# Gauge symmetries in Quantum Gravity and String Theory

Memoria de Tesis Doctoral presentada por


**Miguel Montero Muñoz**

ante el Departamento de Física Teórica
de la Universidad Autónoma de Madrid
para optar al Título de Doctor en Física Teórica

Tesis Doctoral dirigida por el **Dr. Ángel M. Uranga Urteaga**,
Profesor de Investigación del Instituto de Física Teórica UAM-CSIC

UAM UNIVERSIDAD AUTÓNOMA DE MADRID

ift

Departamento de Física Teórica
Universidad Autónoma de Madrid

Instituto de Física Teórica
UAM-CSIC


Junio de 2016

*A tod@s aquell@s que alguna vez*
*me enseñaron algo de física.*

# Agradecimientos

Imposible comenzar estos agradecimientos por otro que Ángel, mi director de tesis. Gracias por toda la física que me has enseñado estos años, primero en el máster y luego en el doctorado, y por tener siempre la puerta abierta o el email disponible para resolver en un momento desde pequeñas dudas hasta preguntas que me habían mareado por semanas. Gracias por escucharme y animarme siempre a explorar mis propias ideas, por tomarlas en serio, incluyendo todas esas veces en las que realmente no tenían sentido. Y gracias por las tuyas, que han hecho mi doctorado tan divertido, por tus mini-reviews en la pizarra y por tus "yo creo que ahí hay algo, míralo" que invariablemente dan en el clavo. ¡Con que aprenda en el postdoc la mitad de lo que he aprendido de ti, me doy por satisfecho!

Mil gracias a Mikel, Ander, e Iñaki, porque sin vosotros la primera parte de esta tesis no existiría. A Mikel y a mi *pisha* Ander por compartir conmigo los primeros pasos en el extraño mundo de las teorías de cuerdas con aún más dimensiones. Ander, gracias por toda la física, las apuestas inteligentes, y las risas en general. Iñaki, cada vez que hablo contigo de física flipo en colores. Ha sido un privilegio currar contigo, uno que espero mantener en el futuro. Gracias por todo lo que me has enseñado, y por lo mucho que disfruté mi visita a Munich.

Muchas gracias a mi amiga-enemiga T-dual, Irene. Por todo lo que me has enseñado, tanto de física (¡y tanto más por lo que cuesta que nos entendamos a veces!) como de todo tipo de cosas, desde cómo no comerme arañas en la jungla hasta las constelaciones, pasando por (amagos de) esquiar. La segunda parte de la tesis no existiría sin ti. ¡No dudo de que seguiremos petándolo en Utrecht!

Gracias también a Luis, primero por insistir tanto con que "en eso del relaxión se puede decir algo" y por ser luego esencial para decir ese algo. Gracias por compartir tus ideas como si nada, por tu entusiasmo por la física tan contagioso, y por todas las charlas. Gracias, junto a Ángel, por escribir El Libro. Gracias también a Fernando por las discusiones de física y las risas.

Thanks to Gary and Pablo for an awesome month in Madison which allowed me to try all the delicacies of American cuisine, and for all the physics. I am sure the best is yet to come!

Grazzzzzie mile a il mio Carissimo Amico, Gianluca Zoccarato. Sei un ragazzo molto gentile se queda corto para describirte, pero es hasta donde alcanza mi italiano (al menos, la parte buena). No creo haber discutido más de física con otra persona que contigo durante mi doctorado, salvo Ángel. Eres un crack de la física además de un tío estupendo, y estoy orgulloso de ser tu amigo.

No puedo agradecer lo suficiente a Gianluca, pero sobre todo a Irene y Ander por llevarme de aventuras con ellos, a cuevas eslovenas y junglas indias. Me habéis despertado el gusanillo de viajar, y eso es una de las cosas más bonitas que me llevo del doctorado.

Víctor, eres un jefe a todos los niveles, te adoro, y lo sabes. Gracias por estar ahí siempre, por lo que me has enseñado de física y de todo lo demás, por esa conesion mental masema que tenemos. ¡Ni siquiera el postdoc puede separarnos!

Víctor y Mario. Con vosotros comparto un vínculo más fuerte que cualquier otro: el de la procrastinación. Juntos somos *Disco Chino*, el despacho más dumb y divertido que jamás se haya visto en el IFT. Gracias por hacer de cada día una fiesta en la que nunca han faltado el café, tuiter, vainas y la ocasional discusión de física. Sabéis lo que os voy a echar de menos, aunque mi índice h probablemente lo agradecerá, al igual que los vuestros. Y gracias a Brian, que estuvo allí antes de que las cosas se nos fueran de las manos.

