion couplings to moduli induced by backreaction, then one could attempt to fine-tune the orientation of the source of monodromy with respect to the configuration of four-cycles in the bulk of the compactification. The idea is that if the leading multipoles of the backreaction can be made to vanish

[16] This fact is responsible for the well-known problem that brane-antibrane potentials are generically too steep to support inflation [38].

on the four-cycle "receiver" by fine-tuning the relative orientation, then the residual effect of the subleading multipoles might be negligible. This approach was proposed and analyzed in [34], where it was shown that the backreaction coefficient $\gamma_{\rm br}$ goes as $(r_\cap/r_{\rm bulk})^m$, with m an $\mathcal{O}(1)$ integer determined by the lowest unsuppressed multipole, and that one could realize $\gamma_{\rm bg} \sim 10^{-2}$ even for relatively small hierarchies between $r_{\rm bulk}$ and $r_\cap$ with moderate fine-tuning. So, for a modest winding number $N \sim 100$, this backreaction on the moduli potential can be ameliorated and the most dangerous couplings can be removed. It is not clear, however, that this method is applicable for the extremely large windings $N \gtrsim 10^6$ that arise in relaxion constructions. Indeed, even if the intersection argument of Appendix 2.B were somehow avoidable and the object generating the relaxion stopping potential could be localized in the bulk of the compactification, $\gamma_{\rm bg} \sim 10^{-6}$ could not be realized without taking $r_\cap \ll r_{\rm bulk}$, which suppresses (c.f. Eq. (2.57)) the relaxion decay constant f and potentially renders the compactification unstable.

Effects on axion decay constants

As discussed in Appendix 2.A, the relaxion decay constant f only depends on the six-dimensional metric $\tilde{g}_{mn}$, both through the explicit factors of $\tilde{g}_{mn}$ in its definition (2.127) and implicitly via the defining equation of the harmonic form $\Delta\Omega = 0$. Additional D3-brane charge will not perturb f: $\tilde{g}_{mn}$ is Ricci-flat in supersymmetric compactifications and additional D3-brane charge will preserve the same supercharges as the three-branes forming the warped throat. However, anti-D3-branes break the remaining supersymmetry. The anti-D3-brane charge

backreacts on the six-dimensional metric and perturbs the axion decay constant,

$$\frac{\delta f^2}{M_{\rm pl}^2} = \frac{g_s}{2\mathcal{V}_E \ell_s^6} \int \mathrm{d}^6 y \sqrt{\tilde{g}}\, \tilde{g}^{mp}\, \tilde{g}^{nq} \times \left(2\Omega_{mn}\delta\Omega_{pq} + \frac{1}{2}\tilde{g}^{rs}\delta\tilde{g}_{rs}\,\Omega_{mn}\Omega_{pq} - 2\Omega_{mn}\Omega_{ps}\delta\tilde{g}_{qr}\tilde{g}^{rs} \right). \quad (2.93)$$

The anti-D3-brane charge does not substantially perturb the four-dimensional Planck mass $M_{\rm pl}^2$, as $\mathcal{V}_{\rm E}$ is dominated by the volume of the bulk Calabi-Yau. In what follows, we estimate the size of each of these terms.

The first term in (2.93) vanishes at first order, as the perturbation $\delta\Omega$ is orthogonal to the unperturbed Ω. To analyze the contribution from the second and third terms in (2.93), we must backreact the anti-D3-brane charge on the metric $\tilde{g}_{mn}$. As detailed in Appendix 2.D, the dominant metric perturbations are

$$\tilde{g}_{mn}\mathrm{d}y^m\,\mathrm{d}y^n \sim \left(1 + \frac{\alpha N}{N_{\rm D3}}\left(\frac{r'}{r}\right)^8\right)\mathrm{d}r^2 + r^2\left(1 + \frac{\beta N}{N_{\rm D3}}\left(\frac{r'}{r}\right)^{19/2}\mathcal{Y}^{\frac{1}{2},\frac{1}{2},1}(\Psi)\right)\breve{g}_{\theta\phi}\,\mathrm{d}\xi^\theta\,\mathrm{d}\xi^\phi \quad (2.94)$$

where the coefficients α and β and the angular function $\mathcal{Y}^{\frac{1}{2},\frac{1}{2},1}(\Psi)$ depend on the details of the compactification. The perturbation to the decay constant (2.93) is then

$$\frac{\delta f^2}{M_{\rm pl}^2} \propto \frac{g_s}{2\mathcal{V}_E\ell_s^6}\frac{N}{N_{\rm D3}}\int_{r_\cup}^{r_\cap}\mathrm{d}r\, r\,\left(\frac{r_\cup}{r}\right)^{19/2} \propto \frac{f^2}{M_{\rm pl}^2}\frac{N}{N_{\rm D3}}\left(\frac{r_\cup}{r_\cap}\right)^{19/2}. \quad (2.95)$$

Because of the large exponent this is a comparatively weak constraint.

Classical annihilation of the dipole

The fivebrane configuration detailed in §2.3.2 is metastable. If the fivebranes were tensionless, the induced D3-brane charge on the NS5-brane would attract the

induced anti-D3-brane charge on the anti-NS5-brane, and these branes would classically annihilate. However, this Coulomb attraction is balanced by the fivebrane tension—in order for the fivebranes to meet, they must stretch over the large two-cycle $\Sigma_\cap$ at the junction of the two warped throats, $r_\cap$ in Figure 2.3, which costs an energy

$$V_t \sim \frac{2\pi}{\ell_s^4}\frac{e^{4A_\cap}}{\sqrt{g_s}}\sqrt{4\left(\frac{\mathrm{vol}\,\Sigma_\cap}{\ell_s^2}\right)^2 + N^2}. \tag{2.96}$$

This potential energy barrier ensures the configuration is metastable, and can be exponentially long-lived. However, for large enough winding, we expect the Coulomb force—which scales as N^2—to overpower this "tension force," allowing the fivebranes to classically annihilate.

The potential energy density of a probe D3-brane is proportional to Φ_-,

$$V = \frac{2\pi}{\ell_s^4}\left(e^{4A} - \alpha\right) = \frac{2\pi}{\ell_s^4}\Phi_-. \tag{2.97}$$

For N D3-branes and N anti-D3-branes, the Coulomb potential energy density is then

$$V_\mathrm{c} \sim \frac{2\pi N}{\ell_s^4}\delta\Phi_-, \tag{2.98}$$

where $\delta\Phi_-$ (cf. Appendix 2.D) is the perturbation to Φ_- due to N anti-D3-branes, measured at the location of the N D3-branes. Because the D3-branes live in a separate warped throat, $\delta\Phi_-$ must first propagate up the antibrane throat from the anti-NS5-brane location $r = r_\cup$ to the surface $r = r_\cap$,

$$\delta\Phi^{I_s}_{-,\overline{\mathrm{D3}}} \sim -\frac{N}{N_{\mathrm{D3}}}e^{4A_\cup}\left(\frac{r_\cap}{r_\cup}\right)^{2+\Delta_s}, \tag{2.99}$$

which then propagates down the D3-brane warped throat via the homogeneous modes

$$\delta\Phi^{I_s}_{-,\mathrm{D3}} = c_1\left(\frac{r_\cap}{r}\right)^{\Delta_s+2} + c_2\left(\frac{r}{r_\cap}\right)^{\Delta_s-2}. \tag{2.100}$$

We may think of the perturbation (2.99) as specifying a boundary value for the perturbation (2.100). Generically, we have

$$c_1, c_2 \sim \frac{N}{N_{\rm D3}} e^{4A_\cup} \left(\frac{r_\cap}{r_\cup}\right)^{2+\Delta_s} \tag{2.101}$$

so that, at the position of the D3-brane charge,

$$\delta\Phi^{I_s}_{-,\rm D3} \sim -\frac{N}{N_{\rm D3}} e^{4A_\cap} + \frac{N}{N_{\rm D3}} e^{4A_\cap} \left(\frac{r_\cap}{r_\cup}\right)^{2\Delta_s} . \tag{2.102}$$

We then find that the Coulomb energy is roughly

$$V_c \sim -\frac{2\pi}{\ell_s^4} \frac{N^2}{N_{\rm D3}} e^{4A_\cap} . \tag{2.103}$$

Requiring that this be much less than the potential energy barrier (2.96) yields the constraint

$$N \ll \frac{N_{\rm D3}}{\sqrt{g_s}} + \frac{2\,\mathrm{vol}\,\Sigma_\cap}{\ell_s^2} + \mathcal{O}\left(\frac{1}{N_{\rm D3}^2}\left(\frac{\mathrm{vol}\,\Sigma_\cap}{\ell_s^2}\right)^4\right) \tag{2.104}$$

We should also account for the interaction energy between the pair of fivebranes. By performing an open string computation in an unwarped toroidal orbifold, [39] found a potential contribution that grows *logarithmically* with the fivebrane separation, and argued that this would apply to warped geometries, with energy scale set by $r_\cap$. While it is not entirely clear that this logarithmic behavior arises in the actual NS5-brane configuration described in §2.3, the corresponding potential energy contribution would take the schematic form

$$V_{5\bar{5}} \sim \frac{2\pi}{\ell_s^4} e^{4A_\cap} N_{\rm NS5}^2 \log\left(\frac{L}{r_\cap}\right) \sim \frac{2\pi}{\ell_s^4} e^{4A_\cap} . \tag{2.105}$$

We can ensure that this energy is much smaller than the uncharged tension energy (2.96) by imposing

$$\mathrm{vol}\,\Sigma_\cap \gg \sqrt{g_s}\ell_s^2 , \tag{2.106}$$

which is necessary in any case to ensure the validity of the supergravity approximation.

Throughout this work we have taken the homology class $[\Sigma_2]$ wrapped by the NS5-brane to be localized in the warped throat, as in Figures 2.3 and 2.4. That is, we assumed that the harmonic two-form dual to $[\Sigma_2]$ is principally supported in the warped region, and every holomorphic representative of $[\Sigma_2]$ is in the warped region. This localization is automatic in the particular construction given in [34], but should also be required in alternative constructions. A key reason is the fivebrane potential energy (2.105): if the lines of three-form flux stretching from the NS5-brane to the anti-NS5-brane passed through an unwarped region, the overall scale of supersymmetry breaking would exceed the string scale, by (2.105), and immediately destabilize the moduli.

Antibrane tunneling and annihilation

Consider a Klebanov-Strassler throat that arises from $N_{\rm D3}$ D3-branes probing a conifold with $M_{\rm KS}$ D5-branes wrapping the shrinking two-cycle. We take $N_{\rm D3} = M_{\rm KS}K_{\rm KS}$; then $K_{\rm KS}$ is the number of units of H_3 flux on the B-cycle.

If N anti-D3-branes are placed at the tip of this throat, they create a metastable, exponentially long-lived state provided that $N \lesssim 0.08M_{\rm KS}$ [40]. With more anti-D3-branes, $N \gtrsim 0.08M_{\rm KS}$, the anti-D3-branes rapidly annihilate [40] against the flux supporting the warped throat, decaying to the state with $K'_{\rm KS} = K_{\rm KS} - 1$. The D3-brane charge carried by flux is then $N'_{\rm D3} = N_{\rm D3} - M_{\rm KS}$, and $M_{\rm KS} - N$ D3-branes appear, but no anti-D3-branes remain.

While $N_{\rm D3}$ sets the overall scale of warping in the throat, $M_{\rm KS}$ sets the warp

factor at the tip,

$$e^{A}\big|_{\cup} \sim \exp\left(-\frac{2\pi N_{\rm D3}}{3g_s M_{\rm KS}^2}\right) \tag{2.107}$$

and

$$g_s M_{\rm KS}^2 \lesssim N_{\rm D3} \tag{2.108}$$

if the warping is non-negligible. To avoid the KPV instability [40], the number of windings cannot exceed

$$N \ll 0.08\, g_s^{-1/2} N_{\rm D3}^{1/2}. \tag{2.109}$$

Light NS5-brane modes

As discussed in §2.3.2, dimensional reduction of the transverse fluctuations in the NS5-brane's position yields Kaluza-Klein excitations whose mass *decreases* as the relaxion is wound up. Intuitively, we may interpret the presence of two-form flux as increasing the effective volume of the NS5-brane. Since Kaluza-Klein masses will inversely scale with this effective volume, we should expect some modes to become light at large windings. As shown in Appendix 2.C.2, to second order the canonically normalized fluctuations are described by the action

$$\begin{aligned} S_{\rm NS5}^{(2)} = \int {\rm d}^4x\, \sqrt{-g_4}\Bigg(& -V(c) - \frac{1}{2} g^{\mu\nu}\partial_\mu Y_I^{\hat{\imath}} \partial^\mu Y_{\hat{\imath}}^I - \frac{1}{2} m_I^2(c) Y_I^{\hat{\imath}} Y_{\hat{\imath}}^I \\ & + g(c)\,(c\partial_\mu c)\left(Y_I^{\hat{\imath}}\partial^\mu Y_{\hat{\imath}}^I\right) - \frac{1}{2} g(c)^2\,(\partial c)^2\, Y_I^{\hat{\imath}} Y_{\hat{\imath}}^I \Bigg) \end{aligned} \tag{2.110}$$

with

$$m_I^2 \sim \frac{4}{g_s \ell_{\rm E}^2} \frac{e^{2A_\cup}}{N^2} \left(\frac{\ell_{\rm E}}{\ell_s}\right)^4 \quad \text{and} \quad g(N) \sim \frac{1}{2N^2} \tag{2.111}$$

at large winding $N^2 \gg 4(\ell_{\rm E}/\ell_s)^4/g_s$, where $\ell_{\rm E}^2$ is the Einstein-frame volume of the two-cycle Σ_2 and the λ_I are eigenvalues of Σ_2's Laplacian, labeled by the multi-index I. As discussed in the introduction, the appearance of light states is generic

in realizations of monodromy in string theory, and one must ensure that these do not drastically affect the phenomenology.

The presence of $\mathcal{O}(N)$ light states in the spectrum, including 3-7 strings and KK excitations of the NS5-branes, can have a range of consequences. For example, modes with mass $m < 3H/2$ can fluctuate during inflation, storing energy and potentially impacting the late-time perturbations. Here we will examine just one effect of the KK modes, which is an enhanced probability of NS5-brane annihilation.

The masses of the canonically normalized fluctuations of the NS5-brane embedding (2.110) decrease with N, and one might worry that these light modes facilitate an additional instability. For example, if these modes are thermally excited by some source of supersymmetry breaking elsewhere in the compact space, then the NS5-branes can more readily reach each other and either classically or quantum-mechanically annihilate.

A complete analysis of this process is beyond the scope of this work. We will instead use an approximate criterion for the onset of instability. The dominant instanton in the four-dimensional field theory responsible for the transition between the metastable and stable states (i.e., the states with and without the NS5/anti-NS5-brane dipole, respectively) will be SO(4)-symmetric, with radius determined by

$$R_* \sim \frac{T_{\rm D}}{\Delta V}, \tag{2.112}$$

where, in the thin-wall approximation, $T_{\rm D}$ is the tension of the domain-wall interpolating between the two vacua and ΔV is their difference in energy. We then assume a loss of control when the typical thermal fluctuations of a spatial region of size R_* are comparable to the distance between the two fivebranes, $r_{\rm RMS} \sim r_\cap$.

The difference in energies, in the probe approximation, is the potential energy contribution from the NS5-branes,

$$\Delta V = \frac{2\pi}{\ell_s^4} N\, e^{4A_\cup} \tag{2.113}$$

while the tension of the domain wall is determined by an NS5-brane winding N times around the minimum-volume three-cycle whose endpoints are Σ_2 and $\bar{\Sigma}_2$. As described in Appendix 2.C.2, the tension of the domain wall follows from the NS5-brane action and is roughly

$$T_{\rm D} \sim \frac{1}{\ell_s^3} N_{\rm D3}^{3/4} \left(e^{4A_\cap} - e^{4A_\cup}\right) \tag{2.114}$$

and so

$$R_* \sim \ell_s \frac{N_{\rm D3}^{3/4}}{N} \left(\frac{r_\cap}{r_\cup}\right)^4 \tag{2.115}$$

If the NS5-brane is in thermal equilibrium at temperature T, then a smooth excitation of size R_* in the canonically normalized fluctuations $Y_I^{\hat{\imath}}$ gains a thermal expectation value

$$\langle Y_I^{\hat{\imath}} Y_J^{\hat{\jmath}} \rangle \sim \frac{T}{m_I^2} \frac{1}{R_*^3} \delta_{IJ} \delta^{\hat{\imath}\hat{\jmath}}. \tag{2.116}$$

The thermal fluctuation in the radial direction, averaged over Σ_2, is roughly

$$\langle \delta r^2 \rangle \sim g_s \ell_s^2 \, (\ell_s T) \, \frac{N^4}{N_{\rm D3}^{3/4}} \left(\frac{\ell_s}{\ell}\right)^2 \left(\frac{r_\cup}{r_\cap}\right)^{12} \tag{2.117}$$

where ℓ^2 is the unwarped volume of the two-cycle Σ_2. The requirement that this is much smaller than the size of the dipole $\langle \delta r^2 \rangle \ll r_\cap^2$ imposes the weak constraint

$$N^4 \ll e^{4A_\cap} \frac{N_{\rm D3}^{5/4}}{g_s^{1/2} (\ell_s T)} \left(\frac{\ell_{\rm E}}{\ell_s}\right)^2 \left(\frac{r_\cap}{r_\cup}\right)^{10}. \tag{2.118}$$

2.5 Discussion and Outlook

We have identified many obstacles to realizing the relaxion mechanism in string theory. Some of these obstacles are extremely general, while others apply only

to NS5-brane monodromy, the particular example we studied in detail. We will now step back and give some perspective on our results, explaining their scope of validity.

Our first observation was that axion monodromy in string theory proceeds by the accumulation of monodromy charge, and the backreaction of this charge substantially changes the couplings of the axion. This applies to any realization of axion monodromy in string theory. Thus, any ultraviolet completion in string theory of a relaxation mechanism that involves axion displacements $\Delta\phi > f$ will be vulnerable to the backreaction of monodromy charge.

The effects of this backreaction will vary from one model to another. We focused on NS5-brane monodromy because this is, to our knowledge, the only scenario where the smallness of the shift symmetry breaking parameter g is natural—in this case, because of warping—while in alternative constructions in string theory, one must fine-tune discrete data to achieve small g. In the NS5-brane model, we found that the barriers in the stopping potential are exponentially small in the winding number $N \equiv \phi/f$, leading to a runaway relaxion. We expect this barrier suppression phenomenon to be rather general, but not universal. However, the particular effects of backreaction on the axion decay constants detailed in §2.4.2, and the constraints from annihilation in §2.4.2 and §2.4.2, could be very different in other models.

Some of the challenges that we have identified might be milder in non-supersymmetric compactifications of string theory. In particular, in compactifications that break all supersymmetry at the Kaluza-Klein scale as in e.g. [41, 42], tadpole constraints on the total charge need not be a serious limitation. On the other hand, ensuring metastability of such a configuration can be very challenging.

Moreover, for an embedding of the relaxion in a non-supersymmetric compactification, the absence of spacetime supersymmetry below the KK scale might require either the KK scale or even the string scale to arise as the regulator of the relaxion setup at the relaxion cutoff scale M.

We assumed that the periodic stopping potential arises from non-perturbative effects that couple locally to the axion. This local coupling then exposes the stopping potential to an exponential suppression from warping. However, the stopping mechanism could instead arise from other effects, for example from heavy states [43] coupled to the relaxion, or from the exponential production of massive particles [17], which are not necessarily susceptible to the same failure modes.[17]

2.5.1 Exact discrete shift symmetries for relaxions

Throughout this work we have considered axion monodromy, in which a source of monodromy completely breaks the shift symmetry of an axion. An important alternative is *alignment* of multiple periodic contributions to the axion potential, leaving an unbroken discrete shift symmetry. We now briefly outline this possibility and mention some of the obstacles to embedding this scenario in string theory.

The essential feature of the relaxion potential (2.1) is the combination of a slowly-varying term Ⓑ, the linear $gM^3\phi$, and a quickly-varying term Ⓒ, the oscillatory $\Lambda_c^3 v\cos(2\pi\phi/f)$. As written, Ⓑ explicitly and completely breaks the discrete shift symmetry $\phi \mapsto \phi + f$. Alternatively, Ⓑ could represent the leading term in the expansion of a function that is invariant under a much larger discrete

[17] We thank E. Silverstein for illuminating discussions of these points.

shift $\phi \mapsto \phi + kf$. For example, we could have

$$V_{\text{Ⓑ}} = gfM^3 \sin\left(\frac{2\pi\phi}{kf}\right), \tag{2.119}$$

where ϕ's field space diameter is actually k times larger than would be naively inferred by only considering small displacements. We refer to these two cases as having an *explicitly broken symmetry* or an *exact discrete symmetry*, respectively. Thus far, we have only concentrated on the former. The explicit breaking is induced by a source of monodromy, an NS5-brane, and we have shown that the accumulation of monodromy charge leads to backreaction effects that spoil the relaxation mechanism. Given this difficulty, one might ask whether the relaxion mechanism could be more readily realized in a solution of string theory with an exact discrete shift symmetry.

As in the models with explicit breaking, the main difficulty in realizing a discrete shift-symmetric relaxion lies in ensuring that the potential has structure over two—and only two—disparate scales. That is, the potential must roughly be the sum of two terms—a slowly varying term with periodicity f that apes the linear term Ⓑ in (2.1), and a quickly varying stopping potential with periodicity $f_s \ll f$. One might take, as a toy model, a potential that is generated only by the instantons with winding number 1 and k, so that the potential takes the schematic form

$$V = M^2 e^{-S_1} \cos\left(\frac{2\pi\phi}{f}\right)|h|^2 + M^4 e^{-S_1}\cos\left(\frac{2\pi\phi}{f}\right) + M^3 e^{-S_k} v \cos\left(\frac{2\pi k\phi}{f}\right) \tag{2.120}$$

For $\phi \ll f$, the potential is approximately a "monomial with modulations" with $f_s = f/k$, and has the Ⓐ Ⓑ Ⓒ structure of (2.1), with $g \propto e^{-S_1}$ and $\Lambda_c^3 \sim M^3 e^{-S_k}$. The analogue of (2.4) in this two period model is then

$$\frac{v}{M} \gtrsim \frac{1}{k}\left(\frac{M}{\Lambda_c}\right)^3 \sim \frac{1}{k} e^{S_k - S_1}. \tag{2.121}$$

Naively, the generated hierarchy grows with k.

However, there are many problems with this toy model. First and foremost, the action for a k-instanton is typically $S_k \geq kS_1$, and we require that $S_1 \gg 1$ in order to trust the instanton expansion. The stopping potential barriers will then shrink with k,

$$|V_{\text{stop}}| \propto e^{-kS_1} \propto g^k \tag{2.122}$$

and, reminiscent of the suppression due to warping discussed in §2.4.2, the maximum achievable hierarchy actually shrinks with k. The stopping potential is too small to stop the evolution near the point where the Higgs is massless. If some mechanism were able to enhance the k-instanton contribution, one must still explain the absence of j-instanton effects, with $1 < j < k$, and we find it implausible that all such effects could be negligible for $k \gg 1$.[18] Furthermore, one would have to explain why the Higgs couples to instantons of winding 1 and k differently, and why the 1-instanton and k-instanton contributions do not *both* vanish when $v = 0$.

Many of these problems may be mitigated in models with multiple axions, as in the Kim-Nilles-Peloso mechanism [44] and kinetic alignment setups [45, 46] in the inflationary context. Scenarios for aligned relaxions have been presented in [47, 48, 49].[19] A general multi-axion Lagrangian can be written (cf. e.g. [46]) as

$$\mathcal{L} = -K^{ij}\partial\phi_i\partial\phi_j - \sum_a \Lambda_a^4 \exp\left(-Q_a{}^i S_i\right)\left[1 - \cos\left(2\pi Q_a{}^i \phi_i\right)\right], \tag{2.123}$$

where K^{ij} is a positive definite kinetic matrix of real numbers, while $Q_a{}^i$ is a charge matrix containing integers. We imagine that there are "slow" and "fast" linear combinations of canonically normalized axions with effective decay constants

[18] Such a situation appears to conflict with the lattice form of the Weak Gravity Conjecture, but is already implausible regardless.

[19] See also [50] for a recent discussion of naturalness constraints on such scenarios.

f and $f_s \ll f$, respectively. The addition of another direction in field space solves several of the problems mentioned previously, at the cost of introducing much more complicated dynamics.

Foremost among the advantages is that the stopping potential is no longer necessarily suppressed. In the single axion model, the hierarchy between f_s and f—and so between v and M—was generated by a hierarchy in the charge matrix $Q_a{}^i$, and a high charge contribution is exponentially suppressed relative to a low charge contribution. In a multi-axion model, $f_s \ll f$ may instead be realized in the kinetic matrix K^{ij}, and this does not impose an exponential hierarchy in the associated barrier heights. Of course, the hierarchy in the kinetic matrix must then be explained, but it is much easier to realize a hierarchy in the eigenvalues of a matrix of real numbers than in a matrix of bounded integers, and one does not need to explain why instantons with winding j, $1 < j < k$, do not contribute.

A very mild degree of alignment has been demonstrated in explicit examples [51], but whether alignment can yield large effective axion decay constants in string theory is an important open question, even for the $\mathcal{O}(100)$ enhancements that could suffice for inflation. It is not obvious to us that the vastly larger enhancements needed for a relaxion scenario are possible in known compactifications. For example, the "clockwork" mechanism [52, 48] requires a specific matrix of axion charges of instantons, and it remains to be seen whether this particular pattern of charges can arise in string theory. However, it is very plausible that linearly independent combinations of axions couple differently to the Higgs.

In summary, relaxion scenarios with exact discrete symmetries, built on the alignment of multiple instanton effects for one or more axions, are qualitatively different from the axion monodromy scenarios, with explicitly broken symmetry,

considered in this work. However, *both* classes of models are vulnerable to ultraviolet physics. Axion monodromy scenarios suffer from the backreaction of monodromy charge, as we have explained. Aligned scenarios could avoid this problem, but require extremely special axion charges. These charges are ultimately topological data dictated by the ultraviolet theory, and it is not clear that string theory allows strong enough alignment to permit relaxation of a large hierarchy. Furthermore, these multi-axion models have much more complicated dynamics, and it is not clear that the dynamical generation of a large hierarchy can proceed in a robust way.

2.5.2 Constraints from the Weak Gravity Conjecture

The Weak Gravity Conjecture (WGC), a class of conjectures asserting that gravity must be the weakest force [53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63], leads to (still conjectural) constraints on axion theories. One could therefore ask whether the WGC constrains relaxion monodromy scenarios. It does [22, 23], as we will briefly explain, but the known WGC constraints are far weaker than the limitations we have exposed in this work, which are independent of WGC considerations.

The WGC constrains monodromy scenarios by placing upper limits on the tension of domain walls. In the four-dimensional description of axion monodromy due to Kaloper, Lawrence, and Sorbo [26, 27, 25], Brown-Teitelboim domain walls connect different branches of the scalar potential. At the same time, when instanton effects lead to modulations of the axion potential, distinct critical points are connected across four-dimensional field theory domain walls, via Coleman-de Luccia tunneling. It turns out that the electric form of the WGC places constraints [22] on the domain walls of the Kaloper-Lawrence-Sorbo model, while the mag-

netic WGC places constraints on the field theory domain walls associated with instanton modulations [23]. In both cases one finds a bound on the domain wall tension [22, 23]

$$T < mfM_{\rm pl}, \tag{2.124}$$

where m is the mass of the axion. For a relaxion model this implies a bound on the relaxion cutoff scale M of roughly the same order as the constraints already given in [16].

We conclude that the constraints arising from very general four-dimensional quantum gravity considerations, such as the WGC, do not automatically capture all of the effects of actual embeddings in quantum gravity. Examining such embeddings is therefore crucial for assessing the viability of the relaxion mechanism in string theory.[20]

2.6 Conclusions

Could a portion of the observed hierarchy between the weak scale and the Planck scale be a consequence of dynamical relaxation of the Higgs mass during cosmological evolution? This striking idea is the core of the relaxion mechanism [16]. In this scenario, the relaxation of the Higgs mass is driven by the slow evolution of an axion field, the relaxion, whose shift symmetry is very weakly broken by a potential term that introduces monodromy. After relaxation over many cycles of monodromy, the Higgs mass passes through zero, causing barriers to appear in

[20] We note that this view concurs with the results of [25], where some leading effects of backreaction on the axion Lagrangian are captured by a series of higher powers of gauge-invariant field strengths, whose coefficients must necessarily be determined in the UV theory, and in which strong backreaction effects can drive one far from the "natural" bottom-up estimates.

the axion potential, and so halt the evolution. In effective field theory, the hierarchy that is generated is determined by the weak breaking parameter, and so is technically natural.

In this work we asked whether the relaxion mechanism survives ultraviolet completion in string theory. Do the essential components for the scenario exist in a well-controlled compactification, and do these components work in concert in string theory as they do in effective field theory?

We found that the key components of the scenario can indeed be realized in string theory. The mechanism of axion monodromy, first developed in the context of large-field inflation in string theory, can produce—in the probe approximation—the secular relaxion potential needed for slow relaxation over many fundamental axion periods. Moreover, the extremely low scale of the secular potential required for the relaxion mechanism can be explained by situating the source of monodromy in a strongly warped region. This is possible in one known scenario for axion monodromy in string theory, the NS5-brane model of [32], in which two-form axions acquire their potential from NS5-branes wrapping curves in a warped region.

However, our main result is that the structures required for monodromy in string theory present formidable and very general obstacles to a successful relaxion scenario in string theory. Monodromy proceeds by the accumulation of monodromy charge on a source of monodromy. As the relaxion rolls over N fundamental axion periods, it necessarily accumulates or discharges N units of monodromy charge. This large quantity of monodromy charge sources backreaction in the internal space, completely invalidating the probe approximation, and changing the couplings in the effective theory. The impact of monodromy charge is visible in a dual description as the appearance of N new light states.

We argued that the backreaction of monodromy charge can lead to disastrously large changes to the secular potential in *any* realization of the relaxion scenario via axion monodromy in string theory. In the specific case of the NS5-brane model, we computed the detailed form of these changes. The accumulation of monodromy charge suppresses the gauge coupling of the D7-brane gauge theory that generates the stopping potential. In the dual description, the N light states are charged under the D7-brane gauge group, and give a large threshold correction to the gauge coupling. The result is that the stopping potential is suppressed by a factor $\exp(-\gamma_{\rm br} N)$, where $\gamma_{\rm br}$ is a constant determined by the geometry. The stopping potential is therefore completely negligible, and cannot halt the evolution when the Higgs mass passes through zero. The Higgs mass indeed relaxes to smaller values, but this process continues far into the tachyonic regime.

While this detailed analysis was performed in the context of a specific model, we repeat that our findings are generic and are expected to apply to any model that relies upon a monodromy over many fundamental axion periods, regardless of the stopping mechanism. However, these constraints do not apply when the discrete shift symmetry remains unbroken, e.g. when the large field range is realized by the alignment of multiple axions.

In summary, we have shown that the physics of ultraviolet completion in string theory does not decouple from the dynamics of the relaxion mechanism. Our results do not exclude the dynamical relaxation of hierarchies in string theory, but in our view they do exclude technically natural dynamical relaxation driven by axion monodromy. It would be valuable to understand whether some of the difficulties we have uncovered result from limitations in existing constructions, or if instead they are consequences of general structures in quantum gravity.

2.A Axions in String Theory

There is an extensive literature on axions in string theory, but for the reader's convenience we now gather a few salient facts. We begin with the example of the Neveu-Schwarz two-form gauge potential B_2.

Integrating a ten-dimensional p-form gauge potential C_p over a non-trivial p-cycle in the compactification manifold will give rise to an axion in four dimensions. The number of independent, non-trivial p-cycles then determines the maximal number of axions arising from C_p. The two-form B_2 has an associated field strength $H_3 \equiv \mathrm{d}B_2$, and appears in the ten-dimensional type II and heterotic supergravity actions as[21]

$$S_{\text{SUGRA}} \supset -\frac{1}{4\kappa_{10}^2}\int \mathrm{d}^{10}X\,\sqrt{-G^S}\,e^{-2\Phi}\,|H_3|^2 \tag{2.125}$$

Reducing this action along a six-dimensional compact space X_6, each non-trivial two-cycle Σ_2^I with its associated harmonic form ω_2^I, $\ell_s^{-2}\int_{\Sigma_2^I}\omega_2^J = \delta^J{}_I$, gives rise to a four-dimensional axion $b_I(x)$,

$$b_I(x) \equiv \frac{1}{\ell_s^2}\int_{\Sigma_2^I} B_2, \tag{2.126}$$

with $B_2 = \sum_I b_I(x)\omega_2^I$. Upon dimensional reduction, the first term of (2.125) yields kinetic terms for the b_I axions,

$$S_{\text{kin}} = -\frac{1}{2}\int \mathrm{d}^4x\,\sqrt{-g}\,\gamma^{IJ}\partial_\mu b_I \partial^\mu b_J \qquad\qquad \frac{\gamma^{IJ}}{M_{\text{pl}}^2} = \frac{g_s}{2\mathcal{V}_{\text{E}}\ell_s^6}\int_{X_6} *_6\,\omega_2^I \wedge \omega_2^J. \tag{2.127}$$

If a basis of harmonic forms ω_2^I is chosen such that γ^{IJ} is diagonal, then $\phi_I = f_I b_I$ (no sum) are the canonically normalized axion fields, whose decay constants are the eigenvalues of γ, $f_I = \mathrm{eig}_I\,\gamma^{JK}$. For example, if the compact space is a product

[21] Normalization conventions appear in Appendix 2.C.

of two-spheres, $X_6 = \mathrm{S}^2 \times \mathrm{S}^2 \times \mathrm{S}^2$, each with volume $L^2\ell_s^2$, then we simply have $f_I^2 = M_{\mathrm{pl}}^2 L^{-4}/2$.

If X_6 is Calabi-Yau, then the axion decay constants for two-form axions b_I and c_I—arising from the two-form potentials B_2 and C_2, respectively, in type IIB string theory—may be simply computed from the intersection numbers κ_{IJK}, the volumes $\ell_s^2 v^I$ of the two-cycles Σ_2^I, and the overall total volume $\mathcal{V}_{\mathrm{E}}$ of X_6. For example, for an axion $c_- = \ell_s^{-2}\int_{\Sigma_2^-} C_2$ associated to an orientifold-odd cycle Σ_2^- the axion decay constant is

$$\frac{f^2}{M_{\mathrm{pl}}^2} \sim g_s \frac{\kappa_{I--} v^I}{\mathcal{V}_{\mathrm{E}}}. \tag{2.128}$$

When fivebranes are introduced to create monodromy, the axion that experiences this monodromy will in general be a linear combination of the ω_2^I, which we call Ω. For example, in a variant of the axion monodromy construction detailed in §2.3, the (rel)axion $c(x)$ arises from a two-form Ω dual to the blowup cycle of an orbifold whose fixed point locus is Σ_{o}, with $\dim_{\mathbb{C}} \Sigma_{\mathrm{o}} = 1$. As shown in [34], the support of $*\Omega \wedge \Omega$ is localized about Σ_{o}. The six-dimensional metric is approximately a cone,

$$\tilde{g}_{mn}\mathrm{d}y^m\,\mathrm{d}y^n \approx \mathrm{d}r^2 + r^2 \breve{g}_{\theta\phi}\,\mathrm{d}\Psi^\theta\,\mathrm{d}\Psi^\phi \tag{2.129}$$

and $\Omega_{mn} \sim \Omega_{\theta\phi}$ to good approximation has its legs along the angular directions, so

$$\frac{f^2}{M_{\mathrm{pl}}^2} \sim \frac{g_s}{\mathcal{V}_{\mathrm{E}}}\frac{1}{\ell_s^6}\int_{r_\cup}^{r_\cap}\mathrm{d}r\, r \int \mathrm{d}^5\Psi \sqrt{\breve{g}}\,\breve{g}^{\theta\phi}\breve{g}^{\psi\chi}\Omega_{\theta\psi}\Omega_{\phi\chi} \approx \frac{g_s}{\mathcal{V}_{\mathrm{E}}}\frac{r_\cap^2}{\ell_s^2}. \tag{2.130}$$

Locally, we may think of the blow-up cycle as an Eguchi-Hanson space fibered over Σ_{o}. Since the integrand is highly localized about Σ_{o}, we have $\int *\Omega \wedge \Omega \approx \mathrm{vol}(\Sigma_{\mathrm{o}})$, and because of the conical nature of the six-dimensional metric, $\mathrm{vol}(\Sigma_{\mathrm{o}})$ is dominated by the contribution at $r_\cap$, so $\int *_6 \Omega \wedge \Omega \propto r_\cap^2$.

The axion enjoys a continuous shift symmetry to all orders in perturbation theory in both g_s and α'. However, this continuous shift symmetry does not survive at the nonperturbative level, and is broken to a discrete shift symmetry by instantons carrying axion charge. In particular, fundamental strings are charged under B_2, via the coupling

$$S(\mathcal{W}) = \ldots + \frac{i}{2\pi\alpha'} \int_{\mathcal{W}} \mathrm{d}^2\sigma \, \sqrt{-h} \, \epsilon^{mn} B_{mn} + \ldots, \tag{2.131}$$

where h is the metric on the string worldsheet $\mathcal{W}$, and m, n are two-dimensional indices tangent to $\mathcal{W}$. The string path integral receives a contribution from a Euclidean string whose worldsheet wraps Σ_2^I, termed a worldsheet instanton. This contribution will be proportional to e^{-S_I}, where $S_I = S(\Sigma_2^I)$. Because of the coupling (2.131),

$$S_I \supset 2\pi i b_I, \tag{2.132}$$

and so the potential generated by these nonperturbative effects is still invariant under discrete shifts $b_I \mapsto b_I + N$, $N \in \mathbb{Z}$, as $e^{-S_I} \mapsto e^{-S_I + 2\pi i N} = e^{-S_I}$. Thus, worldsheet instantons break the perturbative, continuous shift symmetry of b_I to the discrete shift $b_I \mapsto b_I + 1$.

The real part of the action (2.131) is proportional to the volume of Σ_2^I in string units, and worldsheet instanton contributions become more important as the volume shrinks. These contributions are difficult to compute, so requiring computational control of the effective action constrains the sizes of cycles, Σ_2^I, and thus the sizes of the axion decay constants. A standard requirement for control is that the sizes of all cycles are much larger than the string length, $v^\alpha \gg 1$.

However, the two-cycle Σ_2^I may sit in a warped region, with warp factor e^A. For two-form axions, (2.127) is unchanged—there is no explicit dependence on the warp factor. However, a ten-dimensional string will see a warped volume, and in

particular the real part of the worldsheet instanton action is enhanced by factor e^{-2A}. This allows the two-cycle volumes v^α to be smaller by a factor of e^{-2A} without loss of control, and so the axion decay constant can be very small in a highly warped throat.

If we take v^I to measure the warped volume of Σ_2^I in string units, i.e. the volume a ten-dimensional string would measure, then we may write

$$\frac{f^2}{M_{\rm pl}^2} \sim g_s \frac{\kappa_{I--} v^I}{\mathcal{V}_{\rm E}} \, e^{2A}\big|_{\Sigma_2^I} \ll g_s \frac{\kappa_{I--} v^I}{\mathcal{V}_{\rm E}}, \tag{2.133}$$

keeping the constraint that $v^I \gg 1$.

2.B Necessity of the Intersection

In NS5-brane axion monodromy, D3-brane charge accumulates on an NS5-brane that wraps a two-cycle $\Sigma_{\rm NS5}$ (denoted Σ_2 elsewhere in the text). Taking $c(x) \equiv \int_{\Sigma_{\rm NS5}} C_2$ to be the relaxion field, a stopping potential can be generated by strong gauge dynamics in a gauge theory G to which the relaxion has a nonvanishing axionic coupling $\lambda\, c(x) F \wedge F$, with λ a constant. We will take G to be realized on a stack of D7-branes wrapping a holomorphic four-cycle D (denoted Σ_4 elsewhere in the text). The backreaction of the D3-brane charge changes the supergravity background, with the strongest effects occurring near $\Sigma_{\rm NS5}$. In this appendix we show that any D for which $\lambda \neq 0$ *necessarily intersects* $\Sigma_{\rm NS5}$. Thus, one cannot mitigate the backreaction by arranging that D is outside of the warped region.

For our purposes, it suffices to show that D and $\Sigma_{\rm NS5}$ have at least one point in common, even though the intersection number $[D] \cap [\Sigma_{\rm NS5}]$ of the corresponding homology classes may vanish. We will use $\cap_s$ to denote intersection as point sets,

as distinct from the topological intersection $[\Sigma_1] \cap [\Sigma_2]$,[22] and we will show that $\lambda \neq 0$ implies that $D \cap_s \Sigma_{\rm NS5}$.

Consider a D7-brane that fills spacetime and wraps a smooth four-cycle $D \subset X$ in the internal space X. The D7-brane couples to C_2 axions via the Chern-Simons action

$$S_{\rm CS} = \mu_7 \int_{\mathcal{W}} \sum_p \iota^* C_p \wedge e^{\mathcal{F}} \supset \frac{\mu_7}{3!} \int_{\mathcal{W}} \iota^* C_2 \wedge \mathcal{F} \wedge \mathcal{F} \wedge \mathcal{F}, \tag{2.134}$$

where $\mathcal{W} = \mathcal{M}^{3,1} \times D$ is the D7-brane worldvolume, $\iota : D \to X$ is the inclusion map of D into X, ι^* denotes the pullback onto D, F_2 is the field strength of the worldvolume gauge theory, and $\mathcal{F} = \iota^* B_2 + 2\pi\alpha' F_2$. The axionic coupling to the gauge theory G on a stack of D7-branes wrapping D is therefore

$$S_{\rm CS} \supset \frac{\mu_7}{2!} \left(\int_D \iota^* C_2 \wedge \mathcal{F} \right) \left(\int_{\mathcal{M}^{3,1}} \mathrm{tr}\, \mathcal{F} \wedge \mathcal{F} \right) . \tag{2.135}$$

Poincaré duality in D relates the flux $\mathcal{F}$ to a two-cycle $S_{\mathcal{F}} \subset D$ which may further be viewed as a two-cycle $\iota_* S_{\mathcal{F}}$ in X, so that

$$\vartheta \equiv \int_D \iota^* C_2 \wedge \mathcal{F} = \int_{S_{\mathcal{F}}} \iota^* C_2 = \int_{\iota_* S_{\mathcal{F}}} C_2. \tag{2.136}$$

The holomorphic representative $\iota_* S_{\mathcal{F}}$ of the class $[\iota_* S_{\mathcal{F}}]$ is contained, as a point set, in D. So establishing the condition

$$\iota_* S_{\mathcal{F}} \cap_s \Sigma_{\rm NS5} \neq \emptyset \tag{2.137}$$

will imply our desired result $D \cap_s \Sigma_{\rm NS5}$.

Now we choose a basis of nontrivial two-cycle classes, $\{[\Sigma_i]\}$, $I = 1, \ldots, p \equiv h^{1,1}$, to span $H_2(X, \mathbb{Z})$. Without loss of generality, we may take the NS5-brane class

[22] Two submanifolds M, N, of X have $M \cap_s N \neq \emptyset$ if and only if M and N have at least one point in common, without regard to the orientation of M and N.

$[\Sigma_{\rm NS5}]$ to be an element of this basis, say $[\Sigma_1] \equiv [\Sigma_{\rm NS5}]$. There exists a dual basis of harmonic two-forms ω^J such that $\int_{\Sigma_I} \omega^J = \delta_I^J$. Expanding $\ell_s^{-2} C_2(x) = \sum_{i=1}^p c_I(x)\omega^I$, the relaxion field is $c_1(x) \equiv c(x)$. We may also expand

$$\iota_* S_{\mathcal{F}} = a_1[\Sigma_1] + \cdots + a_p[\Sigma_p] + (\text{boundary})\,, \tag{2.138}$$

for some integers a_I. Comparing to (2.136), we see that $\lambda \neq 0$ if and only if $a_1 \neq 0$.

The relation (2.138) with $a_1 \neq 0$ does not, on its own, imply (2.137). For example, consider a basis of homology $\{[\Sigma_1], [\Sigma_2]\}$, with minimum volume representatives $\{\Sigma_1, \Sigma_2\}$ that obey $\Sigma_I \cap_s \Sigma_J = \delta_{IJ}$. If $[S] = a_1[\Sigma_1] + a_2[\Sigma_2]$, then $S \cap_s \Sigma_1 \neq \emptyset \iff a_1 \neq 0$, regardless of the value of a_2. But working in the basis $\{[\Sigma_1], [\Sigma_2'] = [\Sigma_2] - [\Sigma_1]\}$, for $a_1 \neq 0$ and $a_2 = 0$ we again have $S \cap_s \Sigma_1 \neq \emptyset$, while if $a_1 = a_2$ we have instead $S \cap_s \Sigma_1 = \emptyset$. Thus, the condition (2.137) depends on the relation between $[\Sigma_1]$ and $[\Sigma_2], \ldots [\Sigma_p]$, which we have not yet specified.

We may view this issue in a dual picture. The coupling (2.136) can be written as the triple intersection of three divisors in X,

$$\vartheta = [D] \cap [D_{\mathcal{F}}] \cap [\mathrm{D}_{\rm NS5}]\,, \tag{2.139}$$

where $D_{\mathcal{F}} = \mathrm{PD}_X(\iota_* \mathcal{F})$ and $D_{\rm NS5} = \mathrm{PD}_X(\omega^1)$, with PD_X denoting the Poincaré dual in X. The divisor $D_{\rm NS5}$ is dual to the curve $\Sigma_{\rm NS5}$, in that $D_{\rm NS5}$ is Poincaré dual to the two-form ω^1 that is the dual vector to $\Sigma_{\rm NS5}$ with respect to the pairing $\int_{\Sigma_I} \omega^J = \delta_I^J$. It follows that $[\Sigma_{\rm NS5}] \cap [D_{\rm NS5}] = 1$. Moreover, the requirement of a nonvanishing axionic coupling, $\vartheta \neq 0$, implies that $[D_{NS5}] \cap [D] \neq 0$. Now although $[\Sigma_{\rm NS5}] \cap [D_{\rm NS5}] \neq 0$ and $[D_{\rm NS5}] \cap [D] \neq 0$, it appears that $D_{\rm NS5}$ could stretch between D and $\Sigma_{\rm NS5}$, intersecting each, even though D and $\Sigma_{\rm NS5}$ remain widely separated.

To exclude this possibility, we use further facts about $\Sigma_{\rm NS5}$ and $\iota_* S_{\mathcal{F}}$. Preserv-

ing supersymmetry in the D7-brane worldvolume requires that $\mathcal{F} \in H^2(D, \mathbb{Z})$ be of type $(1,1)$, and so its Poincaré dual $\iota_* S_{\mathcal{F}}$ is a holomorphic curve. Heuristically, $\iota_* S_{\mathcal{F}}$ can be viewed as the curve wrapped by a D5-brane dissolved in D: if the D7-brane were annihilated by introducing an anti-D7-brane, a D5-brane on $\iota_* S_{\mathcal{F}}$ would remain. Moreover, $\Sigma_{\rm NS5}$ is itself an irreducible holomorphic curve. (In a construction in which $\Sigma_{\rm NS5}$ is a sum of irreducible holomorphic curves, this argument can be applied to each component.) We can therefore express $\iota_* S_{\mathcal{F}}$ uniquely as a finite sum of distinct irreducible holomorphic curves $\{\sigma_A\}$, $A = 1, \ldots K$ (with $\Sigma_{\rm NS5} \equiv \sigma_1$):

$$\iota_* S_{\mathcal{F}} = a_1 \Sigma_{\rm NS5} + a_2 \sigma_2 + \cdots + a_K \sigma_K, \qquad a_A \in \mathbb{Z} \geq 0\,, \tag{2.140}$$

and we have shown above that a relaxionic coupling, $\lambda \neq 0$, requires $a_1 \neq 0$. Because the σ_A are distinct irreducible holomorphic curves, they intersect each other at most at points. So $\iota_* S_{\mathcal{F}}$ contains all but finitely many of the points of $\Sigma_{\rm NS5}$. The condition (2.137) then follows, and so D must intersect $\Sigma_{\rm NS5}$, which is what we set out to prove.

Note that if the σ_A were simply a set of distinct, irreducible simplicial complexes, the relation (2.137) would not be automatic. If σ_1 intersected some of $\sigma_2, \ldots \sigma_K$ along suitable two-simplices, then $\sum_i a_A \sigma_A$ might have no points in common with σ_1, because adding $a_2\sigma_2 + \ldots a_K \sigma_K$ could subtract all the points of σ_1. For curves intersecting at most at points, this is not possible.

2.C Type IIB Supergravity with Fivebranes

2.C.1 Conventions for type IIB supergravity

The bosonic part of the type IIB supergravity action in Einstein frame is

$$S_{\rm IIB} = \frac{1}{2\kappa_{10}^2}\int \mathrm{d}^{10}X\,\sqrt{-G_{\rm E}}\left(R_{\rm E} - \frac{|\partial\tau|^2}{2(\mathrm{Im}\,\tau)^2} - \frac{|G_3|^2}{2\,\mathrm{Im}\,\tau} - \frac{1}{4}|\tilde{F}_5|^2\right) - \frac{i}{8\kappa_{10}^2}\int \frac{C_4\wedge G_3\wedge \bar{G}_3}{\mathrm{Im}\,\tau} \tag{2.141}$$

with $2\kappa_{10}^2 = \ell_s^8/2\pi$, $G_3 \equiv F_3 - \tau H_3$, $\tau \equiv C_0 + ie^{-\Phi}$, $F_{p+1} = \mathrm{d}C_p$, $H_3 = \mathrm{d}B_2$, $\tilde{F}_5 = F_5 - \frac{1}{2}C_2\wedge H_3 + \frac{1}{2}B_2\wedge F_3$, and $\tilde{F}_5 = *_{10}\tilde{F}_5$ is imposed at the level of the equations of motion.

We define the string length

$$\ell_s^2 = (2\pi)^2\alpha'. \tag{2.142}$$

The actions for extremal Dp-branes and NS5-branes are given by

$$S_{\mathrm{D}p} = -\mu_p\int \mathrm{d}^{p+1}\xi\, e^{-\Phi}\sqrt{-\det\left(G_{ab} + B_{ab} + 2\pi\alpha' F_{ab}\right)} + S_{\rm CS} \tag{2.143}$$

and

$$S_{\rm NS5} = -\mu_5\int \mathrm{d}^6\xi\, e^{-2\Phi}\sqrt{-\det\left(G_{ab} - e^{\Phi}\left(C_{ab} + 2\pi\alpha' F_{ab}\right)\right)} + S_{\rm CS}, \tag{2.144}$$

respectively, where $\mu_p = 2\pi/\ell_s^{p+1}$, and F_{ab} is the gauge field strength on the brane worldvolume. The Chern-Simons term for a Dp-brane reads

$$S_{\rm CS} = \mu_p\int \sum_n C_n\wedge e^{\mathcal{F}}, \tag{2.145}$$

where we have introduced the notation $\mathcal{F} = B_2 + 2\pi\alpha' F$. The Chern-Simons piece sets the flux quantization condition

$$\frac{1}{\ell_s^{p+1}}\int_{\Sigma^p} F_{p+1} \in \mathbb{Z}. \tag{2.146}$$

We expand in a basis of $H^2(X_6)$, denoted $\omega^I(y)$, with normalization

$$\int_{\Sigma^J} \omega^I = \ell_s^2 \delta^I_J. \tag{2.147}$$

We list our index conventions in Table 2.3.

Directions	**Indices**
(3+1)-dim spacetime	$\mu, \nu, \rho, \ldots$
6-dim internal space	$m, n, \ldots$
5-dim angular space	$\theta, \phi, \ldots$
brane worldvolume	$a, b, \ldots$
transverse to worldvolume	$i, j, k, \ldots$
transverse vielbein	$\hat{\imath}, \hat{\jmath}, \hat{k}, \ldots$
along cycle Σ_2	$\alpha, \beta, \ldots$

Table 2.3: A guide to this chapter's index conventions.

2.C.2 Einstein-frame potentials for fivebranes

The DBI action for a D5-brane is

$$S_{\text{DBI}} = -\frac{2\pi}{\ell_s^6} \int \mathrm{d}^6\xi \, e^{-\Phi} \sqrt{-\det\left(G^S_{ab} + \mathcal{F}_{ab}\right)} \tag{2.148}$$

where G^S_{ab} and B_{ab} are the pull-backs onto the brane worldvolume of the ten-dimensional string-frame metric and NS-NS two-form B_2, respectively.

The ten-dimensional string-frame metric is related to the Einstein-frame metric via

$$G^S_{MN} = e^{\Phi/2} G^E_{MN}, \tag{2.149}$$

which we assume takes a warped product form

$$G^E_{MN}\, \mathrm{d}X^M\, \mathrm{d}X^N = e^{2A(y)} g_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu + e^{-2A(y)} \tilde{g}_{mn}\, \mathrm{d}y^m\, \mathrm{d}y^n, \tag{2.150}$$

where $g_{\mu\nu}$ and $\tilde{g}_{mn}$ are metrics on the four-dimensional spacetime and the six-dimensional internal space, respectively.

Defining

$$b(x) = \frac{1}{\ell_s^2}\int_{\Sigma_2} B_2 = \frac{1}{\ell_s^2}\int_{\Sigma_2} \ell_s^2\, b(x)\, \mathrm{d}y \wedge \mathrm{d}z, \tag{2.151}$$

choosing coordinates on Σ_2 such that $\mathrm{d}y \wedge \mathrm{d}z$ is harmonic, and setting $F_{ab} = 0$, we may write

$$e^{\Phi/2} G^E_{ab} + B_{ab} = \begin{pmatrix} e^{\Phi/2} e^{2A(y)} g_{\mu\nu} & 0 \\ 0 & \mathbf{m} \end{pmatrix} \tag{2.152}$$

with

$$\mathbf{m} = \begin{pmatrix} e^{\Phi/2} e^{-2A} \tilde{g}_{yy} & e^{\Phi/2} e^{-2A} \tilde{g}_{yz} + \ell_s^2 b/2 \\ e^{\Phi/2} e^{-2A} \tilde{g}_{yz} - \ell_s^2 b/2 & e^{\Phi/2} e^{-2A} \tilde{g}_{zz} \end{pmatrix}. \tag{2.153}$$

With $\tilde{g}_2 \equiv \tilde{g}_{yy}\tilde{g}_{zz} - \tilde{g}_{yz}^2$, we have

$$\sqrt{-\det\left(G_{ab} + B_{ab}\right)} = e^{\Phi + 4A} \sqrt{-\det g} \sqrt{e^{\Phi} e^{-4A} \tilde{g}_2 + \ell_s^4 b^2/4}. \tag{2.154}$$

Upon integration over Σ_2, we may take $e^{-4A}\tilde{g}_2 \to \ell_{\mathrm{E}}^4$, where ℓ_{E}^2 is the characteristic size of Σ_2 in the ten-dimensional *Einstein frame*. The DBI action, upon dimensional reduction, then yields a four-dimensional potential for the b axion,

$$S_{\mathrm{DBI}} = \int \mathrm{d}^4x\, \sqrt{-g} \left(-\frac{\pi e^{4A}}{\ell_s^4} \sqrt{\frac{g_s}{4}\left(\frac{\ell_{\mathrm{E}}}{\ell_s}\right)^4 + b^2} \right). \tag{2.155}$$

The dimensional reduction of the NS5-brane action follows similarly.

We will also be interested in the spectrum of Kaluza-Klein excitations, which we now compute. Setting $F_{ab} = 0$, the NS5-brane action is

$$S_{\mathrm{NS5}} = -\frac{2\pi}{\ell_s^6} \int \mathrm{d}^6\xi\, e^{-2\Phi} \sqrt{-\det \mathbf{M}} \tag{2.156}$$

with $\mathbf{M}_{ab} = G_{ab} - e^{\Phi} C_{ab}$. Expanding in fluctuations $\delta\mathbf{M}_{ab}$, we have

$$\sqrt{-\det\left(\overline{\mathbf{M}} + \delta\mathbf{M}\right)} = \sqrt{-\det \overline{\mathbf{M}}} \left(1 + \frac{1}{2} \operatorname{tr} \overline{\mathbf{M}}^{-1} \delta\mathbf{M}\right), \tag{2.157}$$

with

$$\overline{\mathbf{M}} = \begin{pmatrix} e^{\Phi/2}e^{2A}g_{\mu\nu} & 0 \\ 0 & \mathbf{m} \end{pmatrix} \tag{2.158}$$

and

$$\mathbf{m} = \begin{pmatrix} e^{\Phi/2}e^{-2A}\tilde{g}_{yy} & e^{\Phi/2}e^{-2A}\tilde{g}_{yz} - e^{\Phi}\ell_s^2 c/2 \\ e^{\Phi/2}e^{-2A}\tilde{g}_{yz} + e^{\Phi}\ell_s^2 c/2 & e^{\Phi/2}e^{-2A}\tilde{g}_{zz} \end{pmatrix}. \tag{2.159}$$

As above,

$$\sqrt{-\det\overline{\mathbf{M}}} = e^{\Phi+4A}\sqrt{-g}\sqrt{e^{\Phi}e^{-4A}\tilde{g}_2 + e^{2\Phi}\ell_s^4 c^2/4}. \tag{2.160}$$

The fluctuation $\delta\mathbf{M}$ arises from allowing the embedding of the NS5-brane to fluctuate. We may explicitly write the pull-back as

$$G_{ab} - e^{\Phi}C_{ab} = \Pi_{ab}^{MN}\left(G_{MN} - e^{\Phi}C_{MN}\right). \tag{2.161}$$

If we take the embedding of the fivebrane to be specified by $X^M(\xi^a)$ and allow the brane to fluctuate in the transverse directions $X^M(\xi^a) = \delta_a^M\xi^a + \delta_j^M X^j(\xi^b)$, the projection operator is then

$$\Pi_{ab}^{MN} \equiv \frac{\partial X^M}{\partial\xi^a}\frac{\partial X^N}{\partial\xi^b} = \delta_a^M\delta_b^N + \delta_a^M\delta_j^N\frac{\partial X^j}{\partial\xi^b} + \delta_b^N\delta_i^M\frac{\partial X^i}{\partial\xi^b} + \delta_i^M\delta_j^N\frac{\partial X^i}{\partial\xi^a}\frac{\partial X^j}{\partial\xi^b}. \tag{2.162}$$

Assuming a product metric $G_{ai}^E = 0$, we have

$$\delta\mathbf{M}_{ab} = e^{\Phi/2}e^{-2A}\tilde{g}_{ij}\frac{\partial X^i}{\partial\xi^a}\frac{\partial X^j}{\partial\xi^b}. \tag{2.163}$$

and the NS5-brane action may be written

$$S_{\rm NS5} = -\frac{2\pi}{\ell_s^6}\int \mathrm{d}^6\xi\, e^{-\Phi}e^{4A}\sqrt{-g}\sqrt{e^{\Phi}e^{-4A}\tilde{g}_2 + e^{2\Phi}\ell_s^4 c^2/4}\times \tag{2.164}$$
$$\left(1 + \frac{1}{2}e^{-4A}g^{\mu\nu}\tilde{g}_{ij}\partial_\mu X^i\partial_\nu X^j + \frac{1}{2}\frac{e^{\Phi}e^{-4A}\tilde{g}_2}{e^{\Phi}e^{-4A}\tilde{g}_2 + e^{2\Phi}\ell_s^4 c^2/4}\tilde{g}_{ij}\tilde{g}^{\alpha\beta}\tilde{\nabla}_\alpha X^i\tilde{\nabla}_\beta X^j\right).$$

We define the canonically normalized fields as

$$Y^{\hat{\imath}} = FE^{\hat{\imath}}{}_j X^j \tag{2.165}$$

with

$$F^2(x^\mu, y, z) = \frac{2\pi}{\ell_s^6}\frac{e^{-\Phi}}{\sqrt{\tilde{g}_2}}\sqrt{e^{\Phi}e^{-4A}\tilde{g}_2 + e^{2\Phi}\ell_s^4 c^2/4} \qquad \text{and} \qquad \tilde{g}_{ij} = \delta_{\hat{i}\hat{j}}E^{\hat{i}}{}_i E^{\hat{j}}{}_j. \tag{2.166}$$

We have assumed that $\tilde{g}_{i\alpha} = 0$, and thus $\tilde{\nabla}_\alpha E^{\hat{j}}{}_j = 0$. Decomposing in real $\tilde{g}_2$ harmonics gives

$$\tilde{\nabla}^2\mathcal{Y}^I = -\frac{e^{-2A}}{\ell_{\rm E}^2}\lambda^I\mathcal{Y}^I \qquad Y^{\hat{\imath}} = \sum_I Y_I^{\hat{\imath}}\mathcal{Y}^I \qquad \int \mathrm{d}^2\sigma\sqrt{\tilde{g}_2}\,\mathcal{Y}^I\,\mathcal{Y}^J = \delta^{IJ}. \tag{2.167}$$

The action is

$$\begin{aligned} S_{\rm NS5} = \int \mathrm{d}^4x\,\sqrt{-g}\Bigg(& -V(c) - \frac{1}{2}g^{\mu\nu}\partial_\mu Y_I^{\hat{\imath}}\partial^\mu Y_{\hat{\imath}}^I - \frac{1}{2}m_I^2(c)Y_I^{\hat{\imath}}Y_{\hat{\imath}}^I \\ & + g(c)\,(c\partial_\mu c)\left(Y_I^{\hat{\imath}}\partial^\mu Y_{\hat{\imath}}^I\right) - \frac{1}{2}g(c)^2\,(\partial c)^2\,Y_I^{\hat{\imath}}Y_{\hat{\imath}}^I\Bigg) \end{aligned} \tag{2.168}$$

with

$$V(c) \equiv \frac{\pi}{\ell_s^4}\frac{e^{4A}}{\sqrt{g_s}}\sqrt{4\left(\frac{\ell_{\rm E}}{\ell_s}\right)^4 + g_s c^2} \tag{2.169a}$$

$$g(c) \equiv \frac{g_s}{2}\left(4\left(\frac{\ell_{\rm E}}{\ell_s}\right)^4 + g_s c^2\right)^{-1} \tag{2.169b}$$

$$m_I^2(c) \equiv \frac{\mu_5^2 e^{-\Phi}}{F^4}\frac{e^{-2A}}{\ell_{\rm E}^2}\lambda_I = 4\lambda_I\frac{e^{2A}}{\ell_{\rm E}^2}\left(\frac{\ell_{\rm E}}{\ell_s}\right)^4\left(4\left(\frac{\ell_{\rm E}}{\ell_s}\right)^4 + g_s c^2\right)^{-1}. \tag{2.169c}$$

Finally, we will be interested in the tension of the domain wall interpolating between the metastable and stable states of the NS5-brane axion monodromy scenario. In the thin-wall approximation, the domain wall corresponds to an NS5-brane winding n times around the minimum volume three-cycle Σ_3 whose endpoints are Σ_2 and $\bar{\Sigma}_2$, the two-cycles wrapped by the NS5-brane and anti-NS5-brane, respectively. The tension can then be read off by reducing the action

$$\begin{aligned} S_{\rm NS5} &= -\frac{2\pi}{\ell_s^6}\int_{\mathcal{D}} \mathrm{d}^6\xi\, e^{-2\Phi}\sqrt{-\det\left(G_{ab}^S - e^{\Phi}C_{ab}\right)} \\ &\to -T_D\int_{\mathcal{D}^{2,1}} \mathrm{d}^3x\,\sqrt{-\det\mathcal{P}(g_4)}. \end{aligned} \tag{2.170}$$

where $\mathcal{P}(g_4)$ denotes the pullback of the spacetime metric $g_{\mu\nu}$ onto the world-volume of the domain wall $\mathcal{D}^{2,1}$.

We can gain intuition for this tension by modeling the three-cycle Σ_3 as

$$\mathrm{d}s^2_{\Sigma_3} = \mathrm{d}r^2 + r^2\left(\mathrm{d}y^2 + \mathrm{d}z^2\right), \tag{2.171}$$

where the two-torus volume form is $\mathrm{d}y \wedge \mathrm{d}z$, and $r \in [r_\cup, r_\cap]$. Then

$$G^S_{ab} - e^{\Phi} C_{ab} = \begin{pmatrix} e^{\Phi/2} e^{2A} \mathcal{P}(g_4)_{\mu\nu} & 0 & 0 \\ 0 & e^{-2A} e^{\Phi/2} & 0 \\ 0 & 0 & \mathbf{m} \end{pmatrix} \tag{2.172}$$

where

$$\mathbf{m} = \begin{pmatrix} e^{-2A} e^{\Phi/2} r^2 & -e^{\Phi} \ell_s^2 c/2 \\ -e^{\Phi} \ell_s^2 c/2 & e^{-2A} e^{\Phi/2} r^2 \end{pmatrix}, \tag{2.173}$$

and so

$$S_{\mathrm{NS5}} = -\left(\frac{2\pi}{\ell_s^6} \int_{r_\cup}^{r_\cap} \mathrm{d}r\, e^{-3\Phi/4} e^{3A} \sqrt{e^{-4A} r^4 + e^{\Phi} \ell_s^4 c^2/4}\right) \int_{\mathcal{D}^{2,1}} \mathrm{d}^3 x \, \sqrt{-\det \mathcal{P}(g_4)}. \tag{2.174}$$

The tension then takes the form

$$T_{\mathrm{D}} = \frac{2\pi}{\ell_s^3} \frac{L^3}{g_s^{3/4} \ell_s^3} \frac{1}{4} \left(e^{4A_\cap} - e^{4A_\cup}\right) \sqrt{1 + \frac{g_s \ell_s^4 c^2}{4L^4}} \sim \frac{1}{\ell_s^3} N_{\mathrm{D3}}^{3/4} \left(e^{4A_\cap} - e^{4A_\cup}\right), \tag{2.175}$$

since we must have $g_s \ell_s^4 c^2/4L^4 \sim N^2/N_{\mathrm{D3}} \ll 1$ to avoid the KPV instability.

2.D Backreaction on the Internal Space

When one introduces a source of monodromy in a compactification, and explicitly breaks supersymmetry, the corresponding stress-energy will backreact on the metric, affecting the parameters in the low-energy effective theory. In the NS5-brane model detailed in §2.3.2, a key source of stress-energy is anti-D3-brane charge

induced on the anti-NS5-brane. Because D3-brane charge preserves the same supersymmetry as the background, it will not backreact on the internal metric at leading order.[23] However, the anti-D3-brane charge will break the remaining supersymmetry of the background and perturb the internal metric. Furthermore, both D3-brane and anti-D3-brane charge will perturb the warp factor e^{4A}. In this appendix we calculate the perturbations to the internal metric and the warp factor.

At the level of the supergravity equations of motion, we may approximate the N units of induced anti-D3-brane charge as N anti-D3-branes smeared about the anti-NS5-brane. These anti-D3-branes do not source the metric directly, but do so through a combination of the warp factor e^{4A} and the $\tilde{F}_5$ field which we denote Φ_-, where $\Phi_\pm \equiv e^{4A} \pm \alpha$, and

$$F_5 = (1 + *_{10})\, \mathrm{d}\alpha(y) \wedge \mathrm{dvol}_{\mathbb{R}^{1,3}}. \tag{2.176}$$

The equations of motion for $\Phi_\pm$ and the internal Einstein equation read

$$\tilde{\nabla}^2 \Phi_+ = \frac{2}{\Phi_+ + \Phi_-}(\tilde{\nabla}\Phi_+)^2 + \frac{1}{2} g_s \ell_s^4 \,(\Phi_+ + \Phi_-)^2 \sum_i \delta(\mathrm{D3}_i) \tag{2.177a}$$

$$\tilde{\nabla}^2 \Phi_- = \frac{2}{\Phi_+ + \Phi_-}(\tilde{\nabla}\Phi_-)^2 + \frac{1}{2} g_s \ell_s^4 \,(\Phi_+ + \Phi_-)^2 \sum_i \delta(\overline{\mathrm{D3}}_i) \tag{2.177b}$$

$$\tilde{R}_{mn} = \frac{2}{(\Phi_+ + \Phi_-)^2} \tilde{\nabla}_{(m} \Phi_+ \tilde{\nabla}_{n)} \Phi_- \tag{2.177c}$$

where we use $\tilde{\nabla}$, etc., to denote quantities related to the unwarped, internal metric $\tilde{g}_{mn}$. We treat the anti-D3-branes as a perturbation to an imaginary self-

[23] In the presence of anti-brane charge, the D3-brane charge will backreact on the internal metric at second order. Similarly, if the D3-brane charge is large enough a better description becomes available in which the D3-branes are dissolved into flux and a new warped throat is formed, corresponding to the analysis of §2.4.2. In what follows, we will take $N \ll N_{\mathrm{D3}}$ and assume that the induced anti-D3-brane charge may be thought of as a small perturbation to the geometry.

dual background, in which

$$\Phi_+ \approx \frac{2r^4}{L^4} \qquad \text{and} \qquad \Phi_- = 0, \tag{2.178}$$

$L^4 \propto g_s N_{\mathrm{D3}} \ell_s^4$, and the internal metric is taken to be the conifold, a cone over $\mathrm{T}^{1,1}$,

$$\tilde{g}_{mn}\mathrm{d}y^m\,\mathrm{d}y^n = \mathrm{d}r^2 + r^2\mathrm{d}s^2_{\mathrm{T}^{1,1}} = \mathrm{d}r^2 + r^2\breve{g}_{ij}\mathrm{d}\Psi^i\,\mathrm{d}\Psi^j. \tag{2.179}$$

We linearize the system of equations (2.177) using the expansions[24]

$$\tilde{g}_{mn} = \tilde{g}^{(0)}_{mn} + \delta\tilde{g}_{mn} \tag{2.180a}$$

$$\delta\tilde{g}_{rr} = \sum_{I_s} \tau^{I_s}(r)\mathcal{Y}^{I_s}(\Psi) \tag{2.180b}$$

$$\delta\tilde{g}_{r\theta} = \sum_{I_v} b^{I_v}(r)\mathcal{Y}^{I_v}_{\theta}(\Psi) \tag{2.180c}$$

$$\delta\tilde{g}_{\theta\phi} = \sum_{I_s} \frac{1}{5}\pi^{I_s}(r)\breve{g}_{\theta\phi}\mathcal{Y}^{I_s}(\Psi) + \sum_{I_t} \phi^{I_t}(r)\mathcal{Y}^{I_t}_{\theta\phi}(\Psi) \tag{2.180d}$$

$$\Phi_- = \Phi^{(0)}_- + \delta\Phi^{(1)}_-(r,\Psi) = \sum_{I_s} \delta\Phi^{I_s}_-(r)\mathcal{Y}^{I_s}(\Psi) \tag{2.180e}$$

where $\mathcal{Y}^{I_s}(\Psi)$, $\mathcal{Y}^{I_v}_{\theta}(\Psi)$ and $\mathcal{Y}^{I_t}_{\theta\phi}(\Psi)$ are the scalar, transverse vector, and transverse traceless tensor harmonics on $\mathrm{T}^{1,1}$, with appropriate Laplacian eigenvalues $\lambda(I_s)$, $\lambda(I_v)$, and $\lambda(I_t)$, respectively.