Gracias a los profesores del máster por enseñarme tanto, y a toda la gente del IFT que han convertido estos años en una experiencia inolvidable. Gracias Xabi, Aitor, Josu, Fede, Oscar, Emilia, Ilaria, Doris, Rocío, Ginevra, Santi, Giorgos, Patrick, Mar, Wieland, Oscar, Paco, Susana, y a tod@s l@s demás. Gracias Berta por tu amistad y por esos desayunos que ya echo de menos. Gracias Carlos y Diego por toda la física que me habéis enseñado. Gracias Juanmi por todos esos acertijos y por ser una fuente inagotable de historias increíbles. Y a Patxi por todo, pero sobre todo por la física y las hamburguesas. Gracias también a Isabel, Mónica y Mónica, Chabe, y María por toda su ayuda y por hacer la hora de la comida tan divertida.

Muchas gracias a la Fundación Obra Social La Caixa, por la beca que ha hecho posible esta tesis.

No estaría escribiendo esto si no fuera por toda la gente de QUINFOG, que me acogió como a uno más cuando apenas sabía diagonalizar una matriz. En especial, gracias a Juan León y a Edu Martín por darme una oportunidad que poca gente tiene, la de hacer investigación durante la carrera, y por convertirlo en algo tan divertido.

Gracias, Jose por poder contar siempre contigo. Muchas gracias a Nere, Marisa, Yure, Guti, Sandra, Pablo, Isa, Lucía, por estar ahí, ser siempre tan divertid@s, y seguir avisándome de las quedadas aunque yo sea un desastre y me olvide de todo. Y en especial a Guti y Paloma, por todo el cariño, el apoyo constante, y por enseñarme siempre cómo ser mejor persona. Tengo mucha suerte de conoceros a tod@s vosotr@s. Gracias a David por las cervezas, los paseos por la montaña, y por tu punto de vista.

Laura, gracias por nuestras chocolatadas. Confío en que sigas manteniéndome al día de las últimas novedades con el máximo glamour. Guille, vecino, gracias por las cerves! A mis amigos Fotineros, Valen, Ele, Sergio y Cofy, gracias por todas las cenas, viajes, disfraces, gatiiiiitos, Sherlock, y mucho más. Gracias Chuso por mantener vivas las cenas de Navidad. Y muchas gracias a Juan Enrique por todos sus libros y por explicarme tantas cosas, desde el tiro parabólico hasta qué es un escalar.

A mis amigos del Briconsejo, Alf, Cobo y Mikel. Gracias por todos estos años, y por los que están por venir. Alf, eres un tío cojonudo y lo vas a petar. Cobo, gracias por las risas y todo lo demás desde la OIME hasta las del otro día. Más te vale visitar. Mikel, tenemos que ir más a tu casa.

Gracias a Juana, mi fan #1, por acogerme en tu casa tantas veces y tratarme siempre como a un hijo más. Decir que eres un sol se queda corto. Javivi, gracias por las series y por los juegos. A Francisco, gracias por tu paciencia y por todos esos viajes en coche. Albertito, cada día que pasa molas más. Gracias por ser siempre tan divertido (incluso cuando te cuelas en cualquier hueco). Y gracias a Cuchifú por ser un felpudito tan mono y hacernos sonreír tanto desde su rincón.

Gracias a mis padres, Eduardo y Mar, por todo. Si he llegado hasta aquí ha sido gracias a vuestro apoyo, cariño, y amor, y por estar ahí siempre que os necesito. Os quiero mucho. Miau, estoy muy, muy, muy orgulloso de ti. Gracias por tu compañía constante, tu cariño, las conversaciones sobre física y videojuegos, y por ser un ejemplo a seguir en tantas cosas. Te quiero.

Y finalmente, gracias a Bubi. Por ser tan divertida, buena persona, graciosa, y mimosa. Por los viajes y todas las series que hemos visto juntos. Por las ensaladas. Por los momentos-bola en los que nos olvidamos de todo. Por los momentos buenos, pero sobre todo por estar ahí en los malos. Por entenderme mejor que yo mismo, por quererme, apoyarme en los momentos difíciles, por luchar siempre por nosotros, por ser fuerte cuando yo no he podido serlo. Gracias por ser mi mejor amiga, por empujarme a enfrentarme a mis miedos, incluso si eso significa estar separados. Y gracias por prestarme el cuarto raro de tu casa, sin el cual aún estaría escribiendo esta tesis. Te quiero.