The Einstein metric on $\mathrm{T}^{1,1}$ is

$$\begin{aligned}\mathrm{d}s^2_{\mathrm{T}^{1,1}} = {}& \frac{1}{6}\left(\mathrm{d}\theta_1^2 + \sin^2\theta_1\,\mathrm{d}\varphi_1^2\right) + \frac{1}{6}\left(\mathrm{d}\theta_2^2 + \sin^2\theta_2\,\mathrm{d}\varphi_2^2\right) \\ & + \frac{1}{9}\left(\cos\theta_1\,\mathrm{d}\varphi_1 + \cos\theta_2\,\mathrm{d}\varphi_2 + \mathrm{d}\psi\right)^2,\end{aligned} \tag{2.181}$$

and in these coordinates a basis of scalar harmonics is given by

$$\mathcal{Y}^{I_s}(\varphi_1,\theta_1,\varphi_2,\theta_2,\psi) = e^{i\frac{R}{2}\psi}e^{im_1\varphi_1}e^{im_2\varphi_2}d^{(j_1)}_{m_1\frac{R}{2}}(\theta_1)d^{(j_2)}_{m_2\frac{R}{2}}(\theta_2) \tag{2.182}$$

[24] We use the conventions of [64], except that angular indices are denoted $\theta, \phi, \ldots$, $\tilde{g}^{\text{here}}_{mn} = g^{\text{there}}_{mn}$, and $\breve{g}^{\text{here}}_{ij} = \tilde{g}^{\text{there}}_{\theta\phi}$.

where $I_s \equiv \{j_1, m_1, j_2, m_2, R\}$ is a multi-index, $d^{(j)}_{m_1 m_2}(\theta)$ is the Wigner (small) d-matrix, and $\theta_i \in [0, \pi)$, $\phi_i \in [0, 2\pi)$, and $\psi \in [0, 4\pi)$. We will only need the scalar harmonics in the following, with eigenvalues $\lambda_s(j_1, j_2, R) = 6(j_1(j_1+1) + j_2(j_2+1) - R^2/8)$.

We will only focus on the radial scaling of the dominant metric perturbation. The presence of the anti-D3-brane charge on the NS5-brane may be interpreted as N anti-D3-branes smeared over the two-cycle wrapped by the NS5-brane. The backreaction is heavily dependent on the geometric details of this smearing, so the reported radial scalings may be reduced by suitable geometric tuning. However, we expect a generic smearing to source all possible angular modes and any order-of-magnitude estimates to be set by the dominant mode.

When placed in the background (2.178), the anti-D3-branes will feel a force towards small r. In the actual configuration, interactions with the anti-NS5-brane provide a stabilizing force that keeps the three-brane charge localized around the two-cycle, but the system (2.177) does not account for this force. The effects of the stabilizing force could be included by sourcing appropriate perturbations in the warped throat, but doing so would leave the radial scaling of the dominant perturbation unchanged, and so our analysis applies in any case.

For an anti-D3-brane at (r', Ψ'), we find that

$$\delta\Phi^{I_s}_{-}(r; r', \Psi') = -\frac{1}{\Delta_s}\frac{g_s \ell_s^4}{L^4}\frac{r'^4}{L^4}\left(\left(\frac{r'}{r}\right)^{2+\Delta_s}\theta(r-r') + \left(\frac{r}{r'}\right)^{\Delta_s-2}\theta(r'-r)\right)\mathcal{Y}^{I_s}(\Psi'), \tag{2.183}$$

so in the area of interest,

$$\delta\Phi^{I_s}_{-}(r; r', \Psi') \propto -\frac{1}{\Delta_s}\frac{N}{N_{\mathrm{D3}}}\frac{r'^4}{L^4}\left(\frac{r'}{r}\right)^{2+\Delta_s}, \tag{2.184}$$

where $\Delta_s \equiv \sqrt{4+\lambda(I_s)}$. This Φ_- profile induces a metric perturbation

$$\pi^0(r) = 0 \qquad \pi^{I_s}(r) \propto r^2 \frac{N}{N_{\rm D3}} \left(\frac{r'}{r}\right)^{6+\Delta_s}, \tag{2.185a}$$

$$\tau^0(r) \propto \frac{N}{N_{\rm D3}} \left(\frac{r'}{r}\right)^{8} \qquad \tau^{I_s}(r) \propto -\frac{N}{N_{\rm D3}} \left(\frac{r'}{r}\right)^{6+\Delta_s}. \tag{2.185b}$$

From the spectroscopy of $\mathrm{T}^{1,1}$, the lowest scalar mode has quantum numbers $(\frac{1}{2},\frac{1}{2},\pm 1)$ and $\Delta_s = 7/2$, so the dominant metric perturbation is

$$\delta\tilde{g}_{mn}\mathrm{d}y^m\,\mathrm{d}y^n \propto \frac{N}{N_{\rm D3}} \left(\frac{r'}{r}\right)^{19/2} \mathcal{Y}^{\frac{1}{2},\frac{1}{2},1}(\Psi)\, r^2 \breve{g}_{\theta\phi}\,\mathrm{d}\Psi^\theta\,\mathrm{d}\Psi^\phi \tag{2.186}$$

where $\mathcal{Y}^{\frac{1}{2},\frac{1}{2},1}(\Psi)$ is some real superposition of angular harmonics with $(j_1,j_2,R) = (\frac{1}{2},\frac{1}{2},\pm 1)$.

The perturbations in both the warp factor and internal metric will alter the gauge coupling function,

$$\frac{8\pi^2}{g_{\rm YM}^2} = \frac{4\pi}{\ell_s^4}\int_{\Sigma_4} \mathrm{d}^4\xi\,\sqrt{\tilde{g}_4}\,e^{-4A}, \tag{2.187}$$

on a stack of D7-branes wrapping a divisor Σ_4, such that

$$\delta\left(\frac{8\pi^2}{g_{\rm YM}^2}\right) = \frac{4\pi}{\ell_s^4}\int_{\Sigma_4} \mathrm{d}^4\xi\,\sqrt{\tilde{g}_4}\Big(-2\Phi_+^{-2}\left(\delta\Phi_+^{(1)}+\delta\Phi_-^{(1)}\right) + \Phi_+^{-1}\tilde{g}^{ab}\delta\tilde{g}_{ab}\Big). \tag{2.188}$$

We proved in Appendix 2.B that, in order for these D7-branes to couple to $c = \ell_s^{-2}\int_{\Sigma_2} C_2$ supersymmetrically, Σ_4 must not only descend into the warped throat but actually intersect Σ_2. We expect the supergravity description to break down near the intersection and a local model to be more apt and, from the open-string picture discussed in §2.1, we expect this contribution to be $\mathcal{O}(N)$. Furthermore, away from the intersection, the supergravity approximation becomes accurate and we have shown above N D3-branes induce an $\mathcal{O}(N/N_{\rm D3})$ fractional perturbation. This contribution is then

$$\delta\left(\frac{8\pi^2}{g_{\rm YM}^2}\right) \propto N_{\rm D3}\int_{\Sigma_4} \mathrm{d}^4\xi\,\sqrt{\tilde{g}_4}\,r^{-4}\underbrace{\Big(-2\Phi_+^{-1}\left(\delta\Phi_+^{(1)}+\delta\Phi_-^{(1)}\right)+\tilde{g}^{ab}\delta\tilde{g}_{ab}\Big)}_{\mathcal{O}(N/N_{\rm D3})} \propto N. \tag{2.189}$$

CHAPTER 3

SYSTEMATICS OF AXION INFLATION IN CALABI-YAU HYPERSURFACES

Abstract[1]

We initiate a comprehensive survey of axion inflation in compactifications of type IIB string theory on Calabi-Yau hypersurfaces in toric varieties. For every threefold with $h^{1,1} \leq 4$ in the Kreuzer-Skarke database, we compute the metric on Kähler moduli space, as well as the matrix of four-form axion charges of Euclidean D3-branes on rigid divisors. These charges encode the possibility of enlarging the field range via alignment. We then determine an upper bound on the inflationary field range $\Delta\phi$ that results from the leading instanton potential, in the absence of monodromy. The bound on the field range in this ensemble is $\Delta\phi \lesssim 0.3 M_{\rm pl}$, in a compactification where the smallest curve volume is $(2\pi)^2\alpha'$, and we argue that the sigma model expansion is adequately controlled. The largest increase resulting from alignment is a factor ≈ 2.6. We also examine a set of threefolds with $h^{1,1}$ up to 100 and characterize their axion charge matrices. While we find modest alignment in this ensemble, the maximum field range is ultimately suppressed by the volume of the internal space, which typically grows quickly with $h^{1,1}$. Furthermore, we find that many toric divisors are rigid—and the corresponding charge matrices are relatively trivial—at large $h^{1,1}$. It is therefore challenging to realize alignment via superpotentials generated only by Euclidean D3-branes, without taking into account the effects of flux, D7-branes, and orientifolding.

[1] This chapter is based on C. Long, L. McAllister and J. Stout, "Systematics of Axion Inflation in Calabi-Yau Hypersurfaces," *JHEP* **02** (2017) 014, [1603.01259].

We thank A. Braun and J. Halverson for discussions, and thank V. Khrulkov, M. Stillman, and B. Sung. We are particularly indebted to M. Stillman for many helpful explanations.

3.1 Introduction

The prospect of detecting or strongly bounding primordial gravitational waves through measurements of CMB B-modes in the next few years makes the question of large-field inflation in quantum gravity an urgent one. Exhibiting a totally explicit model of large-field inflation in string theory, or proving no-go theorems that exclude classes of constructions, remains challenging. A persistent difficulty is establishing control of the theory in the parameter range where large-field inflation would occur: making the inflaton potential flat over a super-Planckian distance often requires adjusting compactification parameters, such as cycle sizes, flux quanta, and numbers of D-branes, away from the weakly coupled limit. While it is easy to speculate that something that appears difficult might in fact be impossible, and some authors have promoted this expectation to a principle, there has been little success in actually establishing that large-field inflation is (im)possible in some corner of string theory, except in very simple settings.[2]

Axion inflation is a promising framework for examining large-field inflation in string theory. As in the original model of natural inflation [65], all-orders shift symmetries give structure to the inflaton potential and sharpen the problem of exhibiting a flat potential over a large range to that of achieving a large axion periodicity. Axions are numerous in Calabi-Yau compactifications of string theory, descending from p-form fields in ten dimensions, reduced on suitable p-cycles. The resulting axion fields inherit perturbatively exact continuous shift symmetries from the higher-dimensional gauge symmetry, provided that the latter is not broken by classical sources such as wrapped D-branes or background fluxes, which would

[2] See [8] for an overview.

introduce monodromy in the axion potential [33, 32].[3] In this work we will consider axion inflation without explicit monodromy: we will investigate inflation driven by the strictly periodic potential generated by Euclidean D-branes.

Although it is difficult to arrange for a single axion in string theory to have periodicity $2\pi f > M_{\rm pl}$ in a regime of perturbative control [67, 68], an appealing alternative is to arrange for a particular linear combination of $N > 1$ axions to have a large effective periodicity. The resulting inflationary model, aligned natural inflation, is a version of assisted inflation [69]. The first such proposal, for the case $N = 2$, is due to Kim, Nilles, and Peloso (KNP) [44], and is known as 'KNP alignment' or 'lattice alignment.'

More recently, generalizations of lattice alignment to $N \gg 1$ have been studied [52, 70, 48], and a distinct alignment phenomenon involving the kinetic term, known as 'kinetic alignment,' has been identified [45]. Related works include [71, 72, 73, 74, 75, 76, 72, 77, 78, 79, 80, 81, 82, 83]. In §2 we will review these alignment effects in more detail. One key point is that the field range enhancement due to lattice alignment is determined by a matrix Q of quantized axion charges carried by instantons, which without loss of generality we can take to be integers. In an effective field theory construction of aligned natural inflation, the axion periodicity can be made arbitrarily large if these integer charges are unbounded. However, quantum gravity theories with conventional black hole thermodynamics are generally thought not to allow exact continuous global internal symmetries. More concretely, any finite class of string compactifications will be characterized by a finite set of integer data—such as intersection numbers, flux quanta, and D-brane charges—which only allows for a finite degree of alignment. While this plausibly excludes *arbitrarily* super-Planckian field ranges in axion theories without

[3] See e.g. the discussions in [66, 42].

monodromy, the question of physical interest is whether the field range $\Delta\phi_{\rm thy}$ allowed by quantum gravity can exceed the upper bound[4] $\Delta\phi_{\rm exp}$ determined by measurements of CMB B-modes.

To determine what quantitative upper bound quantum gravity, and in particular string theory, imposes on the field range in axion inflation, one can ask whether the integer data in an actual string compactification can permit a high degree of alignment, and whether this is sufficient to achieve $\Delta\phi_{\rm thy} > \Delta\phi_{\rm exp}$ in a parametrically controlled construction. In this paper, we answer these questions, in the negative, for a large class of explicit Calabi-Yau compactifications.

We consider inflation driven by the Ramond-Ramond four-form C_4, in compactifications of type IIB string theory on Calabi-Yau threefold hypersurfaces in toric fourfolds. We examine all 5922 threefolds with $h^{1,1} \leq 4$ in the Kreuzer-Skarke database [85], and identify divisors that are rigid and so support Euclidean D3-brane contributions to the superpotential.[5] In 4390 of these compactifications, Euclidean D3-branes wrapping linear combinations of up to three toric divisors suffice to break all continuous axion shift symmetries, and correspondingly lift all flat directions in the Kähler moduli space.[6] The axion fundamental domain is therefore compact in these examples, and we compute its diameter as a function of the Kähler moduli. The geometric field range $\mathcal{R} \approx \Delta\phi_{\rm thy}$, defined in §3.2.1, is a function of the curve volume parameters t_i, and is homogeneous of degree -2 with respect to the overall scaling $t_i \to \lambda t_i$, so the upper bound on $\Delta\phi$ is dictated, in

[4]For single-field natural inflation, the Planck measurements of the tilt also imply a lower bound on $\Delta\phi$ [84].

[5]Our method is applicable for larger $h^{1,1}$, as we show in §3.5, but computing the divisors' topology becomes more expensive.

[6]The flat directions in the remaining examples may well be lifted by more complicated instanton configurations, but we do not analyze those geometries any further.

part, by the smallest curve volumes compatible with control of the α' expansion. We argue that in a region of reasonable perturbative control, where the minimum curve volume is $\ell_s^2 \equiv (2\pi)^2\alpha'$, the upper bound on the geometric field range is $\mathcal{R} \lesssim 0.3\, M_{\rm pl}$, with $M_{\rm pl}$ the four-dimensional reduced Planck mass.

The largest contribution of lattice alignment to $\mathcal{R}$ in our ensemble is a factor of 2.6, in a compactification where $h^{1,1} = 4$ with axion charge matrix

$$\boldsymbol{\mathcal{Q}} = 2\pi \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.1}$$

In this example $\mathcal{R} = 0.08 M_{\rm pl}$, while with $\mathcal{Q} = 2\pi\mathbb{1}$ one would have $\mathcal{R} = 0.03 M_{\rm pl}$.

We make one simplifying assumption that deserves special mention. In determining which divisors D yield Euclidean D3-brane contributions to the superpotential, we examine only the topology of D itself, and require the rigidity condition $h^\bullet(D, \mathcal{O}_D) = (1, 0, 0)$. We do not systematically include corrections to this zero-mode counting due to orientifolding, worldvolume flux, bulk flux, or intersections with seven-branes (see e.g. [86, 87, 88, 89, 90, 91]). While incorporating these effects is beyond the scope of this work, it will be an important next step.

The organization of this paper is as follows. In §3.2 we review how the topological and geometric data of an O3/O7 orientifold compactification determines an effective theory for axions, and we explain how to compute the field range, including the effects of alignment, in such a theory. In §3.3 we recall how to obtain the topological data of a Calabi-Yau threefold hypersurface in a toric variety. In §3.4 we present the results of a complete scan through the Kreuzer-Skarke database at $h^{1,1} \leq 4$, and in §3.5 we describe a few examples at much larger $h^{1,1}$. Our conclusions appear in §3.6.

3.2 Four-Form Axions in O3/O7 Orientifolds

A comparatively well-understood class of four-dimensional $\mathcal{N} = 1$ solutions of string theory are compactifications of type IIB string theory on O3/O7 orientifolds of Calabi-Yau threefolds. Because the full space of $\mathcal{N} = 1$ orientifolds is not known,[7] in this work we will focus on their Calabi-Yau double covers, which can be enumerated systematically in the case of hypersurfaces in toric varieties.

3.2.1 The effective Lagrangian

In type IIB string theory compactified on an O3/O7 orientifold of a Calabi-Yau threefold X, the closed string moduli are the complex structure moduli, axiodilaton, and Kähler moduli. The complex structure moduli and axiodilaton can be completely fixed by a suitable choice of quantized G_3 flux, while the Kähler moduli are unfixed to all orders in perturbation theory due to the gauge symmetry of the Ramond-Ramond four-form. When $h^{1,1}_{-} = 0$, which we will assume in this work, the coordinates on Kähler moduli space are the complexified volumes T^i of four-cycles, defined as

$$T^i = \frac{1}{2}\int_{D^i} J \wedge J + i \int_{D^i} C_4 \equiv \tau^i + i\theta^i, \tag{3.2}$$

where J is the Kähler form, D^i is a basis element of $H_4(X, \mathbb{Z})$, and C_4 is the Ramond-Ramond four-form field. The Kähler potential is given by

$$\mathcal{K} = -2 \log \mathcal{V}, \tag{3.3}$$

[7] However, see [92] for progress in classifying involutions that exchange two coordinates.

where $\mathcal{V}$ is the volume[8] of the internal space,

$$\mathcal{V} = \frac{1}{6}\int_X J \wedge J \wedge J. \tag{3.4}$$

We can write the volume as $\mathcal{V} = \frac{1}{6}\kappa^{ijk}t_i t_j t_k$ by expanding the Kähler form as $J = t_i\omega^i$, where ω^i form a basis for $H^{1,1}(X,\mathbb{Z})$ and the κ^{ijk} are triple intersection numbers among divisors D^i.

The space of Kähler parameters t_i is restricted by the requirement that the metric on field space be positive definite. To identify the resulting conditions on the t_i, we consider the Mori cone of X, Mori(X), which is the cone of holomorphic curves: any holomorphic curve C in X can be written as

$$C = \sum_a n_a C_a \tag{3.5}$$

where the C_a are the generators of Mori(X), and n_a are nonnegative integers. The Kähler cone is the space dual to the Mori cone, i.e. it is the region of Kähler parameters t_i for which $\int_C J > 0$ for every holomorphic curve C.

Everywhere inside the Kähler cone, the axion field space metric K_{ij} obtained from the tree-level Kähler potential (3.3) is positive definite. However, as one approaches the walls of the Kähler cone, (3.3) does not necessarily provide a good approximation to the true Kähler potential that incorporates all α' and g_s corrections. Our computation based on (3.3) is therefore meaningful only when the t_i are restricted to a proper subset of the Kähler cone. To understand the conditions that must be imposed on the t_i, we recall the form of the perturbative and nonperturbative corrections to the effective Lagrangian. The superpotential for the Kähler moduli is purely nonperturbative because of the axion shift symmetry, and we will

[8] All volumes in this work are determined in ten-dimensional Einstein frame in units of $\ell_s = 2\pi\sqrt{\alpha'}$.

compute it directly in this work, modulo some important technical assumptions detailed below. The Kähler potential receives perturbative corrections in the α' and g_s expansions, as well as nonperturbative corrections, and none of these has been fully characterized.

Control of the string loop expansion can be achieved by arranging for $g_s \ll 1$ by a suitable choice of quantized three-form flux. We remark that string loop corrections to the Kähler potential are suppressed not only by powers of g_s, but also by powers of $\mathcal{V}$, so at large threefold volume very small g_s is not necessary for ensuring that string loop corrections are small.[9] Next, as a proxy for control of the α' expansion, we will consider worldsheet instantons wrapping nontrivial curves $C \subset X$: in the region where the g_s and α' expansions are well-controlled, these are generically the leading nonperturbative corrections to K, and are proportional to[10]

$$\Delta\mathcal{K} \sim \mathcal{V}^{-1}\, e^{-2\pi\sqrt{g_s}\, t}, \tag{3.6}$$

where t is the Einstein frame volume of C, i.e. the volume measured with the ten-dimensional Einstein frame metric, in units of ℓ_s^2. (The string frame volume of C is then $\sqrt{g_s}t$.) To ensure that the worldsheet instanton corrections are small, we will require that the volumes of all curves are larger than some threshold value. In this work we take the threshold volume to be ℓ_s^2, so that worldsheet instanton contributions are suppressed by factors of $e^{-2\pi\sqrt{g_s}}$, which is small for $g_s \gtrsim 0.1$. Because the Kähler metric is homogeneous of degree -2 with respect to overall scaling $t_i \to \lambda t_i$, it is trivial to translate our results to any other desired threshold, as might be motivated by examining the form of perturbative corrections in particular examples.

[9] Investigations of axion field ranges at moderately strong coupling include [93, 94].

[10] We adopt the normalizations of [95], as laid out in Appendix A of [95].

In view of the above requirement, we now define the *stretched Kähler cone* as the set of Kähler parameters t_i for which $\int_C J > 1$, for all holomorphic curves C. The condition on curve volumes explained in the previous paragraph corresponds to the requirement that the t_i lie in the stretched Kähler cone. This condition leads to a lower bound on the volumes of divisors, $\tau^i \equiv \partial\mathcal{V}/\partial t_i$, and on the volume $\mathcal{V}$ of X itself.

3.2.2 The axion fundamental domain

At a point in Kähler moduli space that falls inside the stretched Kähler cone, the effective Lagrangian for the $N = h^{1,1}$ axions takes the form, in four-dimensional Einstein frame,

$$\mathcal{L} = \frac{M_{\rm pl}^2}{2}\mathcal{R}_4 - \frac{M_{\rm pl}^2}{2}K_{ij}\partial^\mu\theta^i\partial_\mu\theta^j - \sum_{a=1}^{P}\Lambda_a^4\big(1-\cos(Q^a{}_i\theta^i)\big). \tag{3.7}$$

Here K_{ij} is the Kähler metric on field space, and $(2\pi)^{-1}Q$ is a matrix of rational numbers determined by instanton charges. We will search for examples in which Q is a full-rank (that is, rank N) matrix, so that there are no exactly flat directions in the axion field space, and correspondingly no unstabilized[11] Kähler moduli. In order for Q to have rank N, there must be at least N linearly independent divisors contributing to the superpotential, i.e. we must have $P \geq N$.

The *fundamental domain* $\mathcal{F}$ (cf. [46]) of the axions is the region contained in the intersection of the $2P$ half-plane constraints $-\pi \leq Q^a{}_i\theta^i \leq \pi$, as visualized in Figure 3.1. When Q has rank N, $\mathcal{F}$ is compact.

[11] Strictly speaking, we will not be stabilizing the real part Kähler moduli τ_i, in the sense that we will not minimize the scalar potential with respect to the τ_i. We do, however, ensure that all the τ_i appear in the superpotential, in N linearly independent combinations.

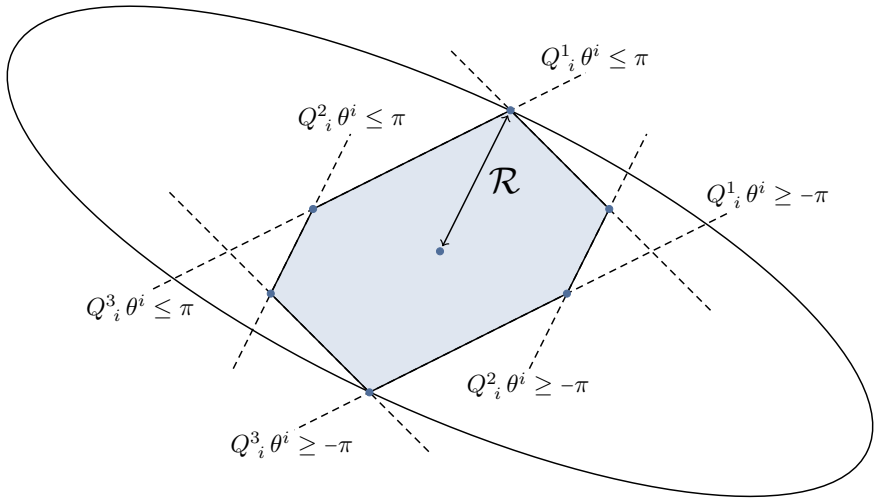

Figure 3.1: The geometric field range $\mathcal{R}$ is the semi-diameter of the fundamental domain $\mathcal{F}$, which is the region contained in the intersection of the $2P$ hyperplane constraints $-\pi \le Q^a{}_i\theta^i \le \pi$. Surfaces of constant distance are ellipsoids with weight matrix K_{ij}.

The fundamental domain is a polytope in field space, and may also be expressed as the convex hull of a set of vertices $\{\mathbf{d}_i\}$. We define the *geometric field range* $\mathcal{R}$ as the distance, measured with respect to the Kähler metric K_{ij}, from the origin to the most distant point on the boundary of $\mathcal{F}$. Equivalently, $\mathcal{R}$ is the distance from the origin to the most distant of the $\mathbf{d}_i$, i.e. $\mathcal{R}$ is the semi-diameter of $\mathcal{F}$.

The length $\Delta\phi$ of an inflationary trajectory driven by a general potential on $\mathcal{F}$ may be larger or smaller than $\mathcal{R}$, but when the initial conditions are arranged so that the trajectory is well-approximated by a straight line, we expect that $\Delta\phi \lesssim \mathcal{R}$. We have verified this expectation by solving for the inflationary evolution that results from the full potential.

The identifications defining $\mathcal{F}$, and hence also the size $\mathcal{R}$ of $\mathcal{F}$, depend on the set of instantons included in the sum in (3.7). Because the Λ_a depend exponentially on four-cycle volumes, there will generally be large hierarchies among the Λ_a, and

so some terms in the axion potential may provide only small ripples that are unimportant in determining the maximum field range. Our approach is to choose the dominant instantons, defined as follows. Given a set of $P > N$ instanton contributions, i.e. P row vectors ${Q^1}_i, \ldots, {Q^P}_i$, one can search for one or more sets of N linearly-independent vectors, corresponding to full-rank square matrices contained in Q. When there are multiple such full-rank sets, we choose the one for which the Λ_a are as large as possible; that is, we identify the $2N$ most important hyperplanes defining the fundamental domain.[12]

Once the dominant rows of Q are identified, the corresponding inequalities define a polytope in field space. The point in this convex polytope furthest from the origin must be one of the vertices $\mathbf{d}_i$. Thus, given a constant Kähler metric $\mathbf{K}$ and a full-rank square matrix ${\mathcal{Q}^i}_j \subset {Q^a}_i$, corresponding to the identifications imposed by the leading instantons, we obtain the axion field range by enumerating the vertices of the associated polytope and computing

$$\mathcal{R}^2 = \max_i \, \mathbf{d}_i^\intercal \cdot \mathbf{K} \cdot \mathbf{d}_i. \tag{3.8}$$

Each choice of $\mathcal{Q}$ will determine a different polytope in field space and thus yield a different value of $\mathcal{R}$. In particular, the semi-diameter of the polytope formed by the $2N$ most important hyperplanes serves as an upper bound on the length of straight-line trajectories that stay within the fundamental domain.

[12]Specifically, we sort the $P > N$ vectors so that the corresponding Λ_a are ordered from largest to smallest. We then select vectors in order from this list, omitting any vector that is not linearly independent of those that have already been selected, and so arrive at a set of N vectors that can be assembled to form a full-rank square matrix $\mathcal{Q}$.

3.2.3 The superpotential

In the type IIB orientifolds considered in this work, the superpotential interactions of the Kähler moduli T^i are generated by nonperturbative effects, either from Euclidean D3-branes on a divisor D in the Calabi-Yau X, or from strong gauge dynamics, such as gaugino condensation, on a stack of seven-branes on a divisor D in X. As explained above, we will restrict our attention to Euclidean D3-branes. Necessary and sufficient conditions for a Euclidean D3-brane contribution to the superpotential were given in [96]. These conditions were derived in the case of M-theory compactified on an elliptically fibered Calabi-Yau fourfold Y_4, with base B_3. Consider a Euclidean M5-brane wrapping a smooth divisor $\hat{D} \subset Y_4$. Two necessary conditions for a superpotential contribution are that $\hat{D}$ is *vertical*, meaning that $\pi(\hat{D})$ is a divisor of B_3, and that $\hat{D}$ is *effective* (see e.g. [97] for the definition). Granting these requirements, a final condition sufficing for a contribution is the rigidity condition

$$h^\bullet(\hat{D}, \mathcal{O}_{\hat{D}}) = (1, 0, 0, 0). \tag{3.9}$$

We will refer to divisors that obey these conditions as *rigid divisors*, with the vertical and effective conditions being implicit.

To translate (3.9) to a condition on smooth divisors $D = \pi(\hat{D}) \subset B_3$, we use the relation [98]

$$h^i(\hat{D}, \mathcal{O}_{\hat{D}}) = h^i(D, \mathcal{O}_D) + h^{i-1}(D, -\Delta|_D), \quad 0 \leq i \leq 3, \tag{3.10}$$

where $h^a \equiv 0$ when $a < 0$. Here $12\Delta = \sum n_i \Sigma^i$, where the Σ^i are the loci where the fiber degenerates, and the n_i denote the type of singularity. Since the $h^{i-1}(D, -\Delta|_D)$ are nonnegative, a necessary condition on D in order for $\hat{D}$ to fulfill the sufficient condition (3.9) is that $h^i(D, \mathcal{O}_D) = 0$, $i = 1, 2$. In the special case that

the degeneration locus of the elliptic fiber does not intersect D—in weak coupling terms, this means that D does not intersect divisors wrapped by D7-branes—we have $h^{i-1}(D, -\Delta|_D) = 0$, so that

$$h^\bullet(D, \mathcal{O}_D) = (1, 0, 0) \tag{3.11}$$

actually suffices to ensure a superpotential contribution. In summary, a divisor $D \subset B_3$ that is effective, does not intersect the discriminant locus 12Δ, and obeys the rigidity condition (3.11) supports a Euclidean D3-brane contribution to the superpotential: its preimage $\hat{D} = \pi^*(D)$ is effective, vertical, and obeys (3.9).

We have emphasized the 'threefold rigidity condition' (3.11) because it depends only on the base B_3, and so can be assessed directly from the combinatorial data in the Kreuzer-Skarke database. A more comprehensive analysis, also applicable to divisors D that intersect Δ, would require information about the elliptic fibration, which in our framework requires specifying an orientifold of the Calabi-Yau threefold X whose image is the non-negatively curved base B_3. A systematic treatment of all $\mathbb{Z}_2$ involutions is beyond the scope of this work. We will work with the Kähler potential $\mathcal{K} = -2\log\left(\frac{1}{2}\mathcal{V}\right)$, where $\mathcal{V}$ is the volume of the double-cover Calabi-Yau manifold. This provides a reasonable proxy for the metric on the Kähler moduli space of the orientifold, at least in the case of orientifolds that flip a single toric coordinate $x_i \to -x_i$. In such a case, we have $h^{1,1}_+ = h^{1,1}(X)$, and we do not expect the orientifold action to significantly change the intersection ring. Beyond the effects of orientifolds themselves, it is worth noting that incorporating D7-branes provides additional freedom to increase the field range $\mathcal{R}$, by factors of the dual Coxeter numbers of the condensing gauge groups.[13]

[13] A string theory embedding of this proposal was considered in [99], where the enhancement was realized by multiply-wound D7-branes.

Let us be very clear on this point: an effective divisor D obeying (3.11) that does not intersect seven-branes (including O7-planes) will yield a Euclidean D3-brane superpotential term; but because we are working directly with threefolds, without either orientifolding or taking a weak-coupling limit from a fourfold, the non-intersection condition is a simplifying assumption that is not verifiable in our framework. We view this approach as an intermediate step between working only with the $\mathcal{N} = 2$ data of a threefold, and performing a full $\mathcal{N} = 1$ analysis complete with explicit orientifolding.

The superpotential that results takes the form

$$W = W_0 + \sum_{\alpha=1}^{p} A_\alpha \, e^{-2\pi q^\alpha{}_i T^i}, \tag{3.12}$$

where W_0 is a flux-dependent constant, and A_α are Pfaffians that depend on the vacuum expectation values of the complex structure moduli. The constant matrix $q^\alpha{}_i$ specifies which Kähler moduli appear in each non-perturbative contribution to the superpotential; at the level of our analysis each of the p linear combinations $\text{Đ}^\alpha \equiv q^\alpha{}_i D^i$ corresponds to a rigid divisor. The full supergravity potential is given by

$$V = e^{\mathcal{K}} \left(\mathcal{D}_i W \overline{\mathcal{D}^i W} - 3|W|^2 \right), \tag{3.13}$$

where $\mathcal{D}_i = \partial_i + K_i$ is the Kähler covariant derivative. In this work, we will assume that the moduli can be stabilized in a vacuum where the cosmological constant is small in string units, and that the dynamics of the real-part saxions may be ignored.[14] The effective Lagrangian density for the axions θ^i is then given by Eq. 3.7, where

$$Q^a{}_i = 2\pi \begin{pmatrix} q^\alpha{}_i \\ q^\beta{}_i - q^\gamma{}_i \end{pmatrix}, \tag{3.14}$$

[14] Ignoring the saxions would be untenable in a construction of an inflationary solution, but is reasonable here because we are simply deriving upper bounds on the geometric diameter.

is a $P \times N$ matrix, with $P = p(p+1)/2$. The last $p(p-1)/2$ rows consist of differences ${q^{\beta}}_i - {q^{\gamma}}_i$ with $\beta > \gamma$, which result from cross terms in (3.13).

In summary, the axion charge matrix Q, whose rows specify the hyperplanes that define the fundamental domain, is given by (3.14), where $\text{Đ}^{\alpha} = {q^{\alpha}}_i D^i$, $\alpha = 1, \ldots, p$ are p effective divisors of the threefold X that fulfill the rigidity condition (3.11), and so support Euclidean D3-brane contributions to the superpotential. We now turn to understanding the impact of the axion charge matrix Q on the size $\mathcal{R}$ of the fundamental domain.

3.2.4 Computing the field range

In §3.2.3 we explained how to obtain the data of the periodic identifications defining the axion fundamental domain $\mathcal{F}$, which are determined by the particular divisors Đ^{α} that are rigid and so support Euclidean D3-brane superpotential terms. These identifications correspond to the hyperplanes in Figure 3.1. We also recalled, in §3.2.1, how to compute the Kähler metric K_{ij} and to determine the region of the two-cycle size parameters t_i for which the α' and g_s expansions are well-controlled (the 'stretched Kähler cone'). This metric corresponds to the ellipse in Figure 3.1. These data completely specify the geometry of $\mathcal{F}$, or more precisely the possible geometries of $\mathcal{F}$: the size $\mathcal{R}$ of $\mathcal{F}$ depends on the Kähler moduli. To determine the maximal field range in a given theory, we must maximize $\mathcal{R}$ subject to the linear constraints on the t_i that define the stretched Kähler cone.

The tree-level metric is a homogeneous function of the t_i, and scales with the overall volume as $\mathcal{V}^{-4/3}$. By using the scaling $t_i \to \lambda t_i$, one finds that the maximal field range is achieved on the boundary of the stretched Kähler cone. If $\text{Mori}(X)$ is

simplicial, the $h^{1,1}$ constraints $\int_{C_i} J = 1$ can be simultaneously fulfilled at the apex of the stretched Kähler cone, where all of the two-cycle volumes are set to unity (in units of ℓ_s), and the maximal field range is achieved at the apex. However, in more general cases the point in the stretched Kähler cone giving the largest $\mathcal{R}$ can occur on a wall, but away from the apex. We therefore searched the stretched Kähler cone numerically to determine the optimal field range. For the purpose of the search, we retained only $h^{1,1}$ terms in the potential, taking

$$V = \sum_{i=1}^{h^{1,1}} \left(1 - \cos\left(\mathcal{Q}^i{}_j \theta^j\right)\right), \tag{3.15}$$

where $\mathcal{Q}^i{}_j$ is the leading-order[15] full rank piece of the full Q. In general there will be further terms that reduce the size the of the fundamental domain, both from additional instantons and from cross terms in the supergravity potential, but because we are quoting an upper bound these can be omitted at this stage.

To search for the maximal $\mathcal{R}$, we computed the four-cycle volumes at a reference point $\mathbf{t}_0$, and then extracted the full rank piece of Q that is leading order at $\mathbf{t}_0$. We then scanned over the stretched Kähler cone for the point $\mathbf{t}_L$ with the largest $\mathcal{R}$. For the reference point, we used the apex of the stretched Kähler cone, defined as the point where the Euclidean norm of the vector $(v_1, \ldots, v_{N_C})$ is minimized, where the $v_a \equiv \int_{C_a} J$ and C_a are the N_C generators of the Mori cone. We then checked that $\mathcal{Q}$ at $\mathbf{t}_L$ is the same as at $\mathbf{t}_0$, meaning that the same instantons remain dominant, and the analysis is self-consistent. In a small fraction of cases we found that the set of dominant instantons changed during the exploration from $\mathbf{t}_0$ to $\mathbf{t}_L$, which we then accounted for in computing the field range.

[15] After the $h^{1,1}$ most important terms have been determined, by comparing their prefactors Λ_a according to the algorithm given in §3.2.2, the problem becomes purely geometric, and we can then set all $\Lambda_a = 1$, as we have done in (3.15).

3.2.5 Alignment

Many authors have argued that quantum gravity will censor super-Planckian field displacements, or at least will do so in sufficiently restrictive circumstances. The large degree of structure imposed on axion theories by all-orders shift symmetries makes these theories a promising setting for directly quantifying the restrictions, if any, that descend from quantum gravity. The objective of the present work is to compute the size[16] $\mathcal{R}$ of the axion fundamental domain $\mathcal{F}$ in an ensemble of string compactifications.

Once the rigid divisors $\text{Đ}^\alpha \equiv q^\alpha{}_i D^i$, the Kähler metric K_{ij}, and the stretched Kähler cone have been determined in a particular theory, the size $\mathcal{R}$ of $\mathcal{F}$ is completely specified, and one could mechanically apply the process described in §3.2.4 to compute $\mathcal{R}$ in a large number of examples, as we shall do in §3.4. However, it will be valuable to first explain that a suitable structure in the axion charge matrix could lead to $\mathcal{R} \gg M_{\rm pl}$, even while the eigenvalues of K_{ij} remain $\ll M_{\rm pl}^2$: this is the celebrated phenomenon of *alignment*, and more precisely of KNP alignment [44], also known as lattice alignment. Here we will attempt to be very precise about the notion of lattice alignment in a Calabi-Yau compactification.

Roughly speaking, an axion theory may be said to manifest lattice alignment when the size of $\mathcal{F}$ is larger than it 'would have been if Q had been trivial,' i.e. the notion of alignment is that of an increase in field range resulting from the structure of the axion charge matrix Q. Heuristically, one might try to define the alignment

[16] We stress that including a source of monodromy, which we will not do here, may ultimately allow displacements $\mathcal{O}(n\mathcal{R})$, where $n \in \mathbb{Z}$ is the number of cycles (also known as windings) of monodromy.

enhancement[17] factor η as

$$\eta_{\text{naive}} \stackrel{?}{=} \frac{\mathcal{R}_{\text{actual}}}{\mathcal{R}_{Q=2\pi\mathbb{1}}} \,. \tag{3.16}$$

The numerator is well-defined in general, but the axion charge matrix *alone* is not invariant under a change of the variables θ^i, so stating that $Q = 2\pi\mathbb{1}$ presupposes a choice of basis. The (physically meaningful) field range $\mathcal{R}$ is of course invariant under the change of variables $\theta^i \to M^i{}_j \theta^j$, with $M \in GL(N, \mathbb{R})$, but K_{ij} and Q separately transform.

Why talk about alignment at all, if a precise definition is subtle (though achievable, see below)? One motivation is that it is generally far easier to compute the classical geometric data determining the metric K_{ij} than it is to determine the nonperturbative, quantum data of Q, which after all is a matrix of axionic charges carried by (D-brane) instantons. As such, one may sometimes know K_{ij} without knowing Q, and it would then be valuable to understand how large an error might be made by approximating $Q \approx 2\pi\mathbb{1}$. In systems of $N \gg 1$ axions, including the ensemble studied here with $2 \leq N \leq 100$, this error can easily be a factor of order N, and in theories with special structure [52, 48] (not established to date in string theory) the error can be exponential in N.

If we were equipped with a canonical choice of basis $\mathcal{B}$, we could define the denominator in (3.16) by taking the 'reference' charge matrix to read $Q = 2\pi\mathbb{1}$ *in the basis* $\mathcal{B}$. In other words, the degree of alignment would be dictated by the extent to which the actual charge matrix Q, expressed in the basis $\mathcal{B}$, differs from $2\pi\mathbb{1}$, as quantified by (3.16).

We are not aware of a natural and fully-specified canonical basis. However,

[17] When $\eta > 1$, we say that the theory manifests alignment, and when $\eta < 1$ the result may be termed anti-alignment.

a natural but (in general) overcomplete set consists of the minimal generators of $\mathrm{Eff}(X)$, the cone of effective divisors in X. The number N_{Eff} of minimal generators $\mathcal{E}_A$, $A = 1, \ldots, N_{\mathrm{Eff}}$ of $\mathrm{Eff}(X)$ often exceeds $h^{1,1}$, and there is then no unique choice of a basis for $H^{1,1}(X)$: there are finitely many choices.

Assume for the moment that $N_{\mathrm{Eff}} = h^{1,1}$, so that the generators $\mathcal{E}_A$ of $\mathrm{Eff}(X)$ define a unique basis $\mathcal{B}$. If each of the $\mathcal{E}_A$ were rigid and supported a Euclidean D3-brane contribution to the superpotential, we would have $\mathcal{Q} = 2\pi\mathbb{1}$ in the basis $\mathcal{B}$ defined by the $\mathcal{E}_A$. Moreover, because every effective divisor is a linear combination of the $\mathcal{E}_A$ with nonnegative integer coefficients, the Euclidean D3-branes supported on the $\mathcal{E}_A$ correspond to the most important instanton contributions in the theory: any additional rigid divisors will have equal or larger action. This simple theory, in which the minimal generators $\mathcal{E}_A$ of $\mathrm{Eff}(X)$ are rigid, serves as a reference case that we define to have trivial alignment ($\eta = 1$).

We now propose that a natural definition of a trivial charge matrix Q is the matrix whose rows are the minimal generators $\mathcal{E}_A$ of $\mathrm{Eff}(X)$, even when $N_{\mathrm{Eff}} > h^{1,1}$. In other words, a well-defined null hypothesis for examining alignment is the assumption that each of the $\mathcal{E}_A$ gives an independent contribution to the non-perturbative superpotential. We may then define the enhancement factor η as

$$\eta = \frac{\mathcal{R}_{\mathrm{actual}}}{\mathcal{R}_{\mathrm{Eff}(X)}}, \tag{3.17}$$

where $\mathcal{R}_{\mathrm{actual}}$ is computed using the Q generated by the rigid divisors $Ð^\alpha$, and $\mathcal{R}_{\mathrm{Eff}(X)}$ is computed using the (by definition) trivial Q generated by assuming that the minimal effective divisors $\mathcal{E}_A$ are rigid. For both the numerator and the denominator only the $h^{1,1}$ most important rows of Q are included, as explained in §3.2.2.

Although we have now given a precise definition of the enhancement η resulting

from lattice alignment, it remains to determine whether η can be large in actual string compactifications. We therefore turn to determining the numbers $q^{\alpha}{}_{i}$ in an ensemble of Calabi-Yau geometries.

3.3 The Topology of Calabi-Yau Hypersurfaces

Calabi-Yau hypersurfaces in toric fourfolds provide a large ensemble of Calabi-Yau threefolds, and allow an efficient combinatorial approach to determining the geometry [100, 85]. We refer the reader to [101], among many others, for an introduction to the subject.

The combinatorial data needed to construct a Calabi-Yau consists of a dual pair of reflexive polytopes Δ and Δ°, and a triangulation of Δ° that defines a fan F. F then defines a toric variety V, and the anticanonical hypersurface $-K$ in V is a Calabi-Yau threefold X. The triangulation of Δ° must be *star* with respect to the origin, meaning that every simplex must contain the origin, in order to define a fan. In addition, the triangulation must be *fine* and *regular*, in order to ensure that the hypersurface is generic and projective.[18] Because a generic hypersurface misses any given point of V, we can allow V to have pointlike singularities without making a generic threefold singular. As a result, points interior to facets can be ignored when triangulating Δ°.

We have made use of several publicly-available software packages to obtain and analyze triangulations. The algebraic software `Sage` [103] provides a useful interface for working with toric varieties. The triangulations can be performed in `TOPCOM` [104], which has been integrated into `Sage`. In addition, we used the

[18] See [102] for a discussion of these points.

program `PALP` [105] for calculations involving reflexive polytopes, and its `Mori` extension [106] is very powerful in computing relevant topological data at small $h^{1,1}$. Most triangulation algorithms are not specialized to compute star triangulations; instead, all triangulations are computed, and then the star ones are selected. When one is mostly concerned with hypersurfaces with small $h^{1,1}$, whose polytopes are readily triangulated in `TOPCOM`, the cost of computing all triangulations is generally not prohibitive. However, since we will describe some preliminary results at large $h^{1,1}$, we will outline how one can begin to probe these geometries. For $h^{1,1} \lesssim 30$, one can use the algorithms given in [107, 102] to get all the triangulations of the polytope by gluing together the triangulations of individual facets, but this quickly becomes expensive as $h^{1,1}$ grows. However, even when computing all triangulations in this way is impractical, it is possible to obtain a single triangulation very quickly. The method was implicit in [100], and was made very clear in [108]: one simply computes a regular and fine (not star) triangulation of the polytope, and then deletes the lines in the strict interior of the polytope. This induces a regular triangulation of the facets, and then a star triangulation is constructed by drawing a line from the origin to each point in the polytope. Using this method it is easy to compute a single triangulation of any polytope in the Kreuzer-Skarke database; for instance, a triangulation of a polytope whose hypersurface X has $h^{1,1}(X) = 400$ takes about ten seconds on a typical laptop.

The tree-level Kähler potential depends only on the classical volume, and can be computed easily via toric methods, as one only needs the intersection ring and the Mori cone $\mathrm{Mori}(X)$. We will consider only *favorable* hypersurfaces X, i.e. those in which all of the divisors of the Calabi-Yau are inherited from divisors of V; in such cases we have $\mathrm{Mori}(X) \subset \mathrm{Mori}(V)$. Computing $\mathrm{Mori}(X)$ from toric data is challenging, so we take the conservative approach of imposing the Mori cone

conditions inherited from V.[19]

Determining the nonperturbative superpotential is more involved, as we need to know the Hodge numbers of divisors in the hypersurface. In favorable Calabi-Yau threefolds, the vanishing loci of the individual homogeneous coordinates, corresponding to rays in the fan, furnish a generating set of $h^{1,1}+4$ divisors $\check{D}^a$ in the Calabi-Yau. To search for a set of $h^{1,1}$ independent rigid divisors we consider the cohomology of these generators and their linear combinations. Recall that the number of independent homology classes of divisors is counted by $h^{1,1}(X,\mathbb{Z})$. Given a choice of a basis $\{D^i\}$ of divisors, the task at hand is to determine whether a divisor $D=\sum_i a_i D^i$ is rigid. To do so, we need to specify what values the a_i can take. In some cases one can choose a basis such that all holomorphic hypersurfaces can be written as sums of the D^i with non-negative integer coefficients, and the problem reduces to scanning over an $(\mathbb{N})^{h^{1,1}}$ lattice. This happens only when $\mathrm{Eff}(X)$ is simplicial.[20] The effective cone is not simplicial in general, so the ranges of the coefficients a_i are not always obvious. However, one can consider non-negative linear combinations of the generators of $\mathrm{Eff}(X)$, which will by definition generate all effective divisors.

The Hodge numbers of the toric divisors $\check{D}^a$, which correspond to rays in the fan F and therefore to points in Δ°, can be computed via polytope data alone [109], in the same fashion that the Hodge numbers of the Calabi-Yau are computed

[19]Note that $\mathrm{Mori}(X)$ can be a proper subset of $\mathrm{Mori}(V)$. In particular, if a curve C is in V but not in X, the sigma model expansion on X is unaffected by taking the volume of C to zero.

[20]The divisors whose rigidity properties we need to examine are all the divisors D that are effective in X. Because we have selected only favorable hypersurfaces X, all divisors of X are inherited from divisors of V. In this work we will consider only effective divisors in X that are inherited from *effective* divisors in V, but more general effective divisors of X are possible. We thank M. Stillman for explaining this point to us.

in [100, 110]. We provide a brief summary of the results. As mentioned above, the polytopes Δ and Δ° are dual, so there is a one-to-one relation between faces of dimension k, $\Theta^{\circ[k]}$, of Δ°, and faces of dimension $3-k$, $\Theta^{[3-k]}$, of Δ. The divisors D^a can be organized according to their corresponding points in Δ°. Let $l^*(\Theta)$ denote the number of interior points of a face Θ; then:

- For divisors D^a that correspond to vertices $\Theta^{\circ[0]}$ of Δ°, we have $h^\bullet(D, \mathcal{O}_D) = (1, 0, n)$, where $n = l^*\big(\Theta^{[3]}\big)$ and $\Theta^{[3]}$ is the three-dimensional face dual to $\Theta^{\circ[0]}$.
- For divisors D^a that correspond to points v_a interior to one-dimensional faces $\Theta^{\circ[1]}$ of Δ°, we have $h^\bullet(D, \mathcal{O}_D) = (1, n, 0)$, where $n = l^*\big(\Theta^{[2]}\big)$ and $\Theta^{[2]}$ is the two-dimensional face dual to $\Theta^{\circ[1]}$.
- For divisors D^a that correspond to points v_a that are interior to two-dimensional faces $\Theta^{\circ[2]}$ of Δ°, we have $h^\bullet(D, \mathcal{O}_D) = (n, 0, 0)$, where $n = l^*\big(\Theta^{[1]}\big) + 1$ and $\Theta^{[1]}$ is the one-dimensional face dual to $\Theta^{\circ[2]}$. If $n > 1$ then these divisors are reducible.

These facts make computing the Hodge numbers of toric divisors $\check{D}^a$ a simple combinatorial process. However, it is often the case that there are fewer than $h^{1,1}$ linearly-independent rigid toric divisors, and therefore to search for instantons leading to a full rank Q one must consider linear combinations of toric divisors that are not linearly equivalent to a toric divisor. Because such combinations do not simply correspond to rays in the fan, obtaining their Hodge diamonds requires more effort. The Koszul sequence allows one to calculate this data, and has been implemented in the program `cohomcalg` [111, 112], which we used extensively. We refer the interested reader to [111, 112] for details.

$h^{1,1}(X)$	2	3	4
Number of polytopes	36	244	1197
Number of favorable polytopes	36	243	1185
Number of favorable triangulations	48	525	5330
Number of full-rank triangulations	24	262	4104
Full-rank with only smooth divisors	9	199	3214

Table 3.1: Results of the scan over reflexive polytopes with $h^{1,1}(X) \leq 4$.

It is worth remarking that a linear combination of toric divisors that is rigid and irreducible is also necessarily singular.[21] Consider a divisor D that is linearly equivalent to $D_x + D_y$, where D_x and D_y are toric divisors defined by the vanishing of toric coordinates x and y, respectively. In order for D to be rigid we need $h^2(D) = 0$, which implies that $h^2(D_x) = h^2(D_y) = 0$, as taking a linear combination will not affect the presence of these deformations. Then the only polynomial one can write to define the divisor is $xy = 0$. This is singular along the intersection of the divisors $x = y = 0$. If the divisor is irreducible then D_x and D_y must have non-zero intersection, and therefore the point $x = y = 0$ is contained in the space, and D is necessarily singular. We find that of the 4390 triangulations in our ensemble that have a full-rank Q, 3422 remain full rank when only smooth toric divisors are included.

3.4 A Complete Scan at Small $h^{1,1}$

Equipped with the results of §3.2 and §3.3, we computed the relevant topological data of all Calabi-Yau hypersurfaces in the Kreuzer-Skarke database with $2 \leq h^{1,1} \leq 4$. We searched for divisors D that are rigid linear combinations of up to

[21] We thank M. Stillman and B. Sung for helpful explanations of this point.

three toric divisors $\check{D}$,

$$D = n_a \check{D}^a \,, \tag{3.18}$$

where n_a are nonnegative integers obeying $\max_a n_a = 3$ and $\sum_a n_a \leq 3$. We computed the topology of individual toric divisors via polytope data, and that of linear combinations with `cohomcalg`. At $h^{1,1} = 2, 3, 4$ we found that 24, 262, and 4104 triangulations, respectively, have full-rank q matrices resulting from Euclidean D3-branes. The results are summarized in Table 3.1.[22]

The combined field space radii for $h^{1,1} = 2, 3, 4$ are plotted in Figure 3.2. We find the maximum to be $\mathcal{R} \approx 0.5 M_{\rm pl}$, in a case with $h^{1,1} = 3$, but in this example the overall volume of the Calabi-Yau is close to unity, and so the compactification is arguably not within the regime of perturbative control. The next largest is an example with $h^{1,1} = 4$ in which $\mathcal{R} \approx 0.3 M_{\rm pl}$, and where the overall volume is ≈ 20. This example is much better controlled, and therefore gives the upper bound that we report.

In Figure 3.3, we show a histogram of enhancements from lattice alignment, η, for the 4390 geometries with $h^{1,1} = 2, 3$, and 4. As seen in the inset, there is a spike at $\eta = 1$ corresponding to a large fraction of geometries—2180 out of 4390—that experience no enhancement from Q. This occurs when the minimal generators of $\mathrm{Eff}(X)$ are rigid and thus the leading order $\mathcal{Q}$ is trivial. In addition, many of the non-trivial Q-matrices actually decrease the geometric field range $\mathcal{R}$. We find a positive enhancement in 494 examples.

[22] It sometimes happens that two isomorphic hypersurfaces are realized as hypersurfaces in different toric varieties corresponding to different polytopes. Since we are simply performing a scan over the geometries, we will not attempt to distinguish whether two Calabi-Yau hypersurfaces are different but will instead only refer to individual triangulations.

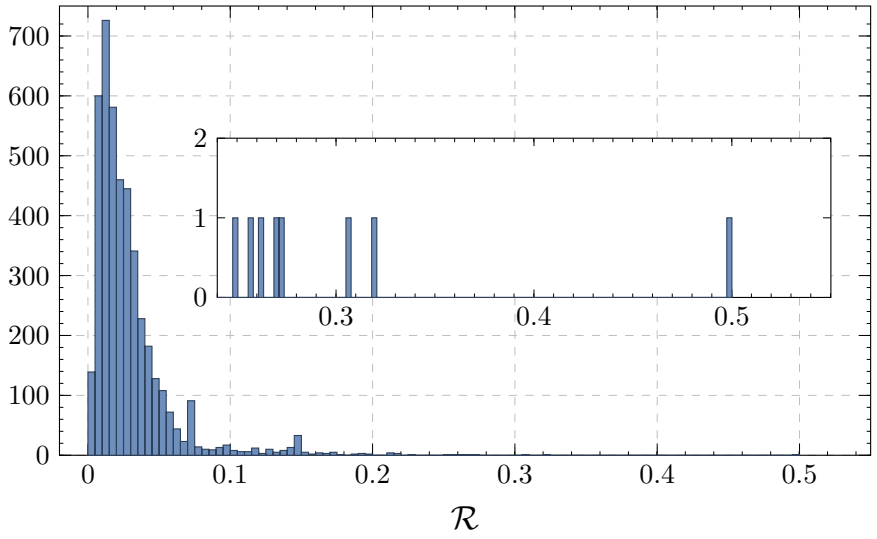

Figure 3.2: Histogram of geometric field ranges $\mathcal{R}$, in units of the reduced Planck mass $M_{\rm pl}$, for $h^{1,1} \leq 4$. The inset shows the tail of the distribution.

In this ensemble, the maximum enhancement from a nontrivial charge matrix is a factor of $\eta = 2.6$ in a threefold with $h^{1,1} = 4$. The vertices of the polytope Δ° are given by

$$\begin{aligned}\mathbf{d}_i = \Big\{&(1,-1,0,0),\ (-1,4,-1,-1),\ (-1,-1,0,0),\ (-1,-1,1,0),\\ &(-1,-1,0,1),\ (-1,2,0,0),\ (-1,-1,1,1)\Big\}.\end{aligned} \tag{3.19}$$

Here we have

$$\boldsymbol{\mathcal{Q}} = 2\pi \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.20}$$

This occurs in an example where the eigenvalues of K_{ij} are quite small, and the geometric field range increases from $0.03 M_{\rm pl}$ to only about $0.079 M_{\rm pl}$. In this example not all of the rigid divisors are smooth. The next largest enhancement is $\eta = 2.55$, which increases the geometric field range from $\mathcal{R} = .05$ to $\mathcal{R} = 0.12$.

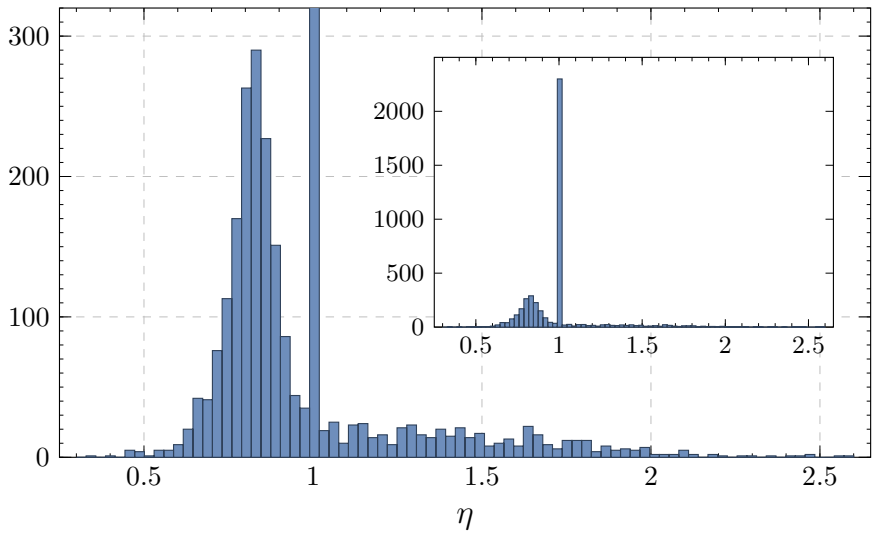

Figure 3.3: Histogram of enhancements η for $h^{1,1} = 2, 3, 4$. Inset demonstrates the large peak at $\eta = 1$, i.e. many geometries see no enhancement in size, or a reduction, from a non-trivial Q.

The vertices of the polytope Δ° are given by

$$\begin{aligned}\mathbf{d}_i = \Big\{&(-1,2,-1,-1),\ (-1,-1,2,1),\ (-1,-1,1,1),\ (1,0,-1,-1),\\ &(-1,-1,1,2),\ (0,-1,1,1),\ (2,1,-2,-2)\Big\},\end{aligned} \tag{3.21}$$

and Q is given by

$$\boldsymbol{\mathcal{Q}} = 2\pi \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 2 & 1 & -1 \end{pmatrix}. \tag{3.22}$$

In this example all of the rigid divisors are smooth.

Although the scan at small $h^{1,1}$ did not yield a geometry that allows a parametrically large fundamental domain, some of the examples exhibit features that could be interesting for inflationary model building. Consider, for instance, the Calabi-Yau hypersurface in the toric variety $(\mathbb{P}^1)^4$. The volume is

$$\mathcal{V} = 2\left(t_1t_2t_3 + t_1t_2t_4 + t_2t_3t_4 + t_1t_3t_4\right), \tag{3.23}$$

and the Mori cone conditions are simply $t_i > 0$, $i = 1, \ldots, 4$. In this geometry, one can make the overall volume arbitrarily large while holding the largest eigenvalue of K fixed, by taking $t_2 = t_3 = t_4 \equiv t_0$ for constant t_0, and letting $t_1 \equiv t \gg 1$. The largest eigenvalue of K is then $1/(144 t_0^4) M_{\rm pl}^2$. This is an appealing feature, as suitably scaling up the volume can provide protection against some perturbative and nonperturbative corrections, while keeping the largest eigenvalue fixed at a sizable value. For instance, by taking $t \to \infty$ and setting $t_0 = 0.2$, the largest eigenvalue of K becomes $4.3 M_{\rm pl}^2$. However, there is a[23] divisor D_s with volume $\tau_s = 6 t_0^2 \approx 0.24$. If there are higher-order instanton contributions[24] $\sim e^{-k\tau_s}$ for $k > 1$, these are not necessarily negligible, e.g. for $k = 2$ their importance relative to the leading term is $e^{-2\pi(0.48)}/e^{-2\pi(0.24)} \sim 0.22$.

3.5 Probing Large $h^{1,1}$

Our analysis thus far has been restricted to small Hodge numbers, $h^{1,1} \leq 4$, but arguments in effective field theory and in random matrix theory suggest that new phenomena will appear in compactifications with $h^{1,1} \gg 1$ [46, 113]. A comparative analysis of these proposals for alignment, and of the requisite degree of fine-tuning at the level of effective field theory, will appear in [113]; here we will briefly summarize the main ideas in order to provide orientation for our search at

[23] The remaining three divisors have large volumes for $t \to \infty$, and their contributions to the superpotential can be neglected.

[24] It is not clear that higher-order contributions from Euclidean D3-branes without flux will be nonvanishing, because $h^0(kD) = h^2(kD) = k$. In fact, in this example are no rigid divisors at all: all of the toric divisors pulled back to the Calabi-Yau hypersurface have the Hodge numbers of K3 surfaces. A superpotential might still be generated if worldvolume fluxes lift the zero modes corresponding to $h^2(D)$ deformations.

large $h^{1,1}$.

An influential early suggestion for alignment of $N \gg 1$ axions was the N-flation proposal [114], where it was observed that the field range of a simple[25] system of N axions is the Pythagorean sum of the ranges of the individual axions, and schematically $\mathcal{R} \propto N^{1/2}$. More recent works have identified stronger enhancements at large N. Multi-axion alignment, the N-dimensional generalization of KNP alignment, yields exponentially large ranges, while plausibly requiring severe fine-tuning [52]. Finally, in [46] it was observed that generic charge matrices could give 'spontaneous' field range enhancements as large as $N^{3/2}$ from a combination of lattice and kinetic alignment. More precisely, the finding of [46] is that for charge matrices Q whose entries are well-approximated as independent and identically distributed (i.i.d.) variables, and are not too sparse, the distribution of field ranges takes the form

$$\mathcal{R} = N^p \zeta \,, \tag{3.24}$$

where $1 \lesssim p \lesssim 3/2$ depends on the sparsity of Q. Here ζ is a positive stochastic variable, varying from one realization of Q to another, that has unit median and a *heavy tail* toward large values: in particular, the mean obeys $\langle \zeta \rangle \gg 1$. The distribution of ζ is computable in special cases. When the entries of Q are such that $QQ^\top$ is a Wishart matrix $\mathcal{W}$, one finds $\zeta \approx \lambda_1(\mathcal{W})^{-1/2}$, with $\lambda_1(\mathcal{W})$ the smallest eigenvalue of $\mathcal{W}$. Because the probability density function of $\lambda_1(\mathcal{W})$ has support near $\lambda_1 = 0$, ζ has a tail toward large positive values. In turn, the range $\mathcal{R}$ has a heavy tail, and one expects to find, after a modest number of independent trials, a range $\mathcal{R}$ that exceeds the median value $\mathcal{R}_{\rm med}$ by orders of magnitude.

Both the engineered N-dimensional alignment of [52], and the spontaneous

[25] The simplifying assumption is that $\mathcal{Q} = 2\pi \mathbb{1}$ in a basis in which K_{ij} is diagonal. This does not hold in generic examples, and in particular is violated in every geometry in our ensemble.

alignment of [46], provide field-theoretic mechanisms for parametrically large field ranges. However, it is clearly necessary to test these ideas in actual string compactifications, in order to understand whether quantum gravity indeed allows these effective theories, and so permits field ranges that are very large in Planck units. To begin exploring this point, we will examine a number of Calabi-Yau hypersurfaces, with $h^{1,1} \in \{50, 60, 70, 80, 90, 100\}$. More systematic results will appear in [115].

3.5.1 Field ranges and volumes

For ten geometries each at $h^{1,1} \in \{50, 60, 70, 80, 90, 100\}$, we computed the relevant topological and metric data and bounded the geometric field range $\mathcal{R}$. Computing the topology of nontrivial linear combinations of toric divisors is computationally expensive at large $h^{1,1}$, so we only searched for rigid divisors among the toric divisors themselves. In many cases the toric divisors suffice to lift all flat directions, and in such cases we bounded the field range. At large $h^{1,1}$, the vertex enumeration problem is computationally taxing and we used alternative methods to obtain the field range.

We may always trivialize $2N$ of the hyperplane constraints via the field transformation

$$\theta^i = \left(\mathcal{Q}^{-1}\right)^i{}_j \Phi^j. \tag{3.25}$$

If $P = N$, this maps the fundamental domain $\mathcal{F}$ into the hypercube of side length 2π. For $P > N$, the $2P$ hyperplane constraints

$$-\pi \leq Q^a{}_i \left(\mathcal{Q}^{-1}\right)^i{}_j \Phi^j \leq \pi \tag{3.26}$$

restrict $\mathcal{F}$ to a hypercube subject to $2(P - N)$ hyperplane 'cuts.' In the Φ^i basis,

distance in field space in Planck units is measured with respect to the metric

$$\boldsymbol{\Xi} = \left(\boldsymbol{\mathcal{Q}}^{-1}\right)^{\top} \cdot \mathbf{K} \cdot \boldsymbol{\mathcal{Q}}^{-1}, \tag{3.27}$$

whose maximum eigenvalue we denote ξ_N^2. If we temporarily ignore the additional constraints (3.26), computing $\mathcal{R}$ via (3.8) involves evaluating the $\boldsymbol{\Xi}$-norm of 2^{N-1} vertices and is thus prohibitively expensive at large N. However, we can always bound the geometric field range by

$$\mathcal{R} \leq \mathcal{R}_{\max} = \pi \sqrt{N} \xi_N M_{\rm pl}\,. \tag{3.28}$$

At large N, eigenvector delocalization generally ensures that the ellipsoid's principal axes are nearly aligned with the diagonals of the hypercube, so that (3.28) is often nearly saturated. Upon including the additional $2(P-N)$ constraints (3.26), the field range will be reduced by the maximally constraining cut, as detailed in [46]. Because we always work with the full-rank square matrix $\mathcal{Q}^i{}_j \subset Q^a{}_i$, we approximate the field range using Eq. 3.28.

In all cases we find $\mathcal{R} \ll M_{\rm pl}$. The mean volume of the Calabi-Yau at the apex of the stretched Kähler cone, as a function of $h^{1,1}$, is plotted in Figure 3.4, and the mean value of ξ_N as a function of $h^{1,1}$ is plotted in Figure 3.5. We also show q_N, the square root of the largest eigenvalue of $\left(\mathcal{Q}\mathcal{Q}^{\top}\right)^{-1}$ in this basis, in Figure 3.6. We find the largest enhancement from lattice alignment occurs at $h^{1,1} = 100$, with $\eta_{\max} = 7.86$. We note that while the effect of alignment can be significant for $h^{1,1} \gg 1$, in our examples this is dwarfed by the growth of the volume with $h^{1,1}$. As $h^{1,1}$ grows, the number of holomorphic curves grows as well, giving more inequalities on the Kähler parameters to stay within the Kähler cone. By demanding that we remain in a regime of control, where all curve volumes are greater than one, the volume is forced to grow quite large (cf. [116, 117]).

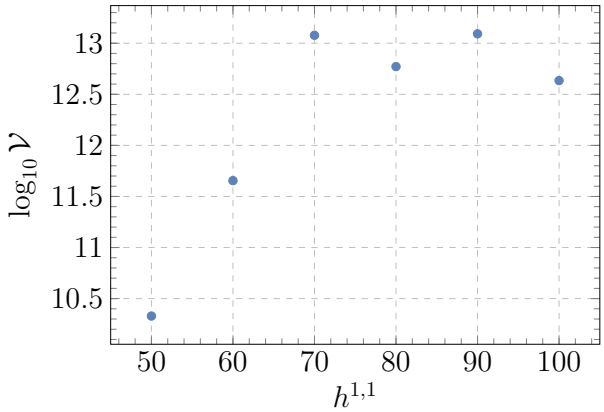

Figure 3.4: $\log_{10}$ of the mean volumes $\mathcal{V}$ as a function of $h^{1,1}$.