# Abstract


This thesis explores the spectrum of states charged under global and gauge symmetries in String Theory, and the constraints that they pose on specific phenomenological models. On one hand, we use supercritical theories (those with more than 10 or 26 dimensions) to embed discrete gauge symmetries of the critical theory into continuous ones in a supercritical extension, and use this description to construct the associated charged objects. Supercritical string theories are related to the usual critical ones via a process of closed-string tachyon condensation. This picture can be extended to construct several other objects as solitons of the supercritical tachyon, such as the NS5-brane in the heterotic theory or the $\mathbb{C}^2/\mathbb{Z}_2$ orbifold in type II theories. In both cases, we unveil a real K-theory structure in the supercritical theory which allows us to classify and construct the the topologically protected solitons of the theory. A supercritical perspective on Gauged Linear Sigma Models and some connections to Matrix Factorizations are pointed out.

We also discuss the constraints on large field inflation models coming from gravitational instantons in the effective field theory. These instantons are related to and motivated by the Weak Gravity Conjecture, which demands the existence of light charged objects in any gauge theory consistently coupled to quantum gravity. We constrain both single and multiple axion models of inflation, and these constraints make the obtention of parametrically large axion field ranges very difficult. We highlight a simple loophole to these arguments which however requires the strong form of the Weak Gravity Conjecture to be false.

Finally, we solve some issues of the relaxion proposal as a solution to the hierarchy problem by embedding it in an axion monodromy setup. Generically, in monodromic models there are membranes which may mediate transitions between the branches of the potential. These can be dangerous to the relaxion proposal, depending on the precise value of the tension of the membranes. The WGC provides an upper bound for this tension, in such a way that the membrane constraints are sufficient to rule out the simplest versions of the relaxation mechanism. We also provide the first steps towards an embedding of the relaxion proposal in D-brane models.

# Resumen

Esta tesis explora el espectro de estados cargados bajo simetrías en teoría de cuerdas, tanto globales como gauge, y las ligaduras que suponen dichos estados en modelos fenomenológicos específicos. Por una parte, se emplean teorías supercríticas (que viven en más de diez o veintiséis dimensiones) para embeber simetrías gauge discretas de la teoría crítica en simetrías gauge continuas de una extensión supercrítica, usando esta descripción para construir también los objetos cargados asociados. Las teorías de cuerdas supercríticas se relacionan con sus compañeras críticas mediante un proceso de condensación de taquión de cuerda cerrada. Esta imagen se utiliza también para construir otros objetos como solitones del taquión supercrítico, como la brana NS5 en la teoría heterótica o el orbifold $\mathbb{C}^2/\mathbb{Z}_2$ en teorías de tipo II. En ambos casos, se descubre una estructura de teoría K real subyacente a la teoría supercrítica, la cual permite clasificar y construir explícitamente solitones del taquión protegidos topológicamente. Se enfatiza también la perspectiva supercrítica en los modelos sigma lineales gaugeados, así como ciertas conexiones con factorizaciones de matrices.

En segundo lugar, se obtienen cotas sobre ciertos parámetros en modelos inflacionarios de campo grande, provenientes de instantones gravitacionales presentes en la teoría efectiva. Estos instantones están motivados por la Conjetura de la Gravedad Débil, que exige la existencia de objetos cargados ligeros en cualquier teoría gauge acoplada de manera consistente a gravedad. Se discuten efectos sobre modelos tanto de uno como de varios axiones, y en ambos casos las cotas obtenidas dificultan significativamente la construcción de inflación de campo grande. Se señala un resquicio en el argumento que sin embargo implicaría la falsedad de la forma fuerte de la Conjetura de la Gravedad Débil.

Finalmente, se resuelven algunos problemas de la propuesta del relaxión para resolver el problema de la jerarquía electrodébil mediante su incorporación a un modelo de axión monodrómico. De manera genérica, los modelos monodrómicos incluyen membranas que pueden mediar transiciones entre las diferentes ramas del potencial, las cuales pueden comprometer la viabilidad del relaxión dependiendo del valor concreto de la tensión de las membranas. La Conjetura de la Gravedad Débil proporciona en este caso una cota superior para dicha tensión, de tal forma que las cotas basadas en membranas son suficientes para descartar las versiones más simples del relaxión. También se presentan los primeros pasos hacia un embebimiento de los modelos relaxiónicos en teoría de cuerdas, en el contexto de modelos de D-branas.

# Contents

# 1
# Introduction

The XX century undoubtedly witnessed impressive progress in every field of physics. While a physicist from the early 1800s could understand many (though by no means all) of the interesting open problems of fin de siècle, a physicist of the first decade of the XX century would be alien to almost any research problem in modern physics. Two of the four fundamental interactions and all of the particles regarded today as the fundamental blocks of matter were discovered starting in 1897 with the discovery of the electron by J. J. Thompson. 19th century cosmology was dominated by steady-state theories, with the discovery of the expansion of the universe and the Big Bang paradigm all belong to the 20th century.

Though all these changes were driven by an intense interplay between theory and experiment, they ultimately rely on two basic paradigm shifts: Quantum mechanics and special relativity. The first comprises a radical change of our understanding of the world, shifting from a deterministic to a probabilistic perspective, while the second tells us that space and time mix in a very specific way. These two basic aspects of the world are most powerful when working together: The formalism of quantum field theory restricts enormously the allowed matter and its interactions. The Standard Model of particle physics, a typical example of a quantum field theory, provides a unified description of three of the four fundamental forces and most of the interactions of matter. It is the most accurate physical theory ever produced.

Luckily for newcomers to physics, this is far from the end of the story. On one hand, there are many interesting open problems whose solution likely lies within the framework of quantum field theory, such as the nature of dark matter, the origin of neutrino masses, or the strong CP problem. Hopefully the LHC and its successors will provide us with new particles and phenomena, for which the most natural explanation will be in terms of a QFT. On the other hand, the framework of quantum field theory, so successful in the description of strong, weak, and electromagnetic interactions, does not work so well for gravity. It only provides meaningful answers when the energies per particle are much lower than the Planck scale $M_P \approx 10^{19}\ GeV$. Technically, an interacting theory of a massless spin 2 boson is non-renormalizable.

A similar problem was encountered in the development of the theory of the weak interactions. Fermi's theory is non-renormalizable, providing an accurate description of the physics only at scales below $\approx 70\ GeV$. At this scale, the theory is superseded by the Standard Model, a different quantum field theory valid on a larger energy range: We say the SM is an ultraviolet completion of Fermi's theory. We do not know where the predictions of the Standard Model start to fail (more precise measurements of the mass of the top quark may put an upper bound to this scale, and heavy neutrino masses suggest another upper bound no bigger than $10^{16}\ GeV$), but the theory itself is mathematically consistent up to the Planck scale and beyond.

The difference between Fermi's theory and gravity is that the former admits an ultraviolet completion which is a weakly coupled QFT, whereas the latter does not: Since any theory weakly coupled theory with a massless fundamental spin 2 field is nonrenormalizable, the only possibility would be to have the graviton to be a composite particle, which is forbidden by the Weinberg-Witten theorem. Therefore to find an ultraviolet completion of gravity we are forced to drop either field theory or weak coupling. The former is the approach taken in the asymptotic safety scenario (see [1] for a review), in which gravity is described by an ultraviolet fixed point under renormalization group flow. So far we only know of one example of a weakly coupled UV completion of gravity, and that is string theory.

String theory is, by any account, one of the most fruitful fields of physics of the last decades. The different vibration modes of the string provide us not only with gravity, but also with every other kind of particle observed and indeed with every kind of particle allowed in quantum field theory. In hindsight, it is reasonable to expect a quantum theory of gravity to include other particles; gravity couples to the stress-energy tensor, which is sourced by every kind of matter. Quantum corrections to e.g. graviton-graviton scattering include loops of every other field in the theory. On top of this, string theory has been a constant source of insight into other branches of physics, such as gauge theory, supersymmetry, extra dimensions, holography and condensed matter physics or even quantum information. It naturally includes gauge coupling unification, provides us with solutions to the dark matter and strong CP problems, neutrino masses, a sensible anthropic explanation of the value of the cosmological constant and the hierarchy problem. There is also a rich interplay with the mathematics community, with some branches of mathematics such as algebraic geometry becoming essential part of the toolkit of the string theorist, and some surprising mathematical results such as the so-called monstrous moonshine being established via string theory.

Whether or not string theory is the right quantum theory of gravity, it is at least one consistent example of such. Unique or not, studying it we can expect to learn a great deal about quantum gravity, in much the same way as we first learnt about many important features of quantum field theories from the study of a particular example, QED. Of the many aspects of string theory still under active research, this thesis focus broadly on the properties of solitonic and charged states in the theory, both at the formal and at a more phenomenological level. Two main tools will be used to achieve this: Supercritical string theories (whose study is also an end in itself), and the Weak Gravity Conjecture.