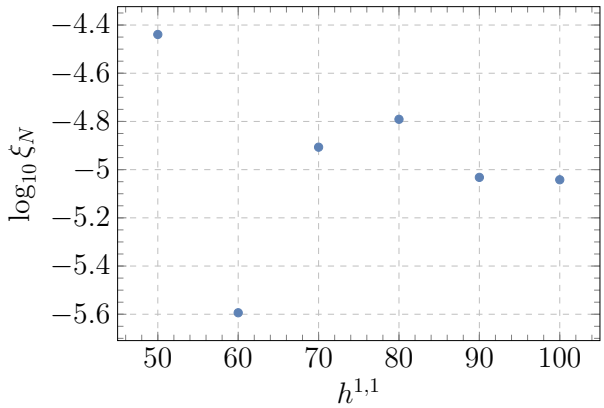

Figure 3.5: Average $\log_{10} \xi_N$ vs. $h^{1,1}$.

These characteristics are in stark contrast with those of the compactification studied by Denef et al. in [95], where $h^{1,1} = 51$, but the volume was stabilized at $\mathcal{V} \sim 50$. In [95] the Kähler moduli were stabilized at a point where the smallest curve volumes were 0.2, but even after scaling up the curve volumes to be ≥ 1, one finds $\mathcal{V} \sim 250$, which is vastly smaller than the volumes we find in hypersurfaces with comparable $h^{1,1}$. A main reason that the volume can be kept small in [95] is that the moduli space is very symmetric. The Calabi-Yau is constructed by taking identical toric patches and gluing them together, so the divisor and curve structure is simply repeated. The result is that the overall volume of the Calabi-Yau does not increase dangerously with the curve volumes. In the two-parameter model

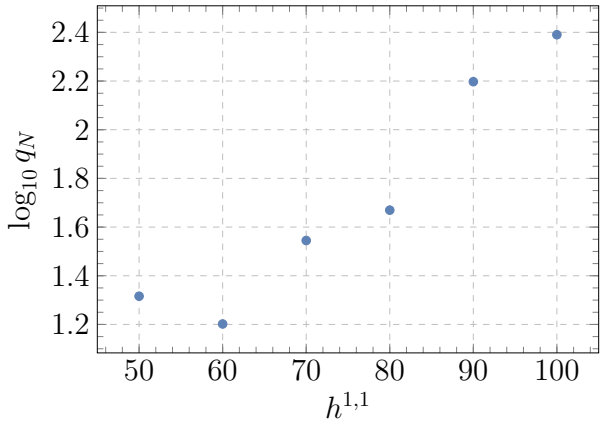

Figure 3.6: Average $\log_{10} q_N$ vs. $h^{1,1}$.

of [95], denoting the volumes of the two classes of curves as s and u, the overall volume takes the form $\mathcal{V} = s^3 + 24s^2u + 96su^2 + 128u^3$, which is simple due to the symmetric intersection structure.

3.5.2 The structure of Q

We have seen that, although the largest eigenvalue $\left(\mathcal{Q}\mathcal{Q}^\top\right)^{-1}$ was often quite large, the field range was still small. Writing $\mathcal{Q}$ as

$$\mathcal{Q} = 2\pi\mathbb{1} + \Delta_{\mathcal{Q}}\,, \tag{3.29}$$

we expect (cf. the analysis in [46, 113]) that if the entries of $\Delta_{\mathcal{Q}}$ are well-approximated as i.i.d. stochastic variables, and if these entries are not too sparse,[26] then $\mathcal{R}$ should manifest a large degree of enhancement from alignment. In the geometries we examined, Q is highly structured, and contains an identity matrix of size at least $h^{1,1} - 1$; the remainder $\Delta_{\mathcal{Q}}$ is then extremely sparse.[27] We found that

[26] Concretely, if e.g. 5% of the entries of a 100×100 matrix $\Delta_{\mathcal{Q}}$ are nonzero, the random matrix analysis yields a heavy tail toward large $\mathcal{R}$.

[27] Between 1% to 7% of the entries in the large $h^{1,1}$ ensemble are populated, but the nonzero off-diagonal entries are restricted to a few rows and columns.

$\left(\mathcal{Q}\mathcal{Q}^\top\right)^{-1}$ can in fact have a very large eigenvalue, but this is only necessary, not sufficient, for a large enhancement of the field range. Indeed, we should interpret q_N as the maximum possible enhancement from lattice alignment. The largest enhancement occurs when the largest-eigenvalue eigenvectors of the Kähler metric K_{ij} and of $\left(\mathcal{Q}\mathcal{Q}^\top\right)^{-1}$ are parallel, such that $\xi_N = f_N q_N$, where f_N^2 is the largest eigenvalue of K_{ij}. If these eigenvectors are misaligned, the enhancement occurs in a different direction in field space—one that is irrelevant to the semi-diameter $\mathcal{R}$—and can compress the polytope, ultimately diminishing the field range.

Let us briefly discuss why Q so consistently contains a large identity matrix. First consider a Calabi-Yau with large $h^{1,1}$ and small $h^{2,1}$. Here the large number of rigid divisors can be understood as a consequence of mirror symmetry. If $h^{2,1}$ is small, then the number of points in the dual polytope Δ is small. Recall that the Hodge numbers of the toric divisors are computed by counting lattice points interior to faces of Δ, so as Δ gets smaller the number of points interior to faces decreases, so more of the toric divisors have a better chance of becoming rigid.[28] For instance, consider the hypersurfaces in the Kreuzer-Skarke database with $h^{1,1} = 404$ and $h^{2,1} = 14$. There are six lattice polytopes corresponding to these Hodge numbers, and in all six at least 402 of the toric divisors are rigid.

On the other hand, this argument does not apply when both $h^{1,1}$ and $h^{2,1}$ are large. For example, we can consider a hypersurface with $h^{1,1} = h^{2,1} = 100$, whose corresponding Δ° polytope has vertices

$$\begin{aligned}\mathbf{d}_i = \Big\{&(1,-1,-1,-1),\ (-1,-1,-1,-1),\ (-1,-1,6,-1),\ (-1,2,-1,-1),\\ &(-1,-1,6,2),\ (-1,-1,4,5),\ (-1,2,-1,0),\ (-1,-1,-1,11),\ (-1,-1,1,9)\Big\}.\end{aligned} \tag{3.30}$$

Here 98 of the 104 toric divisors have $h^\bullet(D,\mathcal{O}_D) = (1,0,0)$, even though the dual

[28] We thank Andreas Braun for inspiration on this point.

polytope has 134 points, only 8 of which are vertices. Therefore most of the dual cones have no interior points, and the non-vertex points are interior to only a few cones. This seems to be a consequence of the shape of Δ, and is likely related to the requirement that the origin is the only interior point of Δ: as the number of points included in the polytope grows, the shape must be more and more skewed. In summary, we find it reasonable to conjecture that in many geometries with large $h^{1,1}$, Q will have a large-dimensional identity block, which does not contribute to lattice alignment.

3.6 Conclusions

In this work we have initiated a systematic analysis of axion field ranges in type IIB compactifications on Calabi-Yau hypersurfaces in toric varieties. For axions descending from the Ramond-Ramond four-form C_4 in the 4390 geometries that we considered, we found a maximum field range of $\mathcal{R}_{\rm max} = 0.3 M_{\rm pl}$. The largest enhancement of $\mathcal{R}$ due to lattice (KNP) alignment in our ensemble was a factor 2.6, in an example with $\mathcal{R} \ll M_{\rm pl}$. The numerical value of $\mathcal{R}_{\rm max}$ should not be overinterpreted, because it can be made smaller or larger by imposing a more or less stringent requirement for control of the α' expansion; the quoted value results from the requirement that the smallest curve has volume $(2\pi)^2\alpha'$. What is clear is that in our examples, with our assumptions, the geometric field range does not parametrically exceed the Planck mass.

To assess the implications of these results, let us reexamine our assumptions and ask which of them might be relaxed. First of all, it is plausible that in some geometries, one or more curves could be taken to have volume $t_i < 1$, while keeping

other volumes large, without invalidating the sigma model expansion. In this work we have followed a conservative, model-independent approach, but a more complete understanding of perturbative and nonperturbative corrections could allow for much larger field ranges.

Second, we considered axion potentials generated by Euclidean D3-branes wrapping divisors D fulfilling the rigidity condition $h^\bullet(D, \mathcal{O}_D) = (1, 0, 0)$. That is, we required that D be a rigid divisor of a smooth threefold, and did not incorporate the effects of orientifolding, worldvolume fluxes, bulk fluxes, and spacetime-filling seven-branes, which could alter the set of instanton contributions to the superpotential. In particular, strong gauge dynamics on seven-branes, such as gaugino condensation on a stack of D7-branes coinciding with an O7-plane, provides a plausible mechanism for allowing larger field ranges, and more significant alignment, than we found in this work. The axion periodicity induced by such branes is increased by a factor of the dual Coxeter number $c_2(G)$ of the condensing gauge group G, and many proposals for lattice alignment in string theory invoke stacks of D7-branes with $c_2 > 1$. Systematically investigating such constructions would be valuable.

Third, we only examined C_4 axions in compactifications of type IIB string theory on Calabi-Yau hypersurfaces X in toric varieties V, and we insisted that X be favorable, meaning that all divisors of X are inherited from V. Each of these restrictions merits further investigation. Two-form axions have a distinct parametric dependence on Kähler moduli, possibly allowing larger field ranges while maintaining control of the α' expansion [118, 117]. We have no evidence to guide speculation about axion field ranges in threefolds that are not favorable hypersurfaces.

Finally, our systematic investigation occurred at small Hodge numbers, $h^{1,1} \leq 4$, and we studied only a handful of examples with $h^{1,1}$ up to 100. An analysis based on random matrix models, with parameters calibrated by the examples found here, suggests that the maximum field range at moderate $h^{1,1}$ could be large. Whether this can occur in actual compactifications depends on a competition between a tendency for the overall volume $\mathcal{V}$ to grow with $h^{1,1}$, which suppresses the entries of the Kähler metric, and the fact that larger axion charge matrices $\mathcal{Q}$ can manifest a greater degree of lattice alignment. We observed a tendency for $\mathcal{Q}$ to be close to the identity in cases with $h^{1,1} \gg 1$, which precludes large enhancements from alignment, due to the prevalence of rigid toric divisors in these examples.

In summary, in compactifications of type IIB string theory on Calabi-Yau hypersurfaces with $h^{1,1} \leq 4$, Euclidean D3-branes wrapping divisors D that do not intersect seven-branes give rise to a potential for C_4 axions that allows for a small degree of lattice alignment, which is insufficient to allow a super-Planckian geometric field range, in the absence of monodromy, in a parameter regime where all curves have volume $\geq (2\pi)^2\alpha'$. Understanding the geometry of axion field space in far more general compactifications is an important problem for the future.

CHAPTER 4

ON CHIRAL MESONS IN ADS/CFT

Abstract[1]

We analyze the spectra of non-chiral and chiral bifundamental mesons arising on intersecting D7-branes in $AdS_5 \times S^5$. In the absence of magnetic flux on the curve of intersection, the spectrum is non-chiral, and the dual gauge theory is conformal in the quenched/probe approximation. For this case we calculate the dimensions of the bifundamental mesonic operators. We then consider magnetization of the D7-branes, which deforms the dual theory by an irrelevant operator and renders the mesons chiral. The magnetic flux spoils the conformality of the dual theory, and induces a D3-brane charge that becomes large in the ultraviolet, where the non-normalizable bifundamental modes are rapidly divergent. An ultraviolet completion is therefore necessary to calculate the correlation functions in the chiral case. On the other hand, the normalizable modes are very well localized in the infrared, leading to new possibilities for local model-building on intersecting D7-branes in warped geometries.

[1]This chapter is based on L. McAllister, P. McGuirk and J. Stout, "On Chiral Mesons in AdS/CFT," *JHEP* **02** (2014) 018, [1311.2577].

It is a pleasure to thank F. Marchesano and G. Shiu for useful discussions of related topics. This work was supported by the NSF under grant PHY-0757868.

The AdS/CFT correspondence [120, 121, 122, 123] is a powerful duality relating conformal field theories (CFTs) in d dimensions to gravitational theories on $(d+1)$-dimensional anti-de Sitter (AdS) spaces. The extension of the duality to include global flavor groups has been well-studied (see [124, 125] for some foundational work) and is well-motivated: it brings the theory closer to phenomenologically viable models, with mesonic bound states serving as prototypes for visible-sector fields. However, to find more realistic models, the flavor group must be extended to a product group, and the resulting mesonic spectrum must be made chiral. Such extensions have been relatively unexplored, and in the present work we report on progress in this direction.

When the gravity side of the duality is a type II string theory, flavor groups are added through the introduction of higher-dimensional Dp-branes that fill AdS and wrap compact cycles [125].[2] The simplest such example is the addition of F D7-branes to type IIB string theory on $AdS_5 \times S^5$, where we take the D7-branes to fill AdS_5 and wrap an S^3 of the S^5. The geometry is supported by N units of D3-brane charge and, without the D7-branes, is dual to $\mathcal{N} = 4$ SU(N) super Yang-Mills. Adding the D7-branes deforms the dual theory to an $\mathcal{N} = 2$ gauge theory with a U(F) flavor group, containing a massless adjoint hypermultiplet as well as a massless quark hypermultiplet that transforms in the bifundamental of SU(N) $\times$ U(F). The brane construction makes this clear, as the open string excitations of the D7-branes give rise to a U(F) gauge theory, and the infinite D7-brane worldvolume transverse to the D3-branes results in a vanishing 4d coupling for this theory. Open strings stretching between the D7-branes and the D3-branes have the same charges as the quarks in the dual theory. We will work in the

[2] The higher-dimensional D-brane need not fill all of AdS; if the brane is characterized by a minimum distance away from the origin of AdS then the dual quarks are massive [126].

standard decoupling limit [120] in which one first takes

$$g_s \to 0, \qquad N \to \infty, \qquad \lambda_t \equiv 4\pi g_s N \quad \text{fixed}, \tag{4.1}$$

and then sends the 't Hooft coupling λ_t to infinity. In this limit, the D3-branes are replaced by their near-horizon backreaction, so that the only open strings are those stretching among the D7-branes. These transform in the adjoint representation of U(F) and are dual to mesonic operators in the gauge theory.

D7-branes are codimension-two objects, and so their backreaction cannot generally be neglected. Correspondingly, the presence of quarks in the dual gauge theory alters the renormalization group flow, which was trivial before the introduction of flavor. Fortunately, the decoupling limit (4.1) simplifies the situation: if we hold fixed the number of flavors, F, while taking the number of colors to be large, then one can consistently neglect the running of quarks in loops. In the dual geometry, many aspects of the D7-brane backreaction scale as F/N and so also vanish in this limit (see [127] and references therein). The flavored gauge theory does have a Landau pole, and so the influence of the quarks on the renormalization group flow cannot be neglected forever, but the scale at which the Landau pole appears grows exponentially with N/F. This so-called quenched approximation, in which the running of quarks in loops is neglected, is equivalent to the limit in which the D7-branes are taken as probes of the dual geometry. In what follows, we will take this approximation without further apology.

The introduction of flavor branes opens up significant possibilities for model-building. Dimensional reduction along the angular directions provides a framework for Randall-Sundrum constructions [128, 129, 130, 131] wherein the Standard Model fields propagating in the bulk [132, 133, 134, 135, 136] descend from the D7-brane fluctuations as in [137]. Upon compactification, the flavor group

on the D7-branes becomes a prototype for the Standard Model gauge group. Of course, the Standard Model gauge group is a product; a corresponding product flavor group results from introducing two separate stacks of D7-branes. The bifundamental fields are then open strings stretching between the stacks, and in order for some of the bifundamentals to be massless, the stacks must intersect.

A further challenge is that the Standard Model spectrum is chiral. In the class of constructions considered here, chirality in the 4d theory can be induced by introducing magnetic flux on the (noncompact) curve where the D7-branes intersect. Upon compactification to 4d, the zero modes of the Dirac operator acquire a net chirality set by the amount of quantized magnetic flux.

Yet another difficulty in embedding fully realistic theories into warped backgrounds of string theory is the fact that the Standard Model is not a supersymmetric theory. In geometries that are characterized by a finite infrared scale, such as the well-studied Klebanov-Strassler solution [138], supersymmetry can be broken in a controllable way by the addition of a small number of anti-D3-branes [139]. The resulting geometry [140, 141, 142, 143, 144, 145, 146, 147, 148] corresponds to the spontaneous breaking of supersymmetry in the dual field theory [40].[3] An alternative is to consider "gluing" the warped geometry to a compact space that does not preserve supersymmetry. The dual field theory is then a non-supersymmetric theory with emergent supersymmetry, as in [149, 150]. Although non-supersymmetric constructions are difficult to control, the filtering provided by the renormalization group means that the influence of the non-supersymmetric bulk, including the effects of moduli stabilization, can be systematically parameterized and incorporated along the lines of [151, 64]. No matter which supersymmetry-breaking mechanism

[3]Some authors have interpreted the singularities of the anti-D3-brane geometry described in [140] as implying that the supersymmetry-breaking state does not exist.

is used,[4] the resulting geometry is considerably more complex after supersymmetry is broken. We will therefore, in this initial work, focus on supersymmetric D7-brane probes of supersymmetric backgrounds.[5]

In this note, we will consider the non-chiral and chiral bifundamental modes existing at the intersections of probe D7-branes in $AdS_5 \times S^5$. We build up to the chiral, warped case through the simpler example of intersecting D7-branes in flat space (§4.1). Although the flat-space analysis of §4.1 has appeared elsewhere in the literature (see e.g. [163, 164, 165, 166]), a detailed treatment is useful here, because the equations of motion are readily generalized from the simple flat-space case to the $AdS_5 \times S^5$ configuration of primary interest.

The organization of this note is as follows. In §4.1 we begin with the simple case of intersecting D7-branes in a flat space background. In §4.1.2 we compute the mass spectrum of the bifundamental modes for the case of vanishing magnetic flux, where the spectrum is non-chiral. Then, in §4.1.3 we calculate the chiral mass spectrum in a configuration with magnetic flux. Next, in §4.2 we consider unmagnetized intersecting D7-branes in $AdS_5 \times S^5$, computing the scaling dimensions of vector-like bifundamental mesonic operators. Finally, in §4.3 we add the simplest possible magnetization to the intersecting D7-branes in $AdS_5 \times S^5$, and show that this magnetization makes the calculation of correlation functions untrustworthy without an ultraviolet completion. Concluding remarks are given in §4.4, while our conventions and a few technical details appear in the appendices.

[4] See [152, 153] for other interesting proposals.

[5] See, for example, [154, 155, 156, 157, 158, 159, 160, 161, 162] for analyses of probe D7-branes in non-supersymmetric backgrounds from the worldvolume and/or worldsheet points of view.

4.1 D7-branes in Flat Space

As a warm-up to the case of strong warping, we will first review the case of intersecting D7-branes probing unwarped flat space, $\mathbb{R}^{9,1} = \mathbb{R}^{3,1} \times \mathbb{C}^3$.

4.1.1 The D7-brane action

As discussed in the introduction, we focus on supersymmetric configurations, and so we take a flat D7-brane probe, which preserves half of the supercharges of flat space. By a choice of orientation and complex structure, the D7-brane worldvolume $\mathcal{W}$ can be taken to be $\mathbb{R}^{3,1} \times \mathbb{C}^2$. The light bosonic degrees of freedom resulting from the open-string excitations of the D7-brane consist of the transverse deformations Φ^i and a U(1) vector potential A_1. We use this potential to construct a Lorentz-invariant[6] and supersymmetric magnetic flux $F_2 = \mathrm{d}A_1$; such a flux satisfies the self-duality condition [167, 168, 169]

$$F_2 = \tilde{*}_4 F_2, \tag{4.2}$$

where $\tilde{*}_4$ is the Hodge star built from the metric on $\mathbb{C}^2$. The condition (4.2) is equivalent to F_2 being $(1,1)$ and primitive with respect to the Kähler form induced on $\mathbb{C}^2$.

To leading order in the α' expansion, the action of the D7-brane in this background is [170, 171, 172]

$$\begin{aligned} S_{\mathrm{D7}} = -\frac{1}{g_8^2}\int_{\mathcal{W}} \mathrm{d}^8\xi^\alpha \sqrt{-\hat{g}}\Big\{ &\frac{1}{2}\hat{g}_{ij}\hat{g}^{\alpha\beta}\partial_\alpha\Phi^i\partial_\beta\Phi^j \\ &+ \frac{1}{4}\hat{g}^{\alpha\beta}\hat{g}^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} + \mathrm{i}\bar{\Theta}P_-^{\mathrm{D7}}\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha\partial_\beta\Theta\Big\}, \end{aligned} \tag{4.3}$$

[6] Here and throughout we will use "Lorentz invariance" to refer to SO(3, 1) invariance.

in which we have omitted a constant term that does not play a role in our analysis. Writing the string tension as $\tau_{\mathrm{F1}}^{-1} = 2\pi\alpha' = \ell_{\mathrm{s}}^2$, the 8d Yang-Mills coupling is $g_8^{-2} = 8\pi^3\ell_{\mathrm{s}}^4 g_{\mathrm{s}}$. Here ξ^α are coordinates on the D7-brane, $\hat{g}_{\alpha\beta}$ is the induced worldvolume metric and $\hat{g}_{ij}$ is the transverse metric. Θ is a 10d double Majorana-Weyl spinor (reviewed in Appendix 4.A) that, as in the Green-Schwarz superstring, redundantly encapsulates the fermionic degrees of freedom of the D7-brane. In particular, Θ is subject to the κ-symmetry identification

$$\Theta \sim \Theta + P_-^{\mathrm{D}p}\kappa, \tag{4.4}$$

in which κ is an arbitrary Majorana-Weyl double spinor. P_-^{D7} is given by

$$P_-^{\mathrm{D7}} = \frac{1}{2}\begin{pmatrix} 1 & -\Gamma_{\mathrm{D7}}^{-1} \\ -\Gamma_{\mathrm{D7}} & 1 \end{pmatrix}, \tag{4.5}$$

in which

$$\Gamma_{\mathrm{D7}} = \mathrm{dv\not{o}l}_{\mathcal{W}} := \frac{1}{8!}\hat{\epsilon}_{\alpha_1\cdots\alpha_8}\hat{\Gamma}^{\alpha_1\cdots\alpha_8} = -\mathrm{i}\Gamma_{(8)}, \tag{4.6}$$

where $\epsilon_{\alpha_1\cdots\alpha_8}$ is the antisymmetric tensor and $\Gamma_{(8)}$ is the $\mathrm{SO}(7,1)$ chirality operator. We use κ-symmetry to set

$$\Theta = \begin{pmatrix} \theta \\ 0 \end{pmatrix}. \tag{4.7}$$

With this choice,

$$S_{\mathrm{D7}} = -\frac{1}{g_8^2}\int_{\mathcal{W}} \mathrm{d}^8\xi^\alpha \left\{ \frac{1}{2}\hat{g}_{ij}\hat{g}^{\alpha\beta}\partial_\alpha\Phi^i\partial_\beta\Phi^j + \frac{1}{4}\hat{g}^{\alpha\beta}\hat{g}^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} + \frac{\mathrm{i}}{2}\bar{\theta}\hat{g}^{\alpha\beta}\hat{\Gamma}_\beta\partial_\alpha\theta \right\}, \tag{4.8}$$

which is the familiar action for maximally supersymmetric 8d U(1) gauge theory.

On a stack of F such D7-branes, the gauge group is enhanced to $\mathrm{U}(F)$, and A_α, Φ^i, and θ are promoted to adjoint-valued fields. The leading-order action is determined by gauge-invariance and supersymmetry to be

$$\begin{aligned} S_{\mathrm{D7}} = -\frac{1}{g_8^2}\int_{\mathcal{W}} \mathrm{d}^8\xi^\alpha\, \mathrm{tr}\Big\{ &\frac{1}{2}\hat{g}_{ij}\hat{g}^{\alpha\beta}D_\alpha\Phi^i D_\beta\Phi^j + \frac{1}{4}\hat{g}^{\alpha\beta}\hat{g}^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} \\ &- \frac{1}{4}\hat{g}_{ij}\hat{g}_{kl}\big[\Phi^i,\Phi^k\big]\big[\Phi^j,\Phi^l\big] + \frac{\mathrm{i}}{2}\bar{\theta}\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha D_\beta\theta - \frac{1}{2}\bar{\theta}\,\hat{\Gamma}_i\big[\Phi^i,\theta\big]\Big\}, \end{aligned} \tag{4.9}$$

in which tr denotes a trace over gauge indices, D_α is a gauge covariant derivative

$$D_\alpha = \partial_\alpha - \mathrm{i}\big[A_\alpha, \cdot\big], \tag{4.10}$$

and $F_2 = \mathrm{d}A - \mathrm{i}A \wedge A$ is the non-Abelian field strength.

Bifundamental modes arise from strings that stretch between stacks of Dp-branes. If the stacks are parallel, then the mass of these modes is proportional to the separation between the branes. Such a configuration still preserves sixteen supercharges and so the action for the bifundamental modes (which provide a full massive vector multiplet) can be fixed by symmetries. Alternatively, the action can be found by Higgsing the theory (4.9). Beginning with a stack of $F_1 + F_2$ D7-branes, the transverse deformations can be treated as $(F_1 + F_2) \times (F_1 + F_2)$ matrices with the ith diagonal element corresponding to a transverse deformation of the ith brane. A vacuum expectation value (vev) with the gauge structure

$$\big\langle \Phi^i \big\rangle = \ell_\mathrm{s}^{-2} \begin{pmatrix} X_1^i \, \mathbb{1}_{F_1} & \\ & X_2^i \, \mathbb{1}_{F_2} \end{pmatrix} \tag{4.11}$$

breaks $\mathrm{U}(F_1 + F_2) \to \mathrm{U}(F_1) \times \mathrm{U}(F_2)$ and describes a separation of the branes $\Delta x^i = |X_1^i - X_2^i|$. The factor of ℓ_s^2 is introduced so that Φ^i has length dimension -1. However, this also has the effect of canceling the factors of ℓ_s that appear in operators correcting the Yang-Mills action. Therefore, in order to trust this effective field theory, we consider cases where $\Delta x^i \ll \ell_\mathrm{s}$. Equivalently, if we are to trust the effective field theory description of the modes stretching between the branes, their mass must be less than that of the massive string states that have been integrated out implicitly.

Writing the fluctuations as

$$\delta\Phi^i = \begin{pmatrix} \phi_1^i & \phi_+^i \\ \phi_-^i & \phi_2^i \end{pmatrix}, \tag{4.12}$$

ϕ_1^i and ϕ_2^i transform as adjoints under $\mathrm{U}(F_1)$ and $\mathrm{U}(F_2)$, respectively, while ϕ_+^i and ϕ_-^i are bifundamentals that acquire masses proportional to the separation. For notational simplicity, in what follows we will consider the case $F_1 = F_2 = 1$, but all of our results generalize easily to higher ranks.

If, instead of being parallel, the branes intersect, some of the bifundamental modes will become massless. The intersection of two D7-branes is generically six-dimensional, and the long-wavelength description of the bifundamental modes can be given in terms of a 6d effective field theory description on this intersection. The 6d masses of the bifundamentals depend on the angles formed by the intersection of the branes. However, the vector bifundamentals never become massless, indicating that the 6d theory is a $\mathrm{U}(1) \times \mathrm{U}(1)$ (rather than the un-Higgsed $\mathrm{U}(2)$) gauge theory, and that fewer than sixteen supercharges are preserved, since the vector multiplet is split. When the intersection is such that both D7-branes fill $\mathbb{R}^{3,1}$ and are holomorphically embedded into $\mathbb{C}^3$, at least minimal supersymmetry is preserved [173] and the 6d theory includes massless scalars and fermions.

4.1.2 Non-chiral modes

In the warped case, the calculation of mass spectra is equivalent to the calculation of scaling dimensions in the dual theory. In this section, we continue our warm-up to the warped case by finding the mass spectrum of non-chiral bifundamental modes in flat space. To this end, we take $z^{I=1,2,3}$ as coordinates on $\mathbb{C}^3$ and consider a pair of D7-branes whose embeddings are specified by

$$\mathrm{D7}_1 : z^3 = tz^2, \quad \mathrm{D7}_2 : z^3 = -tz^2, \qquad t > 0. \tag{4.13}$$

Following the discussion in the previous subsection, we can describe this intersection by considering 8d U(2) SYM along $\mathbb{R}^{3,1} \times \mathbb{C}^2$ (with $\mathbb{C}^2$ spanned by z^1 and z^2), where the vev for the complexified transverse deformation takes the form[7]

$$\Phi = q \begin{pmatrix} z^2 & \\ & -z^2 \end{pmatrix}, \tag{4.14}$$

in which $q = \ell_{\rm s}^{-2} t$. The bifundamental modes are localized on $\mathbb{R}^{3,1} \times \mathbb{C}$, with z^1 the coordinate on the curve of intersection (which in this case is simply $\mathbb{C}$). For the reasons discussed above, we must take $t \ll 1$ in order to trust the effective field theory. Of course, no matter what the value of t, at sufficiently large values of z^2 the branes will be far apart and so one might worry about stringy corrections to the Yang-Mills action. That is, in addition to (4.9), the worldvolume action contains, for example, operators with the schematic form

$$\ell_{\rm s}^{k-4} \left(\Phi\right)^k \sim \left(\frac{t z^2}{\ell_{\rm s}}\right)^{k-4} \varphi_{\pm}^4 + \cdots, \tag{4.15}$$

which might seem to become important at $z^2 \sim t^{-1} \ell_{\rm s}$. However, as we will show below, the bifundamental modes are highly peaked at $z^2 = 0$, and so we anticipate that their physics will be largely insensitive to the corrections at large z^2.

The configuration just described is supersymmetric, so we can find solutions to the bosonic equations of motion by solving the fermionic equations of motion. Although the intersection is $\mathrm{SO}\left(5,1\right)$ symmetric, in anticipation of the magnetization — which preserves only $\mathrm{SO}(3,1)$ and which we discuss below — we will make use of the decomposition $\mathrm{SO}\left(9,1\right) \to \mathrm{SO}(3,1) \times \mathrm{SO}\left(6\right)$, as discussed in Appendix 4.A. It is useful to decompose the 10d fermionic mode θ into modes of different internal chirality (i.e. $\mathrm{SO}\left(6\right)$ weights)

$$\theta = \sum_{m=0}^{3} \left\{ \psi_m \begin{pmatrix} \xi \\ 0 \end{pmatrix} \otimes \eta_m - \psi_m^\dagger \begin{pmatrix} 0 \\ \sigma^2 \xi^* \end{pmatrix} \otimes \tilde{\beta}_6 \eta_m^* \right\}, \tag{4.16}$$

[7] Similar vevs were utilized in [163, 164] to describe brane recombination from non-supersymmetric intersections.

in which ξ is a fixed two-component spinor, η_m are the constant SO (6) positive chirality spinors in (4.112), and $\tilde{\beta}_6$ is the SO (6) Majorana matrix. Writing the U(1) potential as $A_1 = A_\mu \mathrm{d}x^\mu + \sum_{a=1}^{2} \left(a_a \mathrm{d}z^a + a_{\bar{a}} \mathrm{d}\bar{z}^a\right)$, ψ_0 is the fermionic partner of A_μ, $\psi_{1,2}$ are the partners of a_1, and a_2, and ψ_3 is the partner of the complexified transverse deformation Φ. Each of the ψ_m transforms under the adjoint representation of U(2) and we write (cf. (4.12))

$$\psi_m = \begin{pmatrix} & \psi_m^+ \\ \psi_m^- & \end{pmatrix}, \tag{4.17}$$

in which we have set the neutral fields $\psi_m^{1,2}$ to zero since they are not the modes of interest.

The linearized equation of motion for the fermions in this background is

$$0 = \hat{\Gamma}^\alpha \partial_\alpha \theta - \mathrm{i}\, \hat{\Gamma}_i \big[\Phi^i, \theta\big], \tag{4.18}$$

where the transverse fluctuations Φ^i are evaluated on their vev (4.14). From SO$(3,1)$ invariance, we expect that the equation of motion for A_μ should decouple from those of the other bosonic fields, at least for some gauge choice,[8] and thus we can consistently take $\psi_0^\pm$, the superpartner of $A_\mu^\pm$, to vanish. When the 4d momentum is zero we have

$$0 = \bar{\partial}_{\bar{1}} \psi_1^\pm - \bar{\partial}_{\bar{2}} \psi_2^\pm \mp \mathrm{i} q z^2 \psi_3^\pm, \tag{4.19a}$$

$$0 = \partial_2 \psi_3^\pm \mp \mathrm{i} q \bar{z}^{\bar{2}} \psi_2^\pm, \tag{4.19b}$$

$$0 = \partial_1 \psi_3^\pm \pm \mathrm{i} q \bar{z}^{\bar{2}} \psi_1^\pm, \tag{4.19c}$$

$$0 = \partial_1 \psi_2^\pm + \partial_2 \psi_1^\pm. \tag{4.19d}$$

The equations (4.19) also follow from the conditions for supersymmetry [165, 166, 174]. These coupled first-order equations can be turned into largely decou-

[8] One such gauge choice is (4.19a) after simply replacing the fermionic fields with their bosonic partners. See, for example, [166].

pled second-order equations by taking derivatives. For example, application of ∂_1 to (4.19a) and substitution of (4.19c) and (4.19d) yields

$$0 = \partial_1\bar{\partial}_{\bar{1}}\psi_1^{\pm} + \partial_2\bar{\partial}_{\bar{2}}\psi_1^{\pm} - q^2\left|z^2\right|\psi_1^{\pm}. \tag{4.20a}$$

Similarly,

$$0 = \partial_1\bar{\partial}_{\bar{1}}\psi_2^{\pm} + \partial_2\bar{\partial}_{\bar{2}}\psi_2^{\pm} \pm \mathrm{i}q\psi_3^{\pm} - q^2\left|z^2\right|\psi_2^{\pm}, \tag{4.20b}$$

$$0 = \partial_1\bar{\partial}_{\bar{1}}\psi_3^{\pm} + \partial_2\bar{\partial}_{\bar{2}}\psi_3^{\pm} \mp \mathrm{i}q\psi_2^{\pm} - q^2\left|z^2\right|\psi_3^{\pm}. \tag{4.20c}$$

Using (4.19b), (4.20c) gives an equation for $\psi_3^{\pm}$ alone. Writing $\psi_3^{\pm} = \bar{z}^{\bar{2}}\psi_{\pm}$, we have

$$0 = \partial_1\bar{\partial}_{\bar{1}}\psi_{\pm} + \partial_2\bar{\partial}_{\bar{2}}\psi_{\pm} - q^2\left|z^2\right|\psi_{\pm}. \tag{4.21}$$

Once $\psi_{\pm}$ is determined, $\psi_{1,2,3}^{\pm}$ are easily found.