## Supercritical string theory

One of the main breakthroughs in string theory is the realization that it is not a theory of strings. The theory includes a plethora of other extended objects, the branes, essential to define the non-perturbative regime, in which they are on equal footing with the fundamental string. The best known such objects are D-branes, which admit a perturbative description in terms of open string sectors, and which underlie several deep connections, like the microstate interpretation of the Bekenstein-Hawking entropy for certain black holes and the AdS/CFT or gauge/gravity correspondence.

However, D-branes are far from exhausting the non-perturbative sector of string theory. There is also a plethora of other nonperturbative states in string theory, which are often non-BPS. This means that they are out of reach of standard supersymmetric techniques which have proven so useful to analyze D-branes and other supersymmetric solitons. It is important and also uncommon to have a handle on non-supersymmetric sectors of string theory at the nonperturbative level.

One of the more famous features of string theory is that spacetime dimension is not a free parameter, but rather it is fixed by the theory. Bosonic string theories fix the spacetime dimension to be 26 and the more phenomenologically interesting superstring theories fix it to be 10.

However, to constrain the dimension of the theory is essential to assume Lorentz invariance. There is no motivation for Lorentz invariance in $D > 4$ dimensions, since if there are other dimensions present they must be compact, small, and presumably without a Lorentz-invariant metric. Therefore, it is not surprising that noncritical string theories (which live in any $D > 1$ spacetime dimensions) have a long and rich history. Most of the early research in noncritical string theories was focused on 2-dimensional models which, it was hoped, could provide a tractable toy model of quantum gravity.

Another interesting possibility is to consider supercritical string theories, that is, string theories which live in more than 10 or 26 dimensions. These theories are manifestly not Lorentz invariant; the central charge of the sigma model is compensated by a linear dilaton background which explicitly breaks Lorentz invariance. Since the dilaton changes in time, supercritical theories constitute a time-dependent exact solution of string theory, in which the theory becomes strongly coupled in some region of spacetime.

These theories also have a tachyon in their spectrum. This tachyon is similar in nature to the tachyon of the bosonic string in 26 dimensions. This means that the supercritical string theories are unstable, and describe a phase of the theory sitting on the top of a potential, much like the Higgs field in the Standard model before electroweak symmetry breaking. In a similar fashion, the tachyon may condense and reach a new, stable minimum throughout spacetime. This stable minimum turns out to be a critical, ten-dimensional string theory. In other words, different supercritical string theories living in different spacetime dimensions are connected to each other via a process of tachyon condensation. Tachyon condensation also allow us to reach the familiar critical string theories, thus embedding the well-known dualities among these in a much richer web of connections between supercrtitical string theories.

These supercritical theories can be used to acquire new insights about the nonperturbative sector of critical string theory. For instance, Calabi-Yau compactifications often have discrete isometries which descend to discrete symmetries of the low-energy effective field theory. In some cases these isometries can be embedded into continuous ones in the supercritical theory. This has the implication that the low-energy discrete symmetries are actually discrete gauge symmetries, the reminder of a continuous gauge symmetry which is incompletely Higgsed by the supercritical tachyon.

By considering particular profiles for the fields in the supercritical theory, we may reach interesting states in the critical theory. There is no reason to expect that these objects will be supersymmetric in general, so we are studying the nonperturbative, non-supersymmetric sector of string theory. In particular, it is possible to construct objects charged under discrete symmetries of the original theory. Although one expect such objects to exist in a quantum theory of gravity, it is hard to characterize them explicitly. Tachyon condensation provides a systematic way of explicitly constructing new nonperturbative massive states in string theory which would be difficult to describe otherwise.

The classification of these states is helped by the mathematical tool of real K-theory. The supercritical theory of both heterotic and type II theories includes an orbifold action which splits the dimensions into two different classes. In the heterotic case, the tachyon gradient is charged both under the real gauge bundle and the tangent bundle to one of these dimensions, whereas in the type II case it is charged under two geometrical bundles. This means that the tachyon describes a K-theory equivalence class of real bundles. Hence, the task of classifying all the solitons which may arise from tachyon condensation becomes the task of classifying real K-theory groups, a task which was already done in the context of tachyon condensation of $D9-\overline{D9}$ pairs in type I string theory.

## The Weak Gravity Conjecture

Gravity is the weakest of all the fundamental forces. From the point of view of a particle physicist which works with effective field theories, there is no special reason for this; it is equally consistent to have a force weaker than gravity. However, as mentioned above, gravity is beyond the scope of quantum field theory[1]. While at low energies any realistic theory including gravity must resemble a quantum field theory, the converse is not true: There are examples of effective field theories which are perfectly fine as quantum field theories, but which can never arise as the low energy limit of a quantum theory of gravity. It is very easy to write down such theories, which leads to the picture of a landscape of many bad quantum field theories without

[1]At least conventional QFT on the same spacetime; the AdS/CFT correspondence actually instructs us that gravity on an AdS spacetime is equivalent to a quantum field theory in one dimension less.

gravitational UV completion, and a few nice ones with. This picture has often been poetically referred to as the Swampland.

The Swampland can be good or bad news, depending on who you ask. It is a nuisance to the effective field theorist who wants to address questions such as the hierarchy problem or the origin of the cosmological constant purely within the framework of effective field theory, since any candidate solution may lie in the Swampland. But it represents a great opportunity for the string theorist. By studying generic properties of different effective field theories which arise from string theory, one can hope to characterize the set of good theories.

As a particular example, it has been suggested, based on arguments related to black hole evaporation, that gravity does not allow global symmetries. Hence, any effective field theory with a global symmetry is in the Swampland; the global symmetry must be broken at some scale. This is verified in every model coming from string theory, and for a large class of global symmetries one can give an exact proof of this fact within string theory. Another example comes from particles in continuous spin representations. These are perfectly fine to an effective field theorist, but can be shown to be absent in string theory [2].

The above examples, while very useful, are qualitative requirements that a good theory should satisfy. There is, to my knowledge, only one proposed quantitative feature of theories outside the Swampland: The Weak Gravity Conjecture. Roughly speaking, the WGC demands the existence of light enough objects for every weakly coupled interaction in the theory. Equivalently, we can expect any field theory not including these particles to fail (in effective field theory terms, there is a cutoff) at a scale of the order of the tension of this object. This conjecture can be justified by recognizing that in the limit of extremely weak coupling, the theory recovers a global symmetry. This can be reconciled with the absence of global symmetries described above by demanding the existence of such light objects.

Arguments relying on black hole evaporation are often shaky, and this is even more so in the case of the WGC. As in the other cases above, string theory provides most of the concrete evidence for the conjecture: it is true so far in every string model in which we have enough control to test the conjecture, and for a restricted class of models it is possible to provide a formal proof. Significant effort has been devoted recently towards finding a more compelling and general argument for the WGC, but if it is not an actual constraint on the Swampland, it likely is a constraint for the stringy Swampland. As discussed above, this is the only kind of Swampland we can check right now.

For an ordinary $U(1)$ gauge interaction in four dimensions with coupling $g$, the WGC demands the existence of a particle of mass $\lesssim gM_P$. For two such particles, the gauge force is bigger than the gravitational attraction. This explains the name "Weak Gravity Conjecture"; there is always some particle in the theory for which gravity is the weakest force indeed.

Every particle in the Standard Model satisfies the WGC by a huge margin. However, one can also apply the Weak Gravity Conjecture to other kind of interactions arising in four dimensions, and these turn out to have important phenomenological consequences.

A version of the WGC conjecture applies to axions, scalar particles with a discrete shift symmetry. In this case, and with some caveats, the WGC implies that the axion decay constant $f$ of a single axion must be lower than $M_P$. Axions are important in inflationary cosmology, in which they are often used as an inflaton because the shift symmetry provides some degree of protection against unwanted corrections to the potential. This protection is however limited by the WGC, in such a way that inflationary models based on periodic axions are unable to predict a large value of the scalar-to-tensor ratio $r$. There are several experiments, notably BICEP3 and the Keck Array, which will continue to explore the region of large $r$ in the near future. If $r \gtrsim 0.01$ is indeed measured, models of inflation based on a few periodic axions would be unlikely to work, due to the WGC.

Another version of the conjecture applies to three-forms, a more exotic kind of four-dimensional field which does not have propagating degrees of freedom. In spite of this, three-forms are important to model builders. They help solve the cosmological constant problem in string theory in an antropic way, providing a landscape of allowed values for the cosmological constant. They are also important to inflation, since when mixed with an axion they give rise to a non-periodic field which inherits much of the protection of the axion against corrections. These models, collectively known as "axion monodromy", are the main class of single-field inflationary models which can predict a large $r$ while remaining consistent with quantum gravity constraints. Axion monodromy is also an important to give a sensible UV embedding of the relaxation mechanism, a novel proposal to solve the hierarchy problem. In both these models, the WGC predicts the existence of "membranes", domain walls which mediate a sharp change of the value of the field strength of the three form. The effects of these membranes turn out to be very suppressed in axion monodromy inflation scenarios, but they are very significant to the relaxion case.