Equation (4.21) is separable. Performing polar decompositions $z^a = r_a\mathrm{e}^{\mathrm{i}\phi_a}$ and taking the ansatz

$$\psi_{\pm} = \mathrm{e}^{\mathrm{i}(m_1\phi_1 + m_2\phi_2)}\zeta_{\pm}\big(r_1\big)\sigma_{\pm}\big(r_2\big), \tag{4.22}$$

where m_i are integers, we have

$$0 = \zeta_{\pm}'' + \frac{1}{r_1}\zeta_{\pm}' - \frac{m_1^2}{r_1^2}\zeta_{\pm} - 4\lambda\zeta_{\pm}, \tag{4.23a}$$

$$0 = \sigma_{\pm}'' + \frac{1}{r_2}\sigma_{\pm}' - \frac{m_2^2}{r_2^2}\sigma_{\pm} - 4q^2r_2^2\sigma_{\pm} + 4\lambda\sigma_{\pm}, \tag{4.23b}$$

in which λ is a constant to be determined by boundary conditions. Imposing that $\sigma_{\pm} \to 0$ as $r_2 \to 0$ we find

$$\zeta_{\pm}\big(r_1\big) = c_1 I_{|m_1|}\big(\sqrt{2\lambda}\, r_1\big) + c_2 K_{|m_1|}\big(\sqrt{2\lambda}\, r_1\big), \tag{4.24a}$$

$$\sigma_{\pm}\big(r_2\big) = \mathrm{e}^{-qr_2^2}\big(2qr_2^2\big)^{|m_2|/2} L_n^{|m_2|}\big(2qr_2^2\big). \tag{4.24b}$$

in which L_ν^μ are the associated Laguerre polynomials, I_μ and K_μ are the modified Bessel functions of the first and second kinds, and

$$\lambda = q\big(2n + |m_2| + 1\big). \tag{4.25}$$

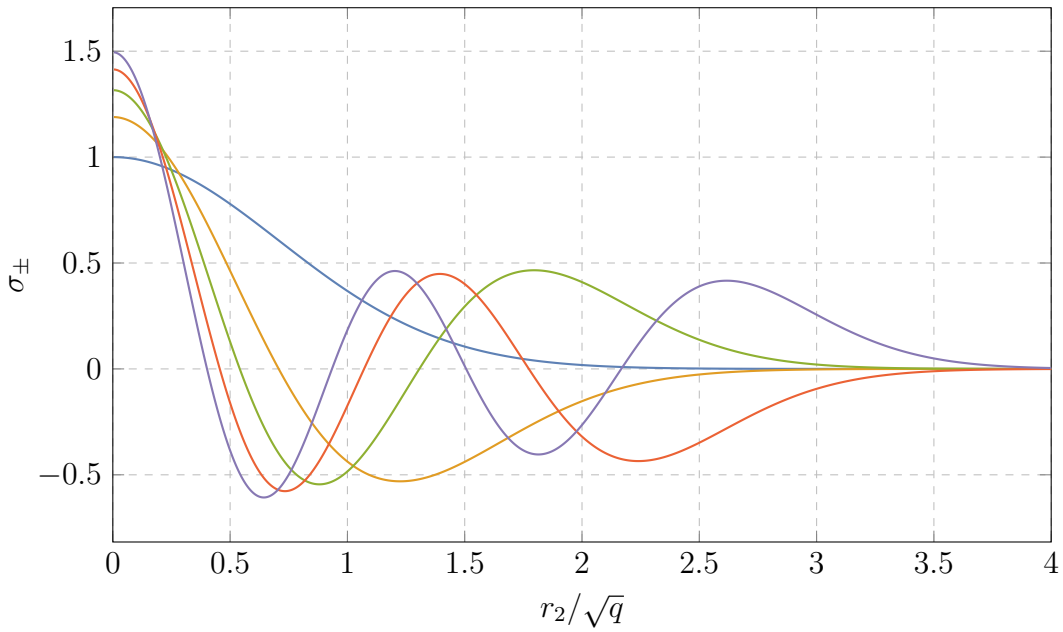

Figure 4.1: Transverse profiles for the flat space vector-like bifundamental modes $\sigma_\pm$ given by (4.24) for $m_2 = 0$ and $n = 0$ (the curve with smallest value at $r_2 = 0$) through $n = 4$ (the curve with the largest value at $r_2 = 0$). The solutions have been normalized to the same value using the inner product $\int \mathrm{d}r^2 f\left(r^2\right) g\left(r^2\right)$.

Regularity of $\sigma_\pm$ requires that n is a non-negative integer. Some of these modes are plotted in Figures 4.1 and 4.2.

One may notice that the system (4.19) also admits a zero mode that depends only on r_2,

$$\psi_3^\pm = \mathrm{e}^{-qr_2^2}. \tag{4.26}$$

It is easy to confirm that this gives a solution to (4.21), but this solution is not normalizable with respect to the norm defined by treating (4.21) as a Sturm-Liouville problem. This is a consequence of the fact that the bifundamental modes are more properly encoded by linear combinations of the $\psi_m^\pm$ rather than by the $\psi_m^\pm$ themselves [163, 164]. Correspondingly, the measure used in integrating over the z^2 and $\bar{z}^{\bar{2}}$ directions is not that defined by (4.21) (see [166]). However, aside from this zero mode, the above equations successfully reproduce the spectrum of

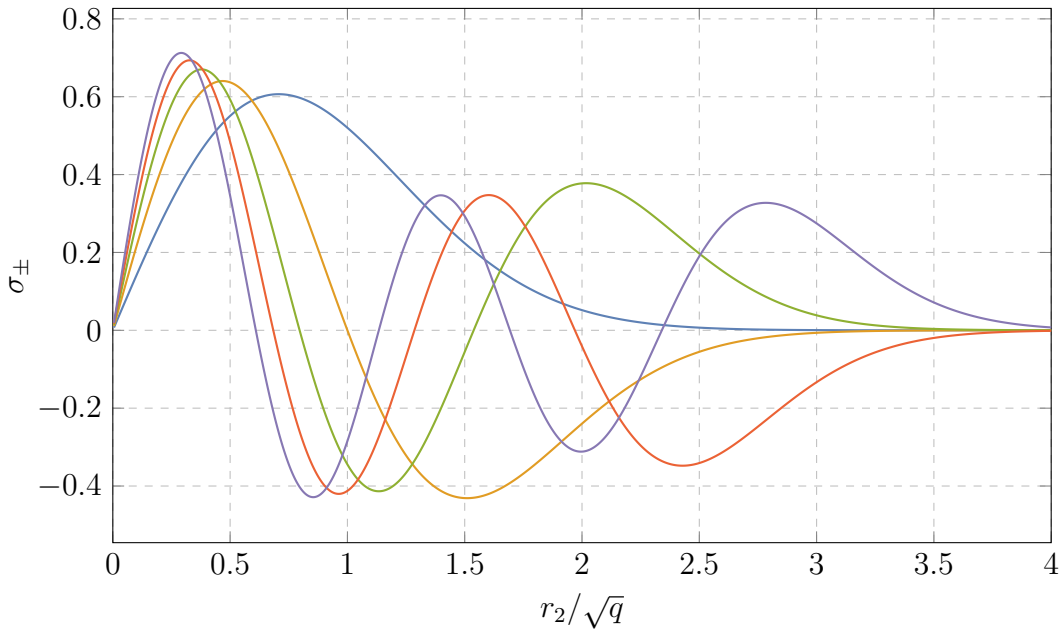

Figure 4.2: Similar plot as Figure 4.1 except with $m_2 = 1$.

6d masses (4.25).

4.1.3 Chiral modes

We now consider magnetized intersections since, upon compactification, such a construction gives a chiral 4d theory. We will again focus on supersymmetric configurations, which implies that F_2 must be $(1,1)$ and primitive. Consider first a single D7-brane on $\mathbb{R}^{3,1} \times \mathbb{C}^2$. The most general $(1,1)$ flux that can be supported by the D7-brane is

$$F_2 = -\frac{\mathrm{i}}{2} f_1 \, \mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^{\bar{1}} - \frac{\mathrm{i}}{2} f_2 \, \mathrm{d}z^2 \wedge \mathrm{d}\bar{z}^{\bar{2}} - \frac{\mathrm{i}}{2} g_1 \, \mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^{\bar{2}} - \frac{\mathrm{i}}{2} g_2 \, \mathrm{d}\bar{z}^{\bar{1}} \wedge \mathrm{d}z^2. \quad (4.27)$$

The $1\bar{1}$ component will describe the magnetization of the intersection, and so we will look for the simplest configurations with $f_1 \neq 0$. The Kähler form on $\mathbb{C}^2$ is

simply

$$J = -\frac{\mathrm{i}}{2}\sum_{I=1}^{2} \mathrm{d}z^I \wedge \mathrm{d}\bar{z}^{\bar{I}}, \tag{4.28}$$

and so primitivity imposes $f_1 = -f_2$. The Bianchi identity implies that f_1 is harmonic,

$$0 = \partial_1\bar{\partial}_1 f_1 + \partial_2\bar{\partial}_2 f_1. \tag{4.29}$$

In the absence of sources, (4.29) requires that f_1 is constant. We can then consistently set $g_1 = g_2 = 0$ and obtain the supersymmetric magnetization

$$F_2 = -\mathrm{i}\, M \left\{\mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^{\bar{1}} - \mathrm{d}z^2 \wedge \mathrm{d}\bar{z}^{\bar{2}}\right\}. \tag{4.30}$$

Compactification would impose a quantization condition on M, but in the non-compact case we can freely take M to be any constant. The above magnetization follows from the gauge choice

$$A_1 = -\frac{\mathrm{i}}{2} M \left\{z^1 \mathrm{d}\bar{z}^{\bar{1}} - \bar{z}^{\bar{1}}\mathrm{d}z^1 - z^2\mathrm{d}\bar{z}^{\bar{2}} + \bar{z}^{\bar{2}}\mathrm{d}z^2\right\}. \tag{4.31}$$

To obtain chiral matter, we again consider the intersection of two D7-branes described by the Higgsing (4.14), and choose a magnetization

$$F_2 = -M \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \left\{\mathrm{d}z^1 \wedge \mathrm{d}\bar{z}^{\bar{1}} - \mathrm{d}z^2 \wedge \mathrm{d}\bar{z}^{\bar{2}}\right\}. \tag{4.32}$$

The corresponding connection is

$$A_1 = -\frac{\mathrm{i}}{2} M \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \left\{z^1 \mathrm{d}\bar{z}^{\bar{1}} - \bar{z}^{\bar{1}}\mathrm{d}z^1 - z^2\mathrm{d}\bar{z}^{\bar{2}} + \bar{z}^{\bar{2}}\mathrm{d}z^2\right\}. \tag{4.33}$$

For simplicity of presentation we will take $M > 0$.

With a non-trivial connection, the equation of motion for the fermions becomes

$$0 = \hat{\Gamma}^\alpha D_\alpha \theta - \mathrm{i}\,\hat{\Gamma}_i \big[\Phi^i, \theta\big], \tag{4.34}$$

where D_α is the gauge-covariant derivative. Following the same decomposition and procedure as for the vector-like case, we again find (4.19) up to the replacements

$$\begin{aligned}
\partial_1\psi_m^\pm &\to \big(\partial_1 \pm M\bar{z}^{\bar{1}}\big)\psi_m^\pm, & \bar{\partial}_{\bar{1}}\psi_m^\pm &\to \big(\bar{\partial}_{\bar{1}} \mp Mz^1\big)\psi_m^\pm,\\
\partial_2\psi_m^\pm &\to \big(\partial_2 \mp M\bar{z}^{\bar{2}}\big)\psi_m^\pm, & \bar{\partial}_{\bar{2}}\psi_m^\pm &\to \big(\bar{\partial}_{\bar{2}} \pm Mz^2\big)\psi_m^\pm.
\end{aligned} \tag{4.35}$$

Again writing $\psi_3^\pm = \bar{z}^{\bar{2}}\psi_\pm$, we find

$$\begin{aligned}
0 = \Big\{ &\partial_1\bar{\partial}_{\bar{1}} + \partial_2\bar{\partial}_{\bar{2}} \pm M\big(\bar{z}^{\bar{1}}\bar{\partial}_{\bar{1}} - z^1\partial_1 - \bar{z}^{\bar{2}}\bar{\partial}_{\bar{2}} + z^2\partial_2\big)\\
&- M^2\big|z^1\big|^2 - \big(M^2+q^2\big)\big|z^2\big|^2\Big\}\psi_\pm.
\end{aligned} \tag{4.36}$$

Due to the self-duality of the magnetic flux, (4.36) is separable. Again using the polar decomposition $z^a = r_a \mathrm{e}^{\mathrm{i}\phi_a}$ and taking the ansatz (4.22), we find the equations

$$0 = \zeta_\pm'' + \frac{1}{r_1}\zeta_\pm' - \frac{m_1^2}{r_1^2}\zeta_\pm - 4M^2r_1^2\zeta_\pm + \big(-4\lambda \pm 4Mm_1\big)\zeta_\pm, \tag{4.37}$$

$$0 = \sigma_\pm'' + \frac{1}{r_2}\sigma_\pm' - \frac{m_2^2}{r_2^2}\sigma_\pm - 4\kappa^2r_2^2\sigma_\pm + \big(4\lambda \mp 4Mm_2\big)\sigma_\pm, \tag{4.38}$$

in which

$$\kappa = \sqrt{M^2+q^2}, \tag{4.39}$$

and λ is again a constant to be determined by boundary conditions. The solutions are

$$\begin{aligned}
\zeta_\pm\big(r_1\big) &= \mathrm{e}^{-Mr_1^2}\big(2Mr_1^2\big)^{|m_1|/2}\Big\{\mathcal{M}\big(\alpha;m_1+1;2Mr_1^2\big) + \mathcal{U}\big(\alpha;m_1+1;2Mr_1^2\big)\Big\}\\
\sigma_\pm\big(r_2\big) &= \mathrm{e}^{-\kappa r_2^2}\big(2\kappa r_2^2\big)^{|m_2|/2}L_{n_2}^{|m_2|}\big(2\kappa r_2^2\big),
\end{aligned} \tag{4.40}$$

with

$$\lambda = \kappa\big(2n_2 + |m_2| + 1\big) \pm Mm_2 \equiv M\big(2\alpha - \big|m_1\big| \pm m_1 - 1\big), \tag{4.41}$$

where the final relation defines α. In (4.40), $\mathcal{M}$ and $\mathcal{U}$ are the confluent hypergeometric functions of the first and second kinds,[9] and regularity requires that n_2 be a non-negative integer.

The chirality of the spectrum is a consequence of the different behavior of the different charges. It is most easily seen by considering the "missing" zero mode [175, 166]

$$\psi_3^{\pm} \sim \mathrm{e}^{-\kappa r_2^2} \mathrm{e}^{\mp M r_1^2} h\big(z^1\big), \tag{4.42}$$

where h is a holomorphic function of z^1. Since we have taken $M > 0$, only the $+$ sector gives rise to normalizable modes, and hence the spectrum is chiral. The fact that h is an arbitrary holomorphic function indicates that there are an infinite number of such chiral modes, as is consistent with the fact that the chiral index, which is proportional to $\int F_2$, is divergent. Upon compactification, further conditions are imposed on $h\big(z^1\big)$ (see e.g. [175]) and the spectrum becomes finite.

4.2 Non-chiral Mesons from D7-branes in AdS

We now consider vector-like mesons arising on intersecting D7-branes in $AdS_5 \times S^5$, building on the groundwork laid in §4.1. As discussed in the introduction, the configuration of interest is the gravity dual of $\mathcal{N} = 4$ SU(N) SYM with a $\mathrm{U}(1) \times \mathrm{U}(1)$ flavor group. The strings stretching between the D7-branes are dual to mesonic operators with charges $(\pm 1, \mp 1)$ under this $\mathrm{U}(1) \times \mathrm{U}(1)$. Our analysis has much in common with the treatment of intersecting D7-branes in weakly warped geometries [166]; however, $AdS_5 \times S^5$ is strongly warped in the sense that no limit

[9] Since the confluent hypergeometric function ${}_1F_1\left(a; b; z\right)$ is not defined when $b = 0, -1, -2, \ldots$, we use the regularized version $\mathcal{M}\left(a; b; z\right) = {}_1F_1\left(a; b; z\right)/\Gamma\left(b\right)$.

of the geometry reproduces a factorized geometry $\mathbb{R}^{3,1} \times X^6$, and so we will need to use different techniques to solve the resulting equations of motion.

4.2.1 Setup and equations of motion

The metric for $AdS_5 \times S^5$ can be written as a warped product of $\mathbb{R}^{3,1}$ and $\mathbb{C}^3$,

$$\mathrm{d}s_{10}^2 = \mathrm{e}^{2\mathcal{A}}\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu + \mathrm{e}^{-2\mathcal{A}}\mathrm{d}z^I \mathrm{d}\bar{z}^{\bar{I}}, \qquad \mathcal{A} = \frac{1}{2}\log\frac{z^I\bar{z}^{\bar{I}}}{L^2}. \tag{4.43}$$

Using hyperspherical coordinates on $\mathbb{C}^3 = \mathbb{R}^6$, this becomes the familiar metric for $AdS_5 \times S^5$,

$$\mathrm{d}s_{10}^2 = \frac{R^2}{L^2}\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu + \frac{L^2}{R^2}\mathrm{d}R^2 + L^2\mathrm{d}s_{S^5}^2, \tag{4.44}$$

where $\mathrm{d}s_{S^5}^2$ is the standard metric on a unit S^5. The geometry is supported by the 5-form flux

$$F_5 = \big(1 + \hat{*}\big)g_\mathrm{s}^{-1}\mathrm{d}\mathrm{e}^{4\mathcal{A}} \wedge \mathrm{dvol}_{\mathbb{R}^{3,1}}, \tag{4.45}$$

where $\hat{*}$ is the 10d Hodge star. In the presence of such flux, the action for a single D7-brane becomes [170, 171, 172]

$$\begin{aligned} S_{\mathrm{D7}} = -\frac{1}{g_8^2}\int_{\mathcal{W}} \mathrm{d}^8\xi^\alpha \sqrt{-\hat{g}}\Big\{ &\frac{1}{2}\hat{g}_{ij}\hat{g}^{\alpha\beta}\partial_\alpha\Phi^i\partial_\beta\Phi^j + \frac{1}{4}\hat{g}^{\alpha\beta}\hat{g}^{\gamma\delta}F_{\alpha\gamma}F_{\beta\delta} \\ &+ \mathrm{i}\bar{\Theta}P_-^{\mathrm{D7}}\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha\hat{\nabla}_\beta\Theta + \frac{g_\mathrm{s}}{8\cdot 4!}\hat{\epsilon}^{\alpha_1\cdots\alpha_8}C_{\alpha_1\cdots\alpha_4}F_{\alpha_5\alpha_6}F_{\alpha_7\alpha_8} \\ &+ \frac{\mathrm{i}g_\mathrm{s}}{16}\bar{\Theta}P_-^{\mathrm{D7}}\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha\hat{\not{F}}_5\hat{\Gamma}_\beta\big(\mathrm{i}\sigma_2\big)\Theta\Big\}, \end{aligned} \tag{4.46}$$

in which

$$\hat{\not{F}}_5 = \frac{1}{5!}F_{M_1\cdots M_5}\hat{\Gamma}^{M_1\cdots M_5}, \tag{4.47}$$

is constructed by contracting all indices of F_5 with $\hat{\Gamma}$-matrices, and not just those along the worldvolume. If the D7-brane fills $\mathbb{R}^{3,1}$ and a cycle $\mathcal{S}^4$ in the other

directions, then after κ-fixing to (4.7) and taking into account the nontrivial spin connection, the fermionic contribution to the action is [176]

$$S^{\mathrm{F}}_{\mathrm{D7}} = -\frac{\mathrm{i}}{2g_8^2}\int \mathrm{d}^8\xi^\alpha \sqrt{-\hat{g}}\,\bar{\theta}\Big\{\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha\partial_\beta + \frac{1}{2}\hat{g}^{\alpha\beta}\hat{\Gamma}_\alpha\partial_\beta\mathcal{A}\big(1+2\hat{\Gamma}_{\mathcal{S}^4}\big)\Big\}\theta, \tag{4.48}$$

in which

$$\hat{\Gamma}_{\mathcal{S}^4} = \mathrm{dv}\hat{\mathrm{o}}\mathrm{l}_{\mathcal{S}^4}, \tag{4.49}$$

is the chirality operator on $\mathcal{S}^4$.

In the non-Abelian case, closed-string fields like the warp factor are interpreted as Taylor series in the adjoint-valued transverse deformations, and thus the closed-string fields are themselves adjoint-valued [177]. However, as in the unwarped case, this fact is only important for higher-dimension operators, and can be neglected to leading order in ℓ_{s}. Similar terms are expected in the non-Abelian fermionic action, but have not been computed explicitly. However, to leading order in ℓ_{s}, supersymmetry and gauge-invariance require that the action take the form [178, 166]

$$\begin{aligned} S^{\mathrm{F}}_{\mathrm{D7}} = &-\frac{\mathrm{i}}{2g_8^2}\int \mathrm{d}^8\xi^\alpha\sqrt{-\hat{g}}\times \\ &\mathrm{tr}\left\{\bar{\theta}\,\hat{\Gamma}^\alpha D_\alpha\theta - \frac{1}{2}\bar{\theta}\,\hat{\Gamma}_i\big[\Phi^i,\theta\big] + \frac{1}{2}\bar{\theta}\,\hat{\Gamma}^\alpha\partial_\alpha\mathcal{A}\big(1+2\hat{\Gamma}_{\mathcal{S}^4}\big)\theta\right\}. \end{aligned} \tag{4.50}$$

The intersection of two D7-branes satisfying

$$\mathrm{D7}_1 : z^3 = \mu + tz^2, \quad \mathrm{D7}_2 : z^3 = \mu - tz^2, \tag{4.51}$$

is described by

$$\Phi = \begin{pmatrix} \ell_{\mathrm{s}}^{-2}\mu + qz^2 & \\ & \ell_{\mathrm{s}}^{-2}\mu - qz^2. \end{pmatrix}. \tag{4.52}$$

When $\mu = 0$, the D7-branes reach the origin of warping and the dual quarks are massless: in the D-brane picture, the D3-branes and D7-branes intersect and

the strings stretching between them have zero length. However, when there is a finite separation between the branes, the quarks have a mass proportional to μ. Consequently, the mesonic spectrum becomes gapped [126]. The warp factor is to be evaluated at this vev, but so long as t is sufficiently small, on the D7-brane we can take

$$\mathcal{A} = \frac{1}{2}\log\frac{|z^1|^2 + |z^1|^2 + \mu^2}{L^2}. \tag{4.53}$$

Decomposing θ as (4.16) and matching terms of internal chirality, we find

$$0 = \big(\bar\partial_{\bar1} - \frac{1}{2}\bar\partial_{\bar1}\mathcal{A}\big)\psi_1^\pm - \big(\bar\partial_{\bar2} - \frac{1}{2}\bar\partial_{\bar2}\mathcal{A}\big)\psi_2^\pm \mp \mathrm{i}q\mathrm{e}^{-2\mathcal{A}}z^2\psi_3^\pm, \tag{4.54a}$$

$$0 = \big(\partial_2 + \frac{3}{2}\partial_2\mathcal{A}\big)\psi_3^\pm \mp \mathrm{i}q\mathrm{e}^{-2A}\bar z^{\bar2}\psi_2^\pm, \tag{4.54b}$$

$$0 = \big(\partial_1 + \frac{3}{2}\partial_1\mathcal{A}\big)\psi_3^\pm \pm \mathrm{i}q\mathrm{e}^{-2A}\bar z^{\bar2}\psi_1^\pm, \tag{4.54c}$$

$$0 = \big(\partial_1 - \frac{1}{2}\partial_1\mathcal{A}\big)\psi_2^\pm + \big(\partial_2 - \frac{1}{2}\partial_2\mathcal{A}\big)\psi_1^\pm, \tag{4.54d}$$

where, as in the flat space analysis of §4.1, we have evaluated the equations at zero 4d momentum and have set $\psi_0^\pm = 0$. Taking, as in [176]

$$\psi_{1,2}^\pm = \mathrm{e}^{\mathcal{A}/2}\varphi_{1,2}^\pm, \quad \psi_3^\pm = \mathrm{e}^{-3\mathcal{A}/2}\varphi_3^\pm, \tag{4.55}$$

and finally writing $\varphi_3^\pm = \bar z^{\bar2}\varphi_\pm$, we find the warped analogue of (4.21)

$$0 = \big\{\partial_1\bar\partial_{\bar1} + \partial_2\bar\partial_{\bar2} - q^2\,|z^2|^2\,\mathrm{e}^{-4\mathcal{A}}\big\}\varphi_\pm. \tag{4.56}$$

Since the warp factor depends on both z^1 and z^2, (4.56) is not separable in those variables. However, writing

$$z^1 = r\cos\beta\,\mathrm{e}^{\mathrm{i}\phi_1}, \quad z^2 = r\sin\beta\,\mathrm{e}^{\mathrm{i}\phi_2}, \tag{4.57}$$

equation (4.56) becomes

$$0 = \left\{\partial_r^2 + \frac{3}{r}\partial_r + \frac{1}{r^2}\breve\nabla^2 - \frac{4q^2r^2L^4\sin^2\beta}{(r^2+\mu^2)^2}\right\}\varphi_\pm, \tag{4.58}$$

in which

$$\breve{\nabla}^2 = \partial_\beta^2 + \big(\cot\beta - \tan\beta\big)\partial_\beta + \frac{1}{\cos^2\beta}\partial_{\phi_1}^2 + \frac{1}{\sin^2\beta}\partial_{\phi_2}^2 \tag{4.59}$$

is the Laplacian on a unit S^3 (see Appendix 4.B).

When $\mu = 0$, (4.58) is completely separable. Indeed, taking

$$\varphi_\pm = \mathrm{e}^{\mathrm{i}(m_1\phi_1 + m_2\phi_2)} f_\pm\big(r\big) Q_\pm\big(\cos 2\beta\big), \tag{4.60}$$

we find that the radial equation satisfies

$$0 = f_\pm'' + \frac{3}{r} f_\pm' - \frac{\lambda}{r^2} f_\pm, \tag{4.61}$$

while the β equation is

$$0 = 4\big(1-x^2\big)Q_\pm'' - 8xQ_\pm' - \frac{2m_1^2}{1+x}Q_\pm - \frac{2m_2^2}{1-x}Q_\pm - 2\xi^2\big(1-x\big)Q_\pm + \lambda Q_\pm, \tag{4.62}$$

in which $x = \cos 2\beta$,

$$\xi^2 \equiv q^2 L^4 = \frac{1}{\pi} t^2 g_{\mathrm{s}} N, \tag{4.63}$$

and λ is a constant to be determined by boundary conditions.[10]

4.2.2 The meson spectrum

When $\xi = 0$, the solutions to (4.62) are the scalar hyperspherical harmonics (see Appendix 4.B)

$$Q_\pm\big(x\big) = c\big(1+x\big)^{m1/2}\big(1-x\big)^{m2/2} P^{(m_2,m_1)}_{\frac{1}{2}(\ell - m_1 - m_2)}\big(x\big), \tag{4.64}$$

where $P_n^{(a,b)}$ are the Jacobi Polynomials, c is the normalization constant (4.132), $\lambda = \ell\,(\ell+2)$, and the quantum numbers must satisfy the inequalities $0 \leq |m_1| + |m_2| \leq \ell$ and the constraint $\frac{1}{2}\,(\ell - m_1 - m_2) \in \mathbb{Z}$.

[10] Note that in this section and the next, λ carries no dimensions, in contrast to the previous section.

We have been unable to find analytic solutions to (4.62) when $\xi \neq 0$. However, since (4.62) is an ordinary differential equation, numerical methods readily apply. We implement a spectral method by expanding the unknown solution in terms of the spherical harmonics. The potential term proportional to ξ does not mix modes of different m_1 and m_2, so we can accomplish the spectral decomposition by writing

$$Q(x) = \sum_{\ell} b_\ell y_\ell(x), \tag{4.65}$$

where y_ℓ are the solutions (4.64) and we have suppressed other indices. Equation (4.62) then becomes

$$0 = \sum_{\ell} \Big\{ \lambda - \ell\,(\ell+2) - 2\xi^2(1-x) \Big\} b_\ell y_\ell. \tag{4.66}$$

Using that at fixed m_1 and m_2,

$$\int_{-1}^{1} \mathrm{d}x\, y_\ell y_{\ell'} = \frac{1}{\pi^2}\delta_{\ell\ell'}, \tag{4.67}$$

and using the recursion relationship (4.133), we can re-express (4.66) as the matrix equation

$$0 = \Big[\lambda - \ell\,(\ell+2) - 2\xi^2 d_0\Big] b_\ell + 2\xi^2 d_- b_{\ell-2} + 2\xi^2 d_+ b_{\ell+2}, \tag{4.68}$$

with

$$\begin{aligned} d_0 =& \left(1 + \frac{m_2^2 - m_1^2}{\ell\,(\ell+2)}\right), \\ d_- =& \sqrt{\frac{\left(\ell^2 - (m_1+m_2)^2\right)\left(\ell^2 - (m_1-m_2)^2\right)}{4\,(\ell+1)\,\ell^2\,(\ell-1)}}, \\ d_+ =& \sqrt{\frac{\left((\ell+2)^2 - (m_1+m_2)^2\right)\left((\ell+2)^2 - (m_1-m_2)^2\right)}{4\,(\ell+1)\,(\ell+2)^2\,(\ell+3)}}. \end{aligned} \tag{4.69}$$

Note that even and odd ℓs do not mix, so that this effectively gives two independent matrix equations where the matrices are each tridiagonal.

Solving (4.62) amounts to diagonalization of the matrix defined by (4.68). Unfortunately, because this is an infinite-dimensional matrix, we cannot perform this

diagonalization exactly. However, to obtain an estimate of the spectrum, we can truncate the matrix to a finite submatrix. A good rule of thumb in such problems is that including the first $2n$ modes determines the first n eigenvalues to an accuracy of a few percent [179]. Accurate eigenvalues will be robust against variations in n, and our strategy will be to increase the number of modes included until the eigenvalues calculated in this way stabilize. The first few eigenvalues at $m_1 = m_2 = 0$ resulting from this process are shown in figures 4.3 and 4.4. As ξ increases, the wavefunctions become increasingly localized on the intersection at $\beta = 0$, as shown in figure 4.5.

Note that when $\xi \ll 1$, (4.68) immediately yields the perturbative result

$$\lambda \approx \ell\,(\ell+2) + 2\xi^2\left[1 + \frac{m_2^2 - m_1^2}{\ell\,(\ell+2)}\right]. \tag{4.70}$$

However, since $\xi^2 = t^2 g_{\rm s} N/\pi$, working at $\xi \ll 1$ requires taking t^2 to be small with respect to the inverse 't Hooft coupling $1/\lambda$. This limit is of little utility in the present investigation, because we are interested in taking $\lambda \to \infty$ to suppress α' corrections to the leading-order supergravity, cf. (4.1).

If instead $\xi \gg 1$, we find that the spectrum is well-approximated by

$$\lambda \approx 4\xi\big(\ell + |m_1| - 1\big). \tag{4.71}$$

At large ξ, ℓ is no longer a good quantum number, as the intersection badly breaks the rotational symmetry of the S^3. Correspondingly, the solutions to (4.62) are linear combinations of many different spherical harmonics. However, m_1 and m_2 remain good quantum numbers, and so we find it more natural to write the spectrum as

$$\lambda \approx 4\xi\big(n + |m_2| + 1\big), \tag{4.72}$$

where $n = \ell - |m_1| - |m_2|$.

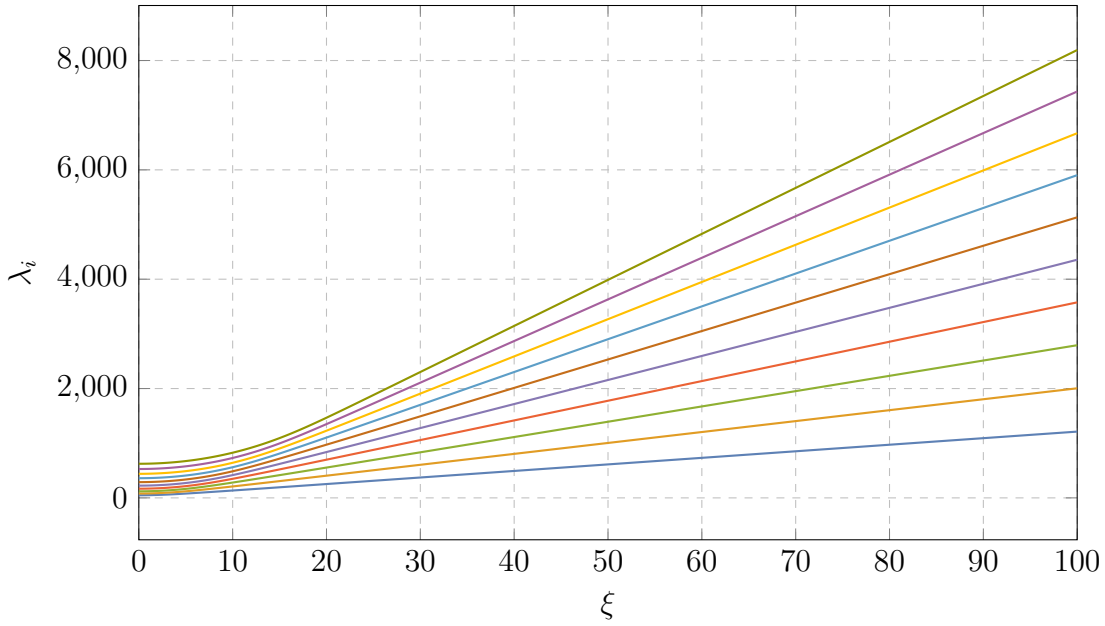

Figure 4.3: The first few eigenvalues of (4.62) found via spectral methods, for $m_1 = m_2 = 0$. The growth continues to be linear as ξ increases.

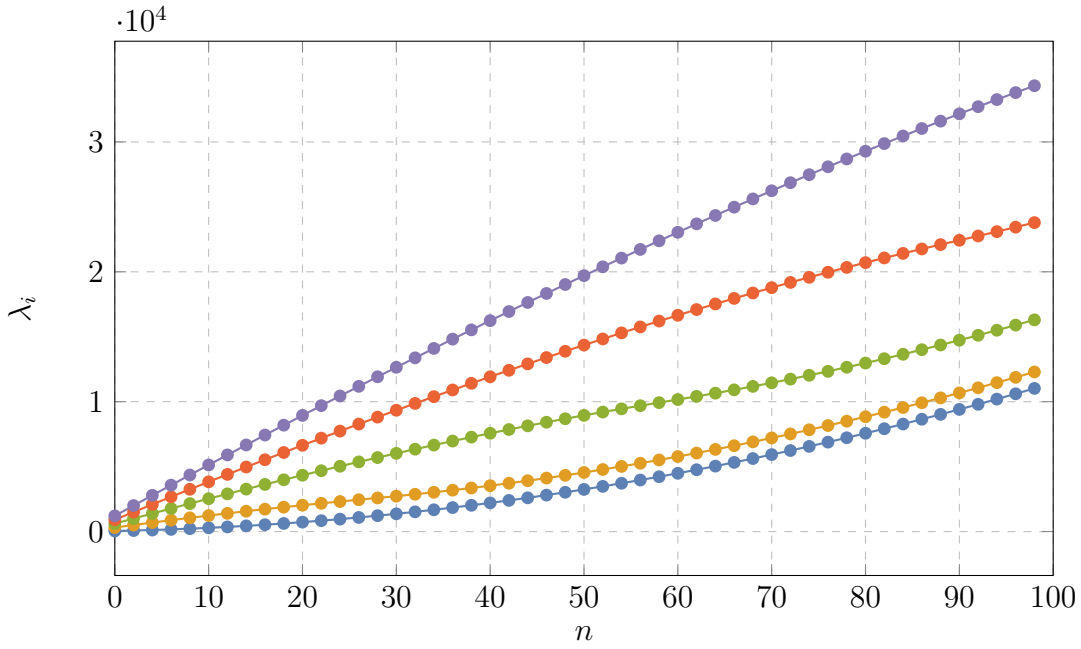

Figure 4.4: The spectrum for $m_1 = m_2 = 0$ (which requires that ℓ be even) for $\xi = 0$ (bottom), 25, 50, 75, and 100 (top).