## Plan of the thesis

This thesis will be divided in two parts, plus conclusions and appendices:

- Part I will begin in Chapter 2 with an introduction to supercritical string theory, describing the supercritical theories which will be the main tool which will be the main tools employed in this part of the thesis. After this, Chapter 3 describes the use of supercritical string theories to embed discrete symmetries into continuous supercritical extensions, as well as the construction of objects charged under such symmetries. Chapter 4 then extends these tools to the construction of the heterotic NS5 brane, and unveils a K-theory structure in supercritical theories. Chapter 5 discusses condensation of nontrivial tachyon profiles in type II theories, describing some topological supercritical couplings as well as another real K-theory structure in the theory.

- Part II is devoted to the applications of the Weak Gravity Conjecture to constrain effective field theories. Chapter 6 introduces the Weak Gravity Conjecture, discussing its justification, the evidence for it, and the different forms of the conjecture currently used in the literature. Chapter 7 discusses constraints posed by gravitational instantons inspired by the WGC to large-field inflation axionic models. Chapter 8 provides a monodromic UV completion of the relaxion proposal for a solution of the hierarchy problem, which then receives additional constraints from nucleation of monodromic membranes.

- Finally, Chapter 9 presents the main conclusions of this thesis. Finally, the Appendices provide relevant technical details of the results presented in the thesis, omitted from the main text to ease its reading. Appendix A provides some partition functions of the supercritical string theories. Appendix B is an informal review of the K-theory focused on its applications to the main text. Appendix C describes the Atiyah-Bott-Saphiro construction essential for Chapters 4 and 5. Appendix D contains a review of axion monodromy models, focusing on the Kaloper-Sorbo protection of the potential. Appendix E provides the essential details behind the bubble nucleation formulae of Chapter 8. Finally, Appendix F discusses the details of the partial embedding of relaxion models in string theory presented in Chapter 8.

# Part I

# Supercritical string theory

# 2

# Supercritical string theories and tachyon condensation

Supercritical string theories [3, 4, 5] provide a significant extension of the familiar duality web between supersymmetric string theories. They are related to these theories via a closed-string tachyon condensation process in which the tachyon "eats up" the extra dimensions. This Chapter provides a review of the supercritical theories which will be of use throughout this thesis as well as the tachyon condensation process. Original references for most of the results in this Chapter are [4, 5].

## 2.1 A survey of supercritical string theories

### 2.1.1 Beyond the critical dimension

As its name suggests, String Theory is a theory of fundamental 1-dimensional objects, the strings, in contrast with the pointlike particles of quantum field theory. The particular string theory is defined by the field content in the two-dimensional worldsheet of the propagating string. One of the most characteristic features of string theory is the prediction of a fixed spacetime dimension, which depends solely on this field content. Some of the most well-known possibilities are:

- Bosonic string theory: The simplest and oldest string theory, it simply takes the embedding functions $X$ : Worldsheet $\rightarrow$ Spacetime describing how the string is embedded into physical spacetime as its worldsheet content. The action of this theory is the Polyakov action

$$S_P = -\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{\det[g]}g^{ab}G^{\mu\nu}\partial_a X^\mu \partial_b X^\nu, \tag{2.1.1}$$

  where we have used latin indices $a, b\ldots$ to indicate worldsheet fields, and greek indices $\mu, \nu\ldots$ to indicate spacetime labels.

- Type II theory: A simple supersymmetrization of the bosonic theory above, to the embedding functions $X^\mu$ it also adds the superpartner fermionic fields $\psi^\mu$, to achieve $(1,1)$ supersymmetry. The action is the supersymmetric version of Polyakov,

$$S_{II} = -\frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{\det[g]}\left[g^{ab}G^{\mu\nu}\partial_a X^\mu\partial_b X^\nu - i\alpha'\bar{\psi}^\mu\Gamma^c\partial_c\psi_\mu\right]. \qquad (2.1.2)$$

Type II theories further split into IIA and IIB theories. Both have the worldsheet content described above, but are distinguished by the chirality of the supersymmetries in spacetime. In worldsheet terms, they correspond to different choices of GSO projection.

- $SO(32)$ heterotic: This theory has half the supersymmetry of type II theories. The worldsheet content in the so-called fermionic formulation consists of the embedding coordinates and its right-moving superpartners $\psi^\mu$, plus a set of 32 left-moving fermions $\lambda$ in the fundamental representation of $SO(32)$.
- Type I theory: This theory contains both closed and open strings. It preserves the same supersymmetry as the heterotic theory and may be regarded as a particular background of type II theory, containing an orientifold $O9^-$ and 32 spacefilling $D9$-branes.

All of these examples are CFT's, invariant under the action of the conformal group. To have meaningful scattering amplitudes, it is necessary that the theory is invariant under the larger diff×Weyl symmetry which includes arbitrary local rescalings as well as diffeomorphisms. However, for a generic CFT the Weyl symmetry is anomalous, with the anomaly being measured by the total central charge $c - c^g$ of the theory, including a universal ghost sector which only depends on the particular version of string theory under consideration. For the bosonic string the central charge $c$ of the matter has to be 26, whereas for type II superstrings it is 15. Heterotic strings have different left and right-moving central charges, which have to be 26 and 15 respectively.

Any CFT with the right central charge is a valid background for string theory. We will focus on a particular class of models, the linear dilaton backgrounds with action given by

$$S = \frac{1}{4\pi\alpha'}\int d^2\sigma\sqrt{-g}\left[g^{ab}\nabla_a X^\mu\nabla_b X_\mu + \alpha' R V_\mu X^\mu\right]. \qquad (2.1.3)$$

This is just like the Polyakov action (2.1.1), but with an extra term coupling the target space coordinates directly to the worldsheet curvature scalar. This coupling is precisely what results from a dilaton vertex operator [6], and so the action (2.1.3) can be interpreted as a linearly growing dilaton background coupled to the string worldsheet. The theory has different spacetime regions in which the string coupling (which is the expectation value of the dilaton) can be as low or high as we want. We may regard (2.1.3) as the action for the Polyakov string in a particular background state.

While (2.1.3) is actually not conformally invariant at the classical level, it becomes a nice CFT once quantum effects are taken into account [6]. The central charge of the bosonic linear dilaton theory, including the ghost sector, is

$$c - c^g = D - 26 + 6\alpha' V_\mu V^\mu = 0. \tag{2.1.4}$$

Therefore, by tuning $V_\mu V^\mu$ to the right value, we can achieve any value of $D$ we are interested in:

- Theories with a timelike ($V_\mu V^\mu < 0$) dilaton gradient live in spacetime dimensions greater than 26, or 10 in the case of supersymmetric theories. These are called *supercritical* string theories. These are cosmological solutions of a sort, since the dilaton of the theory changes monotonically with time in an appropriate coordinate system. They describe a spacetime evolving from strong to weak string coupling or vice-versa.

- Theories with a spacelike dilaton gradient have $D < 26, 10$ and are called *subcritical* string theories. We will not be very much concerned with them in this work. They describe static but spatially inhomogeneous solutions. A particularly interesting example is $D = 2$, for which (2.1.3) reduces to the so-called Liouville field theory [7, 8], which has been used as a model of quantum gravity and appears in the classical geometric problem of uniformizing Riemann surfaces [9].

- Theories with a null ($V_\mu V^\mu = 0$) dilaton gradient do not change the dimension of the theory, and are best described as particular classical states of the critical string theory. Like supercritical backgrounds, these field configurations correspond to cosmological solutions in which the string coupling rolls to either zero or infinite value. However, since in this case there is no reference frame in which the dilaton gradient $V^\mu$ has only time components, the solution is not spatially homogenous either, and describes a domain wall which propagates at the speed of light and acts as a boundary between two regions at strong and weak coupling .

Having described the general idea behind supercritical string theories, we describe those which will be of use in later Chapters. Many of these details can be found in [4, 5].

### 2.1.2 Mass in a linear dilaton background

For our purposes, the relevant information we need to extract from a given supercritical string theory is the spectrum of its lowest energy excitations. In Lorentz-invariant theories this is typically encoded in terms of the mass (and spin) of each excitation. However, the linear dilaton CFT breaks the Lorentz group and in particular no Casimir element $m^2$ can be found. Thus, we need to define mass in another way. We will follow [4]: A field $\phi$ has mass $m$ when the low-energy effective lagrangian describing its dynamics has the right mass term $m^2\phi^2$ (or linear term for fermions).

String theory effective actions encode tree-level string amplitudes [6, 10], and since the string coupling constant for closed strings is $g_s = e^{\Phi}$ in terms of the dilaton $\Phi$, an action for $\phi$ generated via tree-level amplitudes will be of the form

$$\mathcal{L} \propto \frac{e^{-2\Phi}}{2}[-\partial_\mu\phi\partial^\mu\phi + m^2\phi^2], \tag{2.1.5}$$

which has a modified equation of motion

$$-2\partial_\mu\Phi\partial^\mu\phi + \partial^2\phi + m^2\phi = (\partial_\mu - V_\mu)^2\phi + (m^2 - \partial_\mu\Phi\partial^\mu\Phi)\phi = 0. \tag