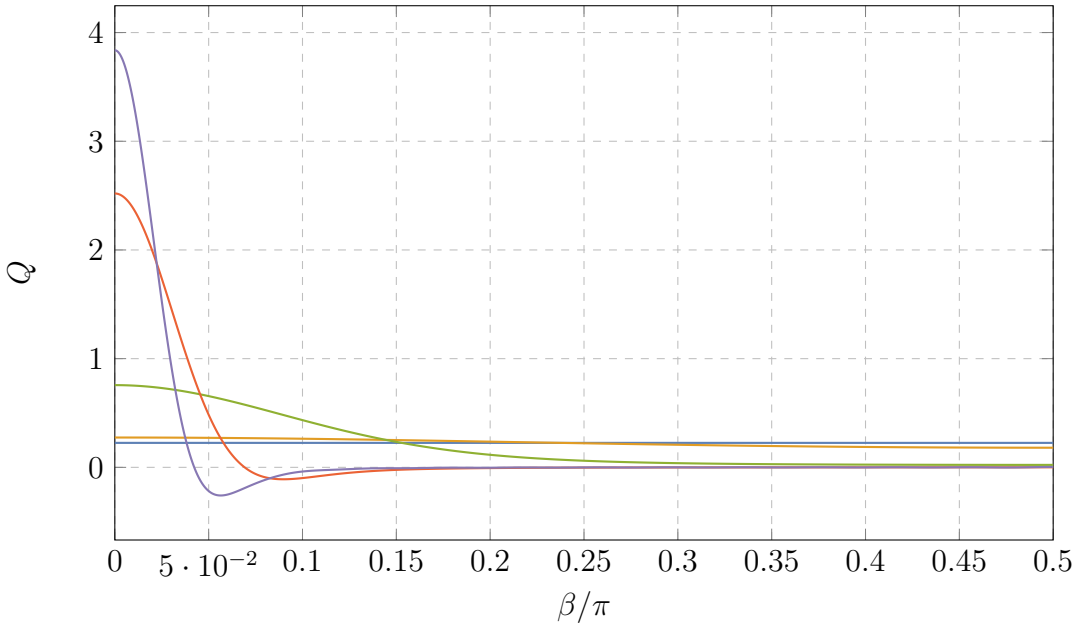

Figure 4.5: The lowest-lying solutions of (4.62) for $\xi = 0, 2.5, 10, 50, 100$. When $\xi = 0$, the solution is a constant zero mode, but as ξ increases, the profile becomes increasingly peaked at $\beta = 0$, the location of the intersection.

With the eigenvalues of (4.62) in hand, the solution to (4.61) is

$$f_{\pm} = c_1 r^{-1-\sqrt{1+\lambda}} + c_2 r^{-1+\sqrt{1+\lambda}}. \tag{4.73}$$

We can compare the solution (4.73) to the well-known result for a canonically normalized scalar at zero momentum,

$$\varphi = \varphi_0 r^{\Delta-4} + \varphi_1 r^{-\Delta}. \tag{4.74}$$

The solution (4.73) does not match the form (4.74), since the transverse deformations are not canonically normalized (see (4.77)). Nevertheless, Δ can be determined by taking the ratio of the two terms in (4.73), and we find the result

$$\Delta = 2 + \sqrt{1+\lambda}. \tag{4.75}$$

This then gives the approximate expressions

$$\Delta \approx \begin{cases} \ell + 3 + \frac{\xi^2}{1+\ell}\left[1 + \frac{m_2^2 - m_1^2}{\ell(\ell+2)}\right] & \xi \ll 1, \\ 2\sqrt{\xi\,(n + |m_2| + 1)} & \xi \gg 1 \end{cases}. \tag{4.76}$$

The fact that the radial modes are simply power laws is an indication that the dual theory is conformal. Indeed, one can confirm that the $\mu = 0$ configuration (4.51) respects the supersymmetry generated by eight supercharges, four of which correspond to the generators of superconformal transformations in the dual theory. Alternatively, when $\mu = 0$, the vev (4.52) corresponds to a strictly marginal deformation of the theory. To see this, it suffices to consider the Abelian action (4.46) and examine only the action of the transverse scalars Φ^i. Using the complexified field Φ and expanding in scalar spherical harmonics gives the 5d action

$$S \sim -\int \mathrm{d}^5x\,\sqrt{-g}\sum_{\ell=0}^{\infty}\left\{\frac{L^2}{r^2}g^{mn}\partial_m\Phi_\ell^\dagger\partial_n\Phi_\ell + \frac{\ell\,(\ell+2)}{r^2}\Phi_\ell^\dagger\Phi_\ell\right\}. \tag{4.77}$$

Defining the canonically normalized scalars $\chi_\ell = \frac{L}{r}\Phi_\ell$ gives

$$S \sim -\int \mathrm{d}^5x\,\sqrt{-g}\sum_{\ell=0}^{\infty}\left\{g^{mn}\partial_m\chi_\ell^\dagger\partial_n\chi_\ell + \frac{\ell\,(\ell+2)-3}{L^2}\chi_\ell^\dagger\chi_\ell\right\}. \tag{4.78}$$

Using the familiar result

$$\Delta = 2 + \sqrt{4 + m^2L^2} \tag{4.79}$$

yields

$$\Delta = \ell + 3. \tag{4.80}$$

With the coordinates of (4.57), the configuration $\Phi = qz^2$ can be expressed as

$$\Phi = qr\,\sin\beta\,\mathrm{e}^{\mathrm{i}\phi_2} = \frac{qr}{\sqrt{2}}\sqrt{1-\cos 2\beta}\,\mathrm{e}^{\mathrm{i}\phi_2}. \tag{4.81}$$

Comparing to (4.64), the mode (4.81) corresponds to $\ell = 1$, $m_1 = 0$, $m_2 = 1$, and hence this configuration is the non-normalizable solution of the $\Delta = 4$ mode, and so describes a marginal deformation of the dual theory.

4.3 Chiral Mesons from D7-branes in AdS

Just as in the flat space case, we can induce chirality into the dual theory through the introduction of a supersymmetric magnetic flux (4.30). However, this magnetic flux will respect only four of the gravity supercharges, and the other four, corresponding to the superconformal charges of the dual theory, will not be preserved. As we shall see, this change has important physical consequences: the calculation of correlation functions will turn out to require counterterms that are super-exponentially sensitive to the ultraviolet completion of the geometry. At the same time, the magnetic flux induces a large amount of D3-brane charge, so that the geometry must be sharply modified in the ultraviolet. In practical terms, this dependence on the ultraviolet behavior presents an obstacle to the calculation of correlation functions. More importantly, it signifies that the magnetization (4.30), and the corresponding appearance of chiral mesons, entails a substantial change in the background.

4.3.1 Setup and equations of motion

We first sketch out the argument regarding the supercharges. A probe D7-brane will preserve the supersymmetry parameterized by a Killing double spinor ϵ if (cf. (4.5))

$$P_{-}^{\mathrm{D7}}\epsilon = 0, \tag{4.82}$$

where, with the presence of a magnetic flux F_2, $\Gamma_{\mathrm{D7}} = -\mathrm{i}\Gamma_{(8)}L\left(F\right)$ with

$$L\big(F\big) = \sqrt{\frac{\det\left(\hat{g}\right)}{\det\left(\hat{g}+\ell_{\mathrm{s}}^2 F\right)}}\left\{1+\frac{\ell_{\mathrm{s}}^2}{2}F_{\alpha_1\alpha_2}\hat{\Gamma}^{\alpha_1\alpha_2}+\frac{\ell_{\mathrm{s}}^4}{8}F_{\alpha_1\alpha_2}F_{\alpha_3\alpha_4}\hat{\Gamma}^{\alpha_1\alpha_2\alpha_3\alpha_4}\right\}. \tag{4.83}$$

The bulk geometry respects the supersymmetry generated by a GKP-like Killing spinor [131], which is independent of the Minkowski coordinates and annihilated by holomorphic γ-matrices. Moreover, such a Killing spinor obeys (4.82) if F_2 is $(1,1)$ and self-dual: the $\not{F}_2$ term annihilates the Killing spinor, and the 1 and $\not{F}_2^2$ terms together are canceled by $\sqrt{\det(g+F)}$ (which takes a simple form because F is self-dual). However, the bulk geometry also supports Killing spinors that depend on the Minkowski coordinates in a particular way (see, e.g., [180]). The existence of such spinors is a special feature of anti-de Sitter space, and the supersymmetry transformations they induce are dual to superconformal transformations. Since the special AdS Killing spinors are not preserved by the magnetized D7-brane configuration, we anticipate that conformality will be lost in the dual theory, even in the probe approximation.

We can also understand the loss of conformality from another point of view. The magnetization that gives rise to chirality follows from the connection (4.31) which, using (4.57), can be written as

$$A_1 = Mr^2\big\{-\cos^2\beta\,\mathrm{d}\phi_1 + \sin^2\beta\,\mathrm{d}\phi_2\big\}, \tag{4.84}$$

in which we are still taking $M > 0$ for simplicity of presentation. Writing $A_1 = Mr^2\omega$, ω satisfies the defining equation of a transverse vector spherical harmonic ϱ, which takes the general form

$$\breve{\nabla}^2\varrho_\theta = -\big[\ell\big(\ell+2\big)-1\big]\varrho_\theta, \quad \breve{\nabla}^\theta\varrho_\theta = 0, \quad \breve{\epsilon}^{\theta\varphi\psi}\breve{\nabla}_\varphi\varrho_\psi = \pm\big(\ell+1\big)\breve{g}^{\theta\psi}\varrho_\psi, \tag{4.85}$$

in which $\breve{g}$ is the metric (4.122) on the unit S^3, $\breve{\nabla}$ is the associated Levi-Civita connection, and $\breve{\epsilon}$ is the associated volume form. The mode ω corresponds to the specific case $\ell = 1$, with the positive sign taken in the third equation in (4.85).[11] Thus,

[11] This sign is independent of the sign in the equation of motion for the bifundamental modes, (4.19).

ω is a transverse vector spherical harmonic [181], and upon dimensional reduction leads to a canonically normalized field with mass $m^2L^2 = 12$ (see [126]). This corresponds to an operator of dimension 6, and A_1 involves the non-normalizable solution. Hence, the introduction of the magnetic flux deforms the dual theory by an irrelevant operator. This implies that not only is conformality lost in the dual theory, but the theory does not even flow from an ultraviolet fixed point.

The addition of this flux modifies the zero-momentum equations to

$$0 = \big(\bar\partial_{\bar1} \mp Mz^1 - \frac{1}{2}\bar\partial_{\bar1}\mathcal{A}\big)\psi_1^\pm - \big(\bar\partial_{\bar2} \pm Mz^2 - \frac{1}{2}\bar\partial_{\bar2}\mathcal{A}\big)\psi_2^\pm \mp \mathrm{i}q\mathrm{e}^{-2\mathcal{A}}z^2\psi_3^\pm, \tag{4.86a}$$

$$0 = \big(\partial_2 \mp M\bar z^{\bar2} + \frac{3}{2}\partial_2\mathcal{A}\big)\psi_3^\pm \mp \mathrm{i}q\mathrm{e}^{-2A}\bar z^{\bar2}\psi_2^\pm, \tag{4.86b}$$

$$0 = \big(\partial_1 \pm M\bar z^{\bar1} + \frac{3}{2}\partial_1\mathcal{A}\big)\psi_3^\pm \pm \mathrm{i}q\mathrm{e}^{-2A}\bar z^{\bar2}\psi_1^\pm, \tag{4.86c}$$

$$0 = \big(\partial_1 \pm M\bar z^{\bar1} - \frac{1}{2}\partial_1\mathcal{A}\big)\psi_2^\pm + \big(\partial_2 \mp M\bar z^{\bar2} - \frac{1}{2}\partial_2\mathcal{A}\big)\psi_1^\pm. \tag{4.86d}$$

Using (4.55), we find

$$\begin{aligned} 0 = \Big\{ &\partial_1\bar\partial_{\bar1} + \partial_2\bar\partial_{\bar2} \pm M\big(\bar z^{\bar1}\bar\partial_{\bar1} - z^1\partial_1 - \bar z^{\bar2}\bar\partial_{\bar2} + z^2\partial_2\big) \\ &- M^2\big|z^1\big|^2 - \big(M^2 + \mathrm{e}^{-4\mathcal{A}}q^2\big)\big|z^2\big|^2\Big\}\varphi_\pm, \end{aligned} \tag{4.87}$$

where $\varphi_\pm = \frac{1}{\bar z^2}\varphi_\pm^3$. With the coordinates (4.57), this becomes

$$\begin{aligned} 0 = \Big\{ &\partial_r^2 + \frac{3}{r}\partial_r \pm 4\mathrm{i}M\big(\partial_{\phi_1} - \partial_{\phi_2}\big) - 4M^2r^2 \\ &+ \frac{1}{r^2}\Big[\partial_\beta^2 + \big(\cot\beta - \tan\beta\big)\partial_\beta + \frac{1}{\cos^2\beta}\partial_{\phi_1}^2 + \frac{1}{\sin^2\beta}\partial_{\phi_2}^2\Big] - \frac{4q^2r^2L^4\sin^2\beta}{\big(r^2+\mu^2\big)^2}\Big\}\varphi_\pm. \end{aligned} \tag{4.88}$$

Again, the relative simplicity of this equation is a consequence of the self-duality constraint imposed by supersymmetry. When $\mu = 0$, the equation is again separable and it is useful to take the ansatz (4.60). $Q_\pm$ satisfies the same eigenvalue problem (4.62) while the radial equation is now

$$0 = f_\pm'' + \frac{3}{r}f_\pm' \mp 4M\big(m_1 - m_2\big)f_\pm - 4M^2r^2f_\pm - \frac{\lambda}{r^2}f_\pm. \tag{4.89}$$

The solutions can be expressed in terms of $\mathcal{M}$ and $\mathcal{U}$, the confluent hypergeometric functions of the first and second kind,

$$f_\pm = \mathrm{e}^{-Mr^2} r^{-\nu} \Big\{ c_1 \, \mathcal{M}\big(\mu;\nu;2Mr^2\big) + c_2 \, \mathcal{U}\big(\mu;\nu;2Mr^2\big) \Big\}, \tag{4.90}$$

in which

$$\nu = 1 + \sqrt{1+\lambda} \qquad \text{and} \qquad \mu = \frac{1}{2}\big(\nu \pm \big(m_1 - m_2\big)\big). \tag{4.91}$$

As anticipated, the solutions are not power laws, and so the dual field theory is no longer conformal even in the probe approximation. Furthermore, noting that the dominant asymptotic behavior at $r \to \infty$ is

$$\mathcal{M}\big(\mu;\nu;2Mr^2\big) \propto e^{2Mr^2}, \tag{4.92}$$

where we have omitted power law factors, we find that the divergent part of (4.90) grows super-exponentially at $r \to \infty$:

$$f_\pm \propto e^{Mr^2}. \tag{4.93}$$

4.3.2 Ultraviolet sensitivity of the correlation functions

To interpret the divergences identified above, it will be helpful to recall the well-established procedure for computing correlation functions in AdS/CFT, focusing on the process of removing divergences of the classical action through the introduction of counterterms, i.e. holographic renormalization (see [182] for a review).

The basic statement of the duality, in the limit (4.1), is the identification of the generating functional of the CFT with the classical supergravity action,

$$\mathcal{Z}_{\mathrm{CFT}} = \mathrm{e}^{-S_{\mathrm{grav}}}. \tag{4.94}$$

An operator $\mathcal{O}$ on the field theory side has a corresponding classical field φ on the gravity side. If $\mathcal{O}$ is a scalar field, then φ also transforms as an $\mathrm{SO}(3,1)$ scalar. The solution for φ at large r can be separated into a dominant term and a subdominant term,

$$\varphi = a_{\mathrm{dom}}\varphi_{\mathrm{dom}} + a_{\mathrm{sub}}\varphi_{\mathrm{sub}}. \tag{4.95}$$

If the geometry is asymptotically anti-de Sitter space, both the dominant and subdominant terms are power laws at large r. Moreover, a_{dom} is dual to a source term for $\mathcal{O}$, and correlation functions of $\mathcal{O}$ are calculated by taking functional derivatives of S_{grav} with respect to a_{dom} and then later taking $a_{\mathrm{dom}} \to 0$.

For finite a_{dom}, the classical action S_{grav} is divergent. This can be addressed by adding counterterms to the action: one first regulates the action by cutting off the space at a large but finite radius r_Λ. The terms that diverge as $r_\Lambda \to \infty$ are canceled by adding terms to the supergravity action that are localized on the boundary at r_Λ. Taking $r_\Lambda \to \infty$ then yields a finite action. The power law behavior of solutions in the AdS case means that such counterterms have power-law (and potentially logarithmic) dependence on r_Λ. However, the super-exponential growth (4.93) of the chiral modes requires the introduction of counterterms that have a similar super-exponential dependence on the cutoff. Since the magnetization required to induce chirality deforms the theory by an irrelevant operator, such strong sensitivity is perhaps not surprising.

If the background remained unaltered by magnetization, the structure of counterterms would represent a technically demanding but potentially surmountable challenge to calculating correlation functions.[12] However, the chirality-inducing

[12]For example, as developed in [183, 184], it is possible to calculate correlation functions in the KT/KS theory [185, 138], even though the theory does not flow from an ultraviolet fixed point.

magnetic flux sources a large amount of D3-brane charge via the Chern-Simons coupling $\int C_4 \wedge F_2 \wedge F_2$: the dissolved D3-brane flux diverges as

$$\int_{R<\rho} F_2 \wedge F_2 \sim M^2 \rho^4. \tag{4.96}$$

This is comparable to the D3-brane charge of the background when

$$\rho \sim \frac{N^{1/4}}{M^{1/2}}, \tag{4.97}$$

at which point the influence of this charge on the geometry must be taken into account. A calculation of correlation functions that fails to incorporate this back-reaction is not physically meaningful.

One might ask whether a different choice of magnetization (still without a localized source) results in a different conclusion. Supersymmetric fluxes supported on the D7-branes are characterized by scalar hyperspherical harmonics — cf. (4.29) — and so the fluxes grow as $F_2 \sim r^{j+2}\Omega^{(j)} + r^{j+1}\mathrm{d}r \wedge \omega^{(j)}$, where $j = 0, 1, 2, \ldots$, and $\Omega^{(j)}$ and $\omega^{(j)}$ are a 2-form and a 1-form on S^3, respectively. Our analysis of the magnetic flux (4.32) corresponds to the case $j = 0$. Other values of j would lead to steeper potentials in (4.87), and so to a greater degree of localization of the bifundamental wavefunctions. However, the charge carried by such flux diverges more quickly than (4.96), growing as r^{2j+4}, and hence the problem of ultraviolet sensitivity is exacerbated.

4.4 Conclusions

In this note we analyzed the spectrum of mesonic operators arising from strings stretching between intersecting D7-branes in $AdS_5 \times S^5$. The dual field theory is an $\mathcal{N} = 1$ deformation of maximally supersymmetric SU(N) SYM, with the

addition of a $\mathrm{U}(F_1) \times \mathrm{U}(F_2)$ flavor group, under which the 7-7′ strings transform as bifundamentals.[13] We considered D7-branes with and without magnetic flux on the curve of intersection, finding sharply different results in these two cases.

The intersection of the D7-branes corresponds to a particular adjoint Higgsing of the U(2) theory arising on coincident D7-branes. In the field theory, the fact that the branes intersect is described by a marginal deformation. If the D7-branes reach the origin of warping, and one furthermore makes the quenched/probe approximation that neglects backreaction of the D7-branes, then the dual theory is conformal. In this case — where magnetization has not yet been incorporated — we computed the spectrum of dual operators. The 7-7′ strings are mixtures of the transverse deformations and the internal components of the gauge field, and as a consequence the equations of motion are difficult to solve analytically. However, conformal symmetry leads to a remarkable simplification of the equations of motion, through which we were able to find numerical solutions. The behavior of the dimensions depends on the value of $\xi \sim \tan\theta\sqrt{g_{\mathrm{s}}N}$, cf. (4.63), where θ is an angle characterizing the intersection. Approximate spectra are given in (4.76). As expected, the modes are well localized along the intersection of the D7-branes and have power-law behavior along the holographic direction.

We then considered introducing magnetic flux on the curve of intersection, leading to a chiral spectrum in the dual theory. The simplest magnetization corresponds to an irrelevant deformation of the theory, by an operator of dimension $\Delta = 6$. As a consequence, the non-normalizable solutions to the bifundamental equations of motion have super-exponential divergence in the ultraviolet, cf. (4.93). Although the limit (4.1) allows us to neglect the backreaction of the D7-branes

[13] For notational simplicity only, we limited our discussion to the case $F_1 = F_2 = 1$, corresponding to a single pair of D7-branes.

themselves, the backreaction of the D3-brane charge induced by the magnetic flux cannot be neglected. Since the calculation of correlation functions, for example through holographic renormalization, requires the use of the non-normalizable modes, the procedure for calculating the correlation functions is unclear. This is a physical limitation rather than a technical one: the divergence of the D3-brane charge induced by magnetization of noncompact D7-branes signals the need for an ultraviolet completion via compactification. In the dual language, the field theory describing magnetized D7-branes does not flow from an ultraviolet fixed point.

On the other hand, we found that the normalizable modes of the chiral bifundamental mesons are very well localized in the infrared. Indeed, at large r,

$$\mathcal{U}\big(\mu;\nu;2Mr^2\big) \sim r^{-\mu}, \tag{4.98}$$

so that, when $c_1 = 0$ in (4.90), the bifundamental modes exhibit a Gaussian localization,

$$f_\pm \propto \mathrm{e}^{-Mr^2}, \tag{4.99}$$

where we have again omitted power law factors and have chosen $M > 0$. Although similar Gaussian peaks appear in flat space (see e.g. [175]), this feature in warped space has the potential to provide a rich playground for model-building. In general, the lack of knowledge of the metric and of related fields often stymies detailed model-building in string compactifications. However, the metrics for infinite families of non-compact (and singular) Calabi-Yau cones are known explicitly. These cones can be used to construct strongly warped geometries that can be attached to compact spaces — see for example the discussion in [131]. Attachment to a compactification modifies the solution in the cone region, by introducing sources for irrelevant perturbations, but these effects can be incorporated systematically, as in [151, 64]. One can therefore build a local model on D3-branes at the apex of the

cone, but also take into account bulk effects, including supersymmetry breaking and moduli stabilization. Constructions in this corner of the landscape are limited to some degree by the possible singularities at the apex. An alternative, toward which the present work is a modest advance, is to consider model-building on intersecting magnetized D7-branes. Although the D7-branes will stretch beyond the warped region into the bulk,[14] we have demonstrated that at least some bifundamental modes are well localized in the infrared. This allows for a combination of the richness of model-building with intersecting D7-branes and the power of local model-building in warped geometries. Although we limited our particular analysis to $AdS_5 \times S^5$, the qualitative result should extend to more general cones and their deformations (though the details, of course, become much more complex).

This localization also implies that although correlation functions are difficult to describe, the mass spectrum of mesons can in principle be calculated with reliable numerical techniques. When the D7-branes move away from the center of AdS_5, the spectrum of mesons becomes gapped even though, in the quenched approximation, the glueball spectrum is continuous [126]. A standard method of finding the meson mass spectrum in the gapped case is to calculate the correlation functions and check for the appearance of poles. However, a practical alternative is to find those solutions that satisfy appropriate infrared boundary conditions and are normalizable in the ultraviolet (see, for example, [126, 186]). Because the equation of motion constitutes a Sturm-Liouville problem, this alternative approach leads to a discrete spectrum, and since the solutions are expected to be exponentially convergent, the resulting spectrum would be reliable. On the other hand, once the spectrum becomes gapped the radial and angular parts of

[14] Indeed, the consistency of embeddings in global models will provide constraints on which models can be built.

the equation of motion no longer separate, even in the unmagnetized case (4.58). This is a significant complication, and so we leave this analysis to future work.

Yet another possibility is to consider alternative magnetizations. The magnetization that we analyzed in this note is the simplest unsourced magnetic flux that is possible in our construction, and other unsourced magnetic fluxes would enhance the bifundamental wavefunction localization that we found, while intensifying the problem of ultraviolet sensitivity. Magnetic flux that is itself localized in the infrared, and produces only normalizable perturbations to the geometry, would require a local source. In particular, it was pointed out in [161] and explicitly shown in [162] that the addition of anti-D3-branes to warped flux backgrounds provides an infrared-localized magnetization. Although the resulting magnetization has a gauge structure that differs from (4.32) — specifically, the induced magnetization is proportional to the identity — this remains an intriguing possibility for future work.

It is a pleasure to thank F. Marchesano and G. Shiu for useful discussions of related topics. This work was supported by the NSF under grant PHY-0757868.

4.A Conventions for Fermions

In this appendix we summarize our conventions for fermions, many of which follow from [187]. We work with a Weyl basis for the $\mathrm{SO}(9,1)$ Γ-matrices and make use of the decomposition $\mathrm{SO}(9,1) \to \mathrm{SO}(3,1) \times \mathrm{SO}(6)$. For $\mathrm{SO}(3,1)$ we take

$$\gamma^0 = \begin{pmatrix} & \mathbb{1}_2 \\ -\mathbb{1}_2 & \end{pmatrix}, \quad \gamma^{i=1,2,3} = \begin{pmatrix} & \sigma^i \\ \sigma^i & \end{pmatrix}, \tag{4.100}$$

in which σ^i are the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.101}$$

For $\mathrm{SO}(2k+1,1)$, we take the chirality matrix to be

$$\gamma_{(2k+2)} = \mathrm{i}^{-k} \mathrm{dv}\not{\mathrm{o}}\mathrm{l}_{\mathbb{R}^{2k+1,1}}, \tag{4.102}$$

where dvol_M is the volume element on M

$$\mathrm{dvol}_M = \frac{1}{d!} \epsilon_{M_1 \cdots M_d} \mathrm{d}x^{M_1} \wedge \cdots \wedge \mathrm{d}x^{M_d}, \tag{4.103}$$

in which $\epsilon_{01\cdots(d-1)} = \sqrt{-\det g}$. For $\mathbb{R}^{3,1}$,

$$\gamma_{(4)} = -\mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} \mathbb{1}_2 & \\ & -\mathbb{1}_2 \end{pmatrix}. \tag{4.104}$$

The 4d Majorana matrix is

$$\beta_4 = \gamma_{(4)}\gamma^2 = \begin{pmatrix} & -\sigma^2 \\ \sigma^2 & \end{pmatrix}. \tag{4.105}$$

For $\mathrm{SO}(6)$, we define

$$\begin{aligned} \tilde{\gamma}^4 &= \sigma^1 \otimes \mathbb{1}_2 \otimes \mathbb{1}_2, & \tilde{\gamma}^7 &= \sigma^2 \otimes \mathbb{1}_2 \otimes \mathbb{1}_2, \\ \tilde{\gamma}^5 &= \sigma^3 \otimes \sigma^1 \otimes \mathbb{1}_2, & \tilde{\gamma}^8 &= \sigma^3 \otimes \sigma^2 \otimes \mathbb{1}_2, \\ \tilde{\gamma}^6 &= \sigma^3 \otimes \sigma^3 \otimes \sigma^1, & \tilde{\gamma}^9 &= \sigma^3 \otimes \sigma^3 \otimes \sigma^2. \end{aligned}$$

For $\mathrm{SO}(2k+4)$, the chirality operator is

$$\gamma_{(2k+4)} = \mathrm{i}^{-k} \mathrm{dv}\not{\mathrm{o}}\mathrm{l}_{\mathbb{R}^{2k+4}}, \tag{4.106}$$

and so

$$\tilde{\gamma}_{(6)} = -\mathrm{i}\,\tilde{\gamma}^1 \cdots \tilde{\gamma}^6 = \sigma^3 \otimes \sigma^3 \otimes \sigma^3. \tag{4.107}$$

The Majorana matrix is

$$\tilde{\beta}_6 = \tilde{\gamma}^7\tilde{\gamma}^8\tilde{\gamma}^9 = \sigma^2 \otimes \mathrm{i}\,\sigma^1 \otimes \sigma^2. \tag{4.108}$$

We will make use of a complex structure

$$z^I = x^{3+I} + \mathrm{i}\, x^{4+I}. \tag{4.109}$$

Defining

$$\sigma^{\pm} = \frac{1}{2}\big(\sigma^1 \pm \mathrm{i}\,\sigma^2\big), \tag{4.110}$$

we have

$$\begin{aligned}
\tilde{\gamma}^1 =& 2\,\sigma^+ \otimes \mathbb{1}_2 \otimes \mathbb{1}_2, & \tilde{\gamma}^{\bar{1}} =& 2\,\sigma^- \otimes \mathbb{1}_2 \otimes \mathbb{1}_2,\\
\tilde{\gamma}^2 =& 2\,\sigma^3 \otimes \sigma^+ \otimes \mathbb{1}_2, & \tilde{\gamma}^{\bar{2}} =& 2\,\sigma^3 \otimes \sigma^- \otimes \mathbb{1}_2,\\
\tilde{\gamma}^3 =& 2\,\sigma^3 \otimes \sigma^3 \otimes \sigma^+, & \tilde{\gamma}^{\bar{3}} =& 2\,\sigma^3 \otimes \sigma^3 \otimes \sigma^-.
\end{aligned}$$

We can construct a basis of positive chirality spinors by first defining

$$\eta_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \eta_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{4.111}$$

The positive chirality spinors are then

$$\eta_0 = \eta_{+++}, \quad \eta_1 = \eta_{+--}, \quad \eta_2 = \eta_{-+-}, \quad \eta_3 = \eta_{--+}, \tag{4.112}$$

in which

$$\eta_{\epsilon_1\epsilon_2\epsilon_3} = \eta_{\epsilon_1} \otimes \eta_{\epsilon_2} \otimes \eta_{\epsilon_3}. \tag{4.113}$$

Note that $\sigma^{\pm}\eta_{\pm} = 0$, so that η_{+++} is annihilated by all contravariant holomorphic $\tilde{\gamma}$-matrices.

Finally, we construct the $\mathrm{SO}(9,1)$ Γ-matrices by

$$\hat{\Gamma}^{\mu} = \gamma^{\mu} \otimes \mathbb{1}_8, \quad \hat{\Gamma}^{m} = \gamma_{(6)} \otimes \tilde{\gamma}^{m}. \tag{4.114}$$

The chirality and Majorana matrices are

$$\begin{aligned}
\hat{\Gamma}_{(10)} =& \hat{\Gamma}^0\hat{\Gamma}^1\cdots\hat{\Gamma}^9 = -\gamma_{(4)} \otimes \tilde{\gamma}_{(6)},\\
\hat{B}_{10} =& \hat{\Gamma}^2\hat{\Gamma}^7\hat{\Gamma}^8\hat{\Gamma}^9 = -\beta_4 \otimes \tilde{\beta}_6.
\end{aligned} \tag{4.115}$$

We will make use of 32-component Majorana-Weyl spinors satisfying

$$\hat{\Gamma}_{(10)}\theta = -\theta, \quad \hat{B}_{10}\theta = \theta^*. \tag{4.116}$$

An example of such a spinor is

$$\theta = \begin{pmatrix} \xi \\ 0 \end{pmatrix} \otimes \eta - \begin{pmatrix} 0 \\ \sigma^2\xi^* \end{pmatrix} \otimes \tilde{\beta}_6\eta^*, \tag{4.117}$$

where $\tilde{\gamma}_{(6)}\eta = +\eta$.

We will also make use of double spinors built from pairs of 10d Majorana-Weyl spinors

$$\Theta = \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix}, \tag{4.118}$$

where both θ^1 and θ^2 satisfy (4.116). $\hat{\Gamma}$-matrices act on double spinors as

$$\hat{\Gamma}^M\Theta = \begin{pmatrix} \hat{\Gamma}^M\theta^1 \\ \hat{\Gamma}^M\theta^2 \end{pmatrix}, \tag{4.119}$$

while explicit Pauli matrices act to mix the elements of the double spinor. For example,

$$\sigma^1 \begin{pmatrix} \theta^1 \\ \theta^2 \end{pmatrix} = \begin{pmatrix} \theta^2 \\ \theta^1 \end{pmatrix}. \tag{4.120}$$

4.B Hyperspherical Harmonics

In this appendix we review a few properties of the hyperspherical harmonics on S^3. A useful parametrization of S^3 is via the usual embedding of S^3 into $\mathbb{R}^4$, $\zeta^i\zeta^i = 1$, where $\zeta^1 \ldots \zeta^4$ are coordinates on $\mathbb{R}^4$. We take (as in, for example, [188])

$$\begin{aligned} \zeta^1 &= r\cos\beta\cos\phi_1, & \zeta^2 &= r\sin\beta\cos\phi_2, \\ \zeta^3 &= r\cos\beta\sin\phi_1, & \zeta^4 &= r\sin\beta\sin\phi_2, \end{aligned} \tag{4.121}$$

with $\beta \in \left[0, \frac{\pi}{2}\right]$ and $\phi_a \in [0, 2\pi)$. The induced metric on S^3 is

$$\mathrm{d}s^2_{S^3} = \breve{g}_{\theta\varphi}\mathrm{d}y^\theta \mathrm{d}y^\varphi = \mathrm{d}\beta^2 + \cos^2\beta\,\mathrm{d}\phi_1^2 + \sin^2\beta\,\mathrm{d}\phi_2^2. \tag{4.122}$$

The volume of S^3 is

$$\mathcal{V}_{S^3} = \int_0^{2\pi} \mathrm{d}\phi_1 \int_0^{2\pi} \mathrm{d}\phi_2 \int_0^{\pi/2} \mathrm{d}\beta \sin\beta\cos\beta = 2\pi^2. \tag{4.123}$$

The scalar spherical harmonics satisfy the eigenvalue problem

$$\breve{\nabla}^2\mathcal{Y} = \frac{\partial^2\mathcal{Y}}{\partial\beta^2} + \big(\cot\beta - \tan\beta\big)\frac{\partial\mathcal{Y}}{\partial\beta} + \frac{1}{\cos^2\beta}\frac{\partial^2\mathcal{Y}}{\partial\phi_1^2} + \frac{1}{\sin^2\beta}\frac{\partial^2\mathcal{Y}}{\partial\phi_2^2} = -\lambda\mathcal{Y}. \tag{4.124}$$

Taking the ansatz

$$\mathcal{Y} = \mathrm{e}^{\mathrm{i}(m_1\phi_1 + m_2\phi_2)} y\big(\cos 2\beta\big) \tag{4.125}$$

gives

$$0 = 4\big(1-x^2\big)y'' - 8xy' - \frac{2m_1^2}{1+x}y - \frac{2m_2^2}{1-x}y + \lambda y, \tag{4.126}$$

in which $x = \cos 2\beta$. Imposing Neumann conditions so that a zero mode is admitted, the solutions are given in terms of Jacobi polynomials $P_r^{(a,b)}$,

$$\begin{aligned}&\mathcal{Y}_{\ell,m_1,m_2}\big(\beta,\phi_1,\phi_2\big)\\ &= c_{\ell,m_1,m_2}\mathrm{e}^{\mathrm{i}(m_1\phi_1+m_2\phi_2)}\big(1+\cos 2\beta\big)^{m_1/2}\big(1-\cos 2\beta\big)^{m_2/2}P_r^{(m_2,m_1)}\big(\cos 2\beta\big),\end{aligned} \tag{4.127}$$

in which $r = \frac{1}{2}\left(\ell - m_1 - m_2\right)$. For these to be non-vanishing regular solutions, r must be an integer and

$$0 \leq \big|m_1\big| + \big|m_2\big| \leq \ell. \tag{4.128}$$

These solutions satisfy

$$\breve{\nabla}^2\mathcal{Y} = -\ell\left(\ell+2\right)\mathcal{Y}, \tag{4.129}$$

and the condition (4.128) gives the expected degeneracy of $\left(\ell+1\right)^2$ (see, for example, [181]).

The Jacobi polynomials are orthogonal in the sense that

$$\int_{-1}^{1} \mathrm{d}x \left(1-x\right)^{a}\left(1+x\right)^{b} P_{r}^{(a,b)} P_{s}^{(a,b)} = \frac{2^{a+b+1}}{2r+a+b+1} \frac{(a+r)!\ (b+r)!}{r!\ (a+b+r)!} \delta_{rs}. \quad (4.130)$$

Therefore the normalization condition

$$\int \mathrm{dvol}_{S^3}\, \mathcal{Y}^{*}_{\ell,m_1,m_2} \mathcal{Y}_{\ell',m_1',m_2'} = \delta_{\ell'\ell} \delta_{m_1' m_1} \delta_{m_2' m_2} \quad (4.131)$$

is satisfied by taking

$$c_{\ell,m_1,m_2} = \frac{1}{\pi} \sqrt{\frac{\ell+1}{2^{m_1+m_2+1}} \frac{\left[\frac{1}{2}\left(\ell+m_1+m_2\right)\right]!\left[\frac{1}{2}\left(\ell-m_1-m_2\right)\right]!}{\left[\frac{1}{2}\left(\ell+m_1-m_2\right)\right]!\left[\frac{1}{2}\left(\ell-m_1+m_2\right)\right]!}}. \quad (4.132)$$

The Jacobi polynomials satisfy the useful recursion relationship

$$\begin{aligned} x\, P_{r}^{(a,b)}(x) = {} & \frac{2\left(a+r\right)\left(b+r\right)}{\left(a+b+2r\right)\left(a+b+2r+1\right)} P_{r-1}^{(a,b)}(x) \\ & + \frac{2\left(r+1\right)\left(a+b+r+1\right)}{\left(a+b+2r+1\right)\left(a+b+2r+2\right)} P_{r+1}^{(a,b)}(x) \\ & + \frac{b^2-a^2}{\left(a+b+2r\right)\left(a+b+2r+2\right)} P_{r}^{(a,b)}(x)\,. \end{aligned} \quad (4.133)$$

BIBLIOGRAPHY

[1] J. R. Espinosa, C. Grojean, G. Panico, A. Pomarol, O. Pujols and G. Servant, "Cosmological Higgs-Axion Interplay for a Naturally Small Electroweak Scale," *Phys. Rev. Lett.* **115** (2015) 251803, [1506.09217].

[2] Thornton, S.T. and Marion, J.B., *Classical Dynamics of Particles and Systems*. Brooks/Cole, 2004.

[3] M. Srednicki, *Quantum field theory*. Cambridge University Press, 2007.

[4] T. Hollowood, *Renormalization Group and Fixed Points: in Quantum Field Theory*. SpringerBriefs in Physics. Springer Berlin Heidelberg, 2013.

[5] K. G. Wilson and J. Kogut, "The renormalization group and the expansion," *Physics Reports* **12** (1974) 75–199.

[6] J. Polchinski, "Renormalization and Effective Lagrangians," *Nucl. Phys.* **B231** (1984) 269–295.

[7] M. Dine, "Naturalness Under Stress," *Ann. Rev. Nucl. Part. Sci.* **65** (2015) 43–62, [1501.01035].

[8] D. Baumann and L. McAllister, *Inflation and String Theory*. Cambridge University Press, 2015.

[9] R. Kallosh, A. D. Linde, D. A. Linde and L. Susskind, "Gravity and global symmetries," *Phys. Rev.* **D52** (1995) 912–935, [hep-th/9502069].

[10] T. Banks and N. Seiberg, "Symmetries and Strings in Field Theory and Gravity," *Phys. Rev.* **D83** (2011) 084019, [1011.5120].

[11] T. Banks and L. J. Dixon, "Constraints on String Vacua with Space-Time Supersymmetry," *Nucl. Phys.* **B307** (1988) 93–108.

[12] T. Banks, M. Johnson and A. Shomer, "A Note on Gauge Theories Coupled to Gravity," *JHEP* **09** (2006) 049, [hep-th/0606277].

[13] L. Susskind, "Trouble for remnants," hep-th/9501106.

[14] CMB-S4 collaboration, K. N. Abazajian et al., "CMB-S4 Science Book, First Edition," 1610.02743.

[15] L. McAllister, P. Schwaller, G. Servant, J. Stout and A. Westphal, "Runaway Relaxion Monodromy," 1610.05320.

[16] P. W. Graham, D. E. Kaplan and S. Rajendran, "Cosmological Relaxation of the Electroweak Scale," *Phys. Rev. Lett.* **115** (2015) 221801, [1504.07551].

[17] A. Hook and G. Marques-Tavares, "Relaxation from particle production," *JHEP* **12** (2016) 101, [1607.01786].

[18] S. P. Patil and P. Schwaller, "Relaxing the Electroweak Scale: the Role of Broken dS Symmetry," *JHEP* **02** (2016) 077, [1507.08649].

[19] S. Di Chiara, K. Kannike, L. Marzola, A. Racioppi, M. Raidal and C. Spethmann, "Relaxion Cosmology and the Price of Fine-Tuning," *Phys. Rev.* **D93** (2016) 103527, [1511.02858].

[20] T. Kobayashi, O. Seto, T. Shimomura and Y. Urakawa, "Relaxion window," 1605.06908.

[21] R. S. Gupta, Z. Komargodski, G. Perez and L. Ubaldi, "Is the Relaxion an Axion?," *JHEP* **02** (2016) 166, [1509.00047].

[22] L. E. Ibanez, M. Montero, A. Uranga and I. Valenzuela, "Relaxion Monodromy and the Weak Gravity Conjecture," *JHEP* **04** (2016) 020, [1512.00025].

[23] A. Hebecker, F. Rompineve and A. Westphal, "Axion Monodromy and the Weak Gravity Conjecture," *JHEP* **04** (2016) 157, [1512.03768].

[24] K. Furuuchi, "Excursions through KK modes," *JCAP* **1607** (2016) 008, [1512.04684].

[25] N. Kaloper and A. Lawrence, "London equation for monodromy inflation," *Phys. Rev.* **D95** (2017) 063526, [1607.06105].

[26] N. Kaloper and L. Sorbo, "A Natural Framework for Chaotic Inflation," *Phys. Rev. Lett.* **102** (2009) 121301, [0811.1989].

[27] N. Kaloper, A. Lawrence and L. Sorbo, "An Ignoble Approach to Large Field Inflation," *JCAP* **1103** (2011) 023, [1101.0026].

[28] N. Kaloper, M. Kleban, A. Lawrence and M. S. Sloth, "Large Field Inflation and Gravitational Entropy," 1511.05119.

[29] K. Choi and S. H. Im, "Constraints on Relaxion Windows," *JHEP* **12** (2016) 093, [1610.00680].

[30] B. Batell, G. F. Giudice and M. McCullough, "Natural Heavy Supersymmetry," *JHEP* **12** (2015) 162, [1509.00834].

[31] J. L. Evans, T. Gherghetta, N. Nagata and Z. Thomas, "Naturalizing Supersymmetry with a Two-Field Relaxion Mechanism," *JHEP* **09** (2016) 150, [1602.04812].

[32] L. McAllister, E. Silverstein and A. Westphal, "Gravity Waves and Linear Inflation from Axion Monodromy," *Phys. Rev.* **D82** (2010) 046003, [0808.0706].

[33] E. Silverstein and A. Westphal, "Monodromy in the CMB: Gravity Waves and String Inflation," *Phys. Rev.* **D78** (2008) 106003, [0803.3085].

[34] R. Flauger, L. McAllister, E. Pajer, A. Westphal and G. Xu, "Oscillations in the CMB from Axion Monodromy Inflation," *JCAP* **1006** (2010) 009, [0907.2916].

[35] M. Aganagic, C. Beem, J. Seo and C. Vafa, "Geometrically Induced Metastability and Holography," *Nucl. Phys.* **B789** (2008) 382–412, [hep-th/0610249].

[36] A. Klemm, B. Lian, S. S. Roan and S.-T. Yau, "Calabi-Yau fourfolds for M theory and F theory compactifications," *Nucl. Phys.* **B518** (1998) 515–574, [hep-th/9701023].

[37] M. Berg, D. Marsh, L. McAllister and E. Pajer, "Sequestering in String Compactifications," *JHEP* **06** (2011) 134, [1012.1858].

[38] C. P. Burgess, M. Majumdar, D. Nolte, F. Quevedo, G. Rajesh and R.-J. Zhang, "The Inflationary brane anti-brane universe," *JHEP* **07** (2001) 047, [hep-th/0105204].

[39] J. P. Conlon, "Brane-Antibrane Backreaction in Axion Monodromy Inflation," *JCAP* **1201** (2012) 033, [1110.6454].

[40] S. Kachru, J. Pearson and H. L. Verlinde, "Brane / flux annihilation and the string dual of a nonsupersymmetric field theory," *JHEP* **06** (2002) 021, [hep-th/0112197].

[41] A. Saltman and E. Silverstein, "A New handle on de Sitter compactifications," *JHEP* **01** (2006) 139, [hep-th/0411271].

[42] L. McAllister, E. Silverstein, A. Westphal and T. Wrase, "The Powers of Monodromy," *JHEP* **09** (2014) 123, [1405.3652].

[43] R. Flauger, M. Mirbabayi, L. Senatore and E. Silverstein, "Productive Interactions: heavy particles and non-Gaussianity," 1606.00513.

[44] J. E. Kim, H. P. Nilles and M. Peloso, "Completing natural inflation," *JCAP* **0501** (2005) 005, [hep-ph/0409138].

[45] T. C. Bachlechner, M. Dias, J. Frazer and L. McAllister, "Chaotic inflation with kinetic alignment of axion fields," *Phys. Rev.* **D91** (2015) 023520, [1404.7496].

[46] T. C. Bachlechner, C. Long and L. McAllister, "Planckian Axions in String Theory," *JHEP* **12** (2015) 042, [1412.1093].

[47] K. Choi and S. H. Im, "Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry," *JHEP* **01** (2016) 149, [1511.00132].

[48] D. E. Kaplan and R. Rattazzi, "Large field excursions and approximate discrete symmetries from a clockwork axion," *Phys. Rev.* **D93** (2016) 085007, [1511.01827].

[49] N. Fonseca, L. de Lima, C. S. Machado and R. D. Matheus, "Large field excursions from a few site relaxion model," *Phys. Rev.* **D94** (2016) 015010, [1601.07183].

[50] T. Flacke, C. Frugiuele, E. Fuchs, R. S. Gupta and G. Perez, "Phenomenology of relaxion-Higgs mixing," 1610.02025.

[51] C. Long, L. McAllister and J. Stout, "Systematics of Axion Inflation in Calabi-Yau Hypersurfaces," *JHEP* **02** (2017) 014, [1603.01259].

[52] K. Choi, H. Kim and S. Yun, "Natural inflation with multiple sub-Planckian axions," *Phys. Rev.* **D90** (2014) 023545, [1404.6209].

[53] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, "The String landscape, black holes and gravity as the weakest force," *JHEP* **06** (2007) 060, [hep-th/0601001].

[54] C. Cheung and G. N. Remmen, "Naturalness and the Weak Gravity Conjecture," *Phys. Rev. Lett.* **113** (2014) 051601, [1402.2287].

[55] A. de la Fuente, P. Saraswat and R. Sundrum, "Natural Inflation and Quantum Gravity," *Phys. Rev. Lett.* **114** (2015) 151303, [1412.3457].

[56] T. Rudelius, "Constraints on Axion Inflation from the Weak Gravity Conjecture," *JCAP* **1509** (2015) 020, [1503.00795].

[57] M. Montero, A. M. Uranga and I. Valenzuela, "Transplanckian axions!?," *JHEP* **08** (2015) 032, [1503.03886].

[58] J. Brown, W. Cottrell, G. Shiu and P. Soler, "Fencing in the Swampland: Quantum Gravity Constraints on Large Field Inflation," *JHEP* **10** (2015) 023, [1503.04783].

[59] T. C. Bachlechner, C. Long and L. McAllister, "Planckian Axions and the Weak Gravity Conjecture," *JHEP* **01** (2016) 091, [1503.07853].

[60] J. Brown, W. Cottrell, G. Shiu and P. Soler, "On Axionic Field Ranges, Loopholes and the Weak Gravity Conjecture," *JHEP* **04** (2016) 017, [1504.00659].

[61] B. Heidenreich, M. Reece and T. Rudelius, "Weak Gravity Strongly Constrains Large-Field Axion Inflation," *JHEP* **12** (2015) 108, [1506.03447].

[62] B. Heidenreich, M. Reece and T. Rudelius, "Sharpening the Weak Gravity Conjecture with Dimensional Reduction," *JHEP* **02** (2016) 140, [1509.06374].

[63] P. Saraswat, "Weak gravity conjecture and effective field theory," *Phys. Rev.* **D95** (2017) 025013, [1608.06951].

[64] S. Gandhi, L. McAllister and S. Sjors, "A Toolkit for Perturbing Flux Compactifications," *JHEP* **12** (2011) 053, [1106.0002].

[65] K. Freese, J. A. Frieman and A. V. Olinto, "Natural inflation with pseudo - Nambu-Goldstone bosons," *Phys. Rev. Lett.* **65** (1990) 3233–3236.

[66] F. Marchesano, G. Shiu and A. M. Uranga, "F-term Axion Monodromy Inflation," *JHEP* **09** (2014) 184, [1404.3040].

[67] T. Banks, M. Dine, P. J. Fox and E. Gorbatov, "On the possibility of large axion decay constants," *JCAP* **0306** (2003) 001, [hep-th/0303252].

[68] P. Svrcek and E. Witten, "Axions In String Theory," *JHEP* **06** (2006) 051, [hep-th/0605206].

[69] A. R. Liddle, A. Mazumdar and F. E. Schunck, "Assisted inflation," *Phys. Rev.* **D58** (1998) 061301, [astro-ph/9804177].

[70] T. Higaki and F. Takahashi, "Natural and Multi-Natural Inflation in Axion Landscape," *JHEP* **07** (2014) 074, [1404.6923].

[71] M. Czerny, T. Higaki and F. Takahashi, "Multi-Natural Inflation in Supergravity," *JHEP* **05** (2014) 144, [1403.0410].

[72] S. H. H. Tye and S. S. C. Wong, "Helical Inflation and Cosmic Strings," 1404.6988.

[73] R. Kappl, S. Krippendorf and H. P. Nilles, "Aligned Natural Inflation: Monodromies of two Axions," *Phys. Lett.* **B737** (2014) 124–128, [1404.7127].

[74] I. Ben-Dayan, F. G. Pedro and A. Westphal, "Hierarchical Axion Inflation," *Phys. Rev. Lett.* **113** (2014) 261301, [1404.7773].

[75] X. Gao, T. Li and P. Shukla, "Combining Universal and Odd RR Axions for Aligned Natural Inflation," *JCAP* **1410** (2014) 048, [1406.0341].

[76] I. Ben-Dayan, F. G. Pedro and A. Westphal, "Towards Natural Inflation in String Theory," *Phys. Rev.* **D92** (2015) 023515, [1407.2562].

[77] H. Abe, T. Kobayashi and H. Otsuka, "Towards natural inflation from

weakly coupled heterotic string theory," *PTEP* **2015** (2014) 063E02, [1409.8436].

[78] T. Ali, S. S. Haque and V. Jejjala, "Natural Inflation from Near Alignment in Heterotic String Theory," *Phys. Rev.* **D91** (2015) 083516, [1410.4660].

[79] C. Burgess and D. Roest, "Inflation by Alignment," *JCAP* **1506** (2015) 012, [1412.1614].

[80] I. Garca-Etxebarria, T. W. Grimm and I. Valenzuela, "Special Points of Inflation in Flux Compactifications," *Nucl. Phys.* **B899** (2015) 414–443, [1412.5537].

[81] G. Shiu, W. Staessens and F. Ye, "Large Field Inflation from Axion Mixing," *JHEP* **06** (2015) 026, [1503.02965].

[82] E. Palti, "On Natural Inflation and Moduli Stabilisation in String Theory," *JHEP* **10** (2015) 188, [1508.00009].

[83] R. Kappl, H. P. Nilles and M. W. Winkler, "Modulated Natural Inflation," *Phys. Lett.* **B753** (2016) 653–659, [1511.05560].

[84] Planck collaboration, P. A. R. Ade et al., "Planck 2015 results. XX. Constraints on inflation," *Astron. Astrophys.* **594** (2016) A20, [1502.02114].

[85] M. Kreuzer and H. Skarke, "Complete classification of reflexive polyhedra in four-dimensions," *Adv. Theor. Math. Phys.* **4** (2002) 1209–1230, [hep-th/0002240].

[86] R. Blumenhagen, M. Cvetic, R. Richter and T. Weigand, "Lifting D-Instanton Zero Modes by Recombination and Background Fluxes," *JHEP* **10** (2007) 098, [0708.0403].

[87] R. Donagi and M. Wijnholt, "MSW Instantons," *JHEP* **06** (2013) 050, [1005.5391].

[88] T. W. Grimm, M. Kerstan, E. Palti and T. Weigand, "On Fluxed Instantons and Moduli Stabilisation in IIB Orientifolds and F-theory," *Phys. Rev.* **D84** (2011) 066001, [1105.3193].

[89] M. Bianchi, A. Collinucci and L. Martucci, "Magnetized E3-brane instantons in F-theory," *JHEP* **12** (2011) 045, [1107.3732].

[90] M. Cvetic, R. Donagi, J. Halverson and J. Marsano, "On Seven-Brane Dependent Instanton Prefactors in F-theory," *JHEP* **11** (2012) 004, [1209.4906].

[91] M. Bianchi, G. Inverso and L. Martucci, "Brane instantons and fluxes in F-theory," *JHEP* **07** (2013) 037, [1212.0024].

[92] X. Gao and P. Shukla, "On Classifying the Divisor Involutions in Calabi-Yau Threefolds," *JHEP* **11** (2013) 170, [1307.1139].

[93] T. W. Grimm, "Axion Inflation in F-theory," *Phys. Lett.* **B739** (2014) 201–208, [1404.4268].

[94] J. P. Conlon and S. Krippendorf, "Axion decay constants away from the lamppost," *JHEP* **04** (2016) 085, [1601.00647].

[95] F. Denef, M. R. Douglas, B. Florea, A. Grassi and S. Kachru, "Fixing all moduli in a simple f-theory compactification," *Adv. Theor. Math. Phys.* **9** (2005) 861–929, [hep-th/0503124].

[96] E. Witten, "Nonperturbative superpotentials in string theory," *Nucl. Phys.* **B474** (1996) 343–360, [hep-th/9604030].

[97] P. A. Griffiths and J. Harris, *Principles of Algebraic Geometry*. Wiley, New York, 1978.

[98] A. Grassi, "Divisors on elliptic Calabi-Yau four folds and the superpotential in F theory. 1.," *J. Geom. Phys.* **28** (1998) 289–319.

[99] C. Long, L. McAllister and P. McGuirk, "Aligned Natural Inflation in String Theory," *Phys. Rev.* **D90** (2014) 023501, [1404.7852].

[100] V. V. Batyrev, "Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties," *J. Alg. Geom.* **3** (1994) 493–545, [alg-geom/9310003].

[101] D. Cox, J. Little and H. Schenck, *Toric Varieties*. Graduate studies in mathematics. American Mathematical Soc., 2011.

[102] R. Altman, J. Gray, Y.-H. He, V. Jejjala and B. D. Nelson, "A Calabi-Yau Database: Threefolds Constructed from the Kreuzer-Skarke List," *JHEP* **02** (2015) 158, [1411.1418].

[103] The Sage Developers, *Sage Mathematics Software (Version 6.10)*, 2016.

[104] J. Rambau, "TOPCOM: Triangulations of point configurations and oriented matroids," in *Mathematical Software—ICMS 2002* (A. M. Cohen, X.-S. Gao and N. Takayama, eds.), pp. 330–340, World Scientific, 2002.

[105] M. Kreuzer and H. Skarke, "PALP: A Package for analyzing lattice polytopes with applications to toric geometry," *Comput. Phys. Commun.* **157** (2004) 87–106, [math/0204356].

[106] A. P. Braun and N.-O. Walliser, "A New offspring of PALP," 1106.4529.

[107] C. Long, L. McAllister and P. McGuirk, "Heavy Tails in Calabi-Yau Moduli Spaces," *JHEP* **10** (2014) 187, [1407.0709].

[108] A. P. Braun and T. Watari, "The Vertical, the Horizontal and the Rest: anatomy of the middle cohomology of Calabi-Yau fourfolds and F-theory applications," *JHEP* **01** (2015) 047, [1408.6167].

[109] A. P. Braun, "Tops as Building Blocks for G2 Manifolds," 1602.03521.

[110] V. I. Danilov and A. G. Khovanskii, "Newton Polyhedra and an Algorithm for Computing Hodge-Deligne Numbers," *Mathematics of the USSR-Izvestiya* **29** (1987) 279.

[111] R. Blumenhagen, B. Jurke, T. Rahn and H. Roschy, "Cohomology of Line Bundles: A Computational Algorithm," *J. Math. Phys.* **51** (2010) 103525, [1003.5217].

[112] R. Blumenhagen, B. Jurke, T. Rahn and H. Roschy, "Cohomology of Line Bundles: Applications," *J. Math. Phys.* **53** (2012) 012302, [1010.3717].

[113] C. Long, L. McAllister and J. Stout, *to appear*.

[114] S. Dimopoulos, S. Kachru, J. McGreevy and J. G. Wacker, "N-flation," *JCAP* **0808** (2008) 003, [hep-th/0507205].

[115] V. Khrulkov, C. Long, L. McAllister, M. Stillman and B. Sung, *work in progress.*

[116] F. Denef, M. R. Douglas and B. Florea, "Building a better racetrack," *JHEP* **06** (2004) 034, [hep-th/0404257].

[117] T. Rudelius, "On the Possibility of Large Axion Moduli Spaces," *JCAP* **1504** (2015) 049, [1409.5793].

[118] T. W. Grimm, "Axion inflation in type II string theory," *Phys. Rev.* **D77** (2008) 126007, [0710.3883].

[119] L. McAllister, P. McGuirk and J. Stout, "On Chiral Mesons in AdS/CFT," *JHEP* **02** (2014) 018, [1311.2577].

[120] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Int. J. Theor. Phys.* **38** (1999) 1113–1133, [hep-th/9711200].

[121] E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2** (1998) 253–291, [hep-th/9802150].

[122] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys. Lett.* **B428** (1998) 105–114, [hep-th/9802109].

[123] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rept.* **323** (2000) 183–386, [hep-th/9905111].

[124] O. Aharony, A. Fayyazuddin and J. M. Maldacena, "The Large N limit of N=2, N=1 field theories from three-branes in F theory," *JHEP* **07** (1998) 013, [hep-th/9806159].

[125] A. Karch and E. Katz, "Adding flavor to AdS / CFT," *JHEP* **06** (2002) 043, [hep-th/0205236].

[126] M. Kruczenski, D. Mateos, R. C. Myers and D. J. Winters, "Meson spectroscopy in AdS / CFT with flavor," *JHEP* **07** (2003) 049, [hep-th/0304032].

[127] C. Nunez, A. Paredes and A. V. Ramallo, "Unquenched Flavor in the Gauge/Gravity Correspondence," *Adv. High Energy Phys.* **2010** (2010) 196714, [1002.1088].

[128] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373, [hep-ph/9905221].

[129] H. L. Verlinde, "Holography and compactification," *Nucl. Phys.* **B580** (2000) 264–274, [hep-th/9906182].

[130] B. R. Greene, K. Schalm and G. Shiu, "Warped compactifications in M and F theory," *Nucl. Phys.* **B584** (2000) 480–508, [hep-th/0004103].

[131] S. B. Giddings, S. Kachru and J. Polchinski, "Hierarchies from fluxes in string compactifications," *Phys. Rev.* **D66** (2002) 106006, [hep-th/0105097].

[132] H. Davoudiasl, J. L. Hewett and T. G. Rizzo, "Bulk gauge fields in the Randall-Sundrum model," *Phys. Lett.* **B473** (2000) 43–49, [hep-ph/9911262].

[133] A. Pomarol, "Gauge bosons in a five-dimensional theory with localized gravity," *Phys. Lett.* **B486** (2000) 153–157, [hep-ph/9911294].

[134] S. Chang, J. Hisano, H. Nakano, N. Okada and M. Yamaguchi, "Bulk standard model in the Randall-Sundrum background," *Phys. Rev.* **D62** (2000) 084025, [hep-ph/9912498].

[135] Y. Grossman and M. Neubert, "Neutrino masses and mixings in nonfactorizable geometry," *Phys. Lett.* **B474** (2000) 361–371, [hep-ph/9912408].

[136] T. Gherghetta and A. Pomarol, "Bulk fields and supersymmetry in a slice of AdS," *Nucl. Phys.* **B586** (2000) 141–162, [hep-ph/0003129].

[137] T. Gherghetta and J. Giedt, "Bulk fields in AdS(5) from probe D7 branes," *Phys. Rev.* **D74** (2006) 066007, [hep-th/0605212].

[138] I. R. Klebanov and M. J. Strassler, "Supergravity and a confining gauge theory: Duality cascades and chi SB resolution of naked singularities," *JHEP* **08** (2000) 052, [hep-th/0007191].

[139] S. Kachru, R. Kallosh, A. D. Linde and S. P. Trivedi, "De Sitter vacua in string theory," *Phys. Rev.* **D68** (2003) 046005, [hep-th/0301240].

[140] I. Bena, M. Grana and N. Halmagyi, "On the Existence of Meta-stable Vacua in Klebanov-Strassler," *JHEP* **09** (2010) 087, [0912.3519].

[141] O. DeWolfe, S. Kachru and M. Mulligan, "A Gravity Dual of Metastable Dynamical Supersymmetry Breaking," *Phys. Rev.* **D77** (2008) 065011, [0801.1520].

[142] P. McGuirk, G. Shiu and Y. Sumitomo, "Non-supersymmetric infrared perturbations to the warped deformed conifold," *Nucl. Phys.* **B842** (2011) 383–413, [0910.4581].

[143] A. Dymarsky, "On gravity dual of a metastable vacuum in Klebanov-Strassler theory," *JHEP* **05** (2011) 053, [1102.1734].

[144] I. Bena, G. Giecold, M. Grana, N. Halmagyi and S. Massai, "On Metastable Vacua and the Warped Deformed Conifold: Analytic Results," *Class. Quant. Grav.* **30** (2013) 015003, [1102.2403].

[145] I. Bena, G. Giecold, M. Grana, N. Halmagyi and S. Massai, "The backreaction of anti-D3 branes on the Klebanov-Strassler geometry," *JHEP* **06** (2013) 060, [1106.6165].

[146] S. Massai, "A Comment on anti-brane singularities in warped throats," 1202.3789.

[147] I. Bena, M. Grana, S. Kuperstein and S. Massai, "Anti-D3 Branes: Singular to the bitter end," *Phys. Rev.* **D87** (2013) 106010, [1206.6369].

[148] A. Dymarsky and S. Massai, "Uplifting the baryonic branch: a test for backreacting anti-D3-branes," *JHEP* **11** (2014) 034, [1310.0015].

[149] T. Gherghetta and A. Pomarol, "The Standard model partly supersymmetric," *Phys. Rev.* **D67** (2003) 085018, [hep-ph/0302001].

[150] R. Sundrum, "SUSY Splits, But Then Returns," *JHEP* **01** (2011) 062, [0909.5430].

[151] D. Baumann, A. Dymarsky, S. Kachru, I. R. Klebanov and L. McAllister,

"D3-brane Potentials from Fluxes in AdS/CFT," *JHEP* **06** (2010) 072, [1001.5028].

[152] S. Kachru, D. Simic and S. P. Trivedi, "Stable Non-Supersymmetric Throats in String Theory," *JHEP* **05** (2010) 067, [0905.2970].

[153] A. Dymarsky and S. Kuperstein, "Non-supersymmetric Conifold," *JHEP* **08** (2012) 033, [1111.1731].

[154] P. G. Camara, L. E. Ibanez and A. M. Uranga, "Flux-induced SUSY-breaking soft terms on D7-D3 brane systems," *Nucl. Phys.* **B708** (2005) 268–316, [hep-th/0408036].

[155] D. Lust, S. Reffert and S. Stieberger, "Flux-induced soft supersymmetry breaking in chiral type IIB orientifolds with D3 / D7-branes," *Nucl. Phys.* **B706** (2005) 3–52, [hep-th/0406092].

[156] D. Lust, S. Reffert and S. Stieberger, "MSSM with soft SUSY breaking terms from D7-branes with fluxes," *Nucl. Phys.* **B727** (2005) 264–300, [hep-th/0410074].

[157] C. P. Burgess, P. G. Camara, S. P. de Alwis, S. B. Giddings, A. Maharana, F. Quevedo et al., "Warped Supersymmetry Breaking," *JHEP* **04** (2008) 053, [hep-th/0610255].

[158] D. Lust, F. Marchesano, L. Martucci and D. Tsimpis, "Generalized non-supersymmetric flux vacua," *JHEP* **11** (2008) 021, [0807.4540].

[159] P. McGuirk, G. Shiu and Y. Sumitomo, "Holographic gauge mediation via strongly coupled messengers," *Phys. Rev.* **D81** (2010) 026005, [0911.0019].

[160] P. McGuirk, "Hidden-sector current-current correlators in holographic gauge mediation," *Phys. Rev.* **D85** (2012) 045025, [1110.5075].

[161] F. Benini, A. Dymarsky, S. Franco, S. Kachru, D. Simic and H. Verlinde, "Holographic Gauge Mediation," *JHEP* **12** (2009) 031, [0903.0619].

[162] P. McGuirk, "Falling flavors in AdS/CFT," *JHEP* **07** (2013) 102, [1212.2210].

[163] K. Hashimoto and S. Nagaoka, "Recombination of intersecting D-branes by local tachyon condensation," *JHEP* **06** (2003) 034, [hep-th/0303204].

[164] S. Nagaoka, "Higher dimensional recombination of intersecting D-branes," *JHEP* **02** (2004) 063, [hep-th/0312010].

[165] S. Cecotti, M. C. N. Cheng, J. J. Heckman and C. Vafa, "Yukawa Couplings in F-theory and Non-Commutative Geometry," 0910.0477.

[166] F. Marchesano, P. McGuirk and G. Shiu, "Chiral matter wavefunctions in warped compactifications," *JHEP* **05** (2011) 090, [1012.2759].

[167] M. Marino, R. Minasian, G. W. Moore and A. Strominger, "Nonlinear instantons from supersymmetric p-branes," *JHEP* **01** (2000) 005, [hep-th/9911206].

[168] J. Gomis, F. Marchesano and D. Mateos, "An Open string landscape," *JHEP* **11** (2005) 021, [hep-th/0506179].

[169] L. Martucci and P. Smyth, "Supersymmetric D-branes and calibrations on general N=1 backgrounds," *JHEP* **11** (2005) 048, [hep-th/0507099].

[170] D. Marolf, L. Martucci and P. J. Silva, "Fermions, T duality and effective actions for D-branes in bosonic backgrounds," *JHEP* **04** (2003) 051, [hep-th/0303209].

[171] D. Marolf, L. Martucci and P. J. Silva, "Actions and Fermionic symmetries for D-branes in bosonic backgrounds," *JHEP* **07** (2003) 019, [hep-th/0306066].

[172] L. Martucci, J. Rosseel, D. Van den Bleeken and A. Van Proeyen, "Dirac actions for D-branes on backgrounds with fluxes," *Class. Quant. Grav.* **22** (2005) 2745–2764, [hep-th/0504041].

[173] M. Berkooz, M. R. Douglas and R. G. Leigh, "Branes intersecting at angles," *Nucl. Phys.* **B480** (1996) 265–278, [hep-th/9606139].

[174] L. Aparicio, A. Font, L. E. Ibanez and F. Marchesano, "Flux and Instanton Effects in Local F-theory Models and Hierarchical Fermion Masses," *JHEP* **08** (2011) 152, [1104.2609].

[175] D. Cremades, L. E. Ibanez and F. Marchesano, "Computing Yukawa couplings from magnetized extra dimensions," *JHEP* **05** (2004) 079, [hep-th/0404229].

[176] F. Marchesano, P. McGuirk and G. Shiu, "Open String Wavefunctions in Warped Compactifications," *JHEP* **04** (2009) 095, [0812.2247].

[177] R. C. Myers, "Dielectric branes," *JHEP* **12** (1999) 022, [hep-th/9910053].

[178] B. Wynants, "Supersymmetric Actions for Multiple D-Branes on D-Brane Backgrounds," Master's thesis, Katholieke Universiteit Leuven, 2006.

[179] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*. Dover Publications, Inc., 2001.

[180] E. Shuster, "Killing spinors and supersymmetry on AdS," *Nucl. Phys.* **B554** (1999) 198–214, [hep-th/9902129].

[181] A. Chodos and E. Myers, "Gravitational Contribution to the Casimir Energy in Kaluza-Klein Theories," *Annals Phys.* **156** (1984) 412.

[182] K. Skenderis, "Lecture notes on holographic renormalization," *Class. Quant. Grav.* **19** (2002) 5849–5876, [hep-th/0209067].

[183] M. Krasnitz, "Correlation functions in a cascading N=1 gauge theory from supergravity," *JHEP* **12** (2002) 048, [hep-th/0209163].

[184] O. Aharony, A. Buchel and A. Yarom, "Holographic renormalization of cascading gauge theories," *Phys. Rev.* **D72** (2005) 066003, [hep-th/0506002].

[185] I. R. Klebanov and A. A. Tseytlin, "Gravity duals of supersymmetric SU(N) x SU(N+M) gauge theories," *Nucl. Phys.* **B578** (2000) 123–138, [hep-th/0002159].

[186] M. Berg, M. Haack and W. Mueck, "Glueballs vs. Gluinoballs: Fluctuation Spectra in Non-AdS/Non-CFT," *Nucl. Phys.* **B789** (2008) 1–44, [hep-th/0612224].

[187] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*. Cambridge University Press, 2007.

[188] A. Meremianin, "Hyperspherical Harmonics with Arbitrary Arguments," *J.Math.Phys.* **50** (2009) 013526, [0807.2128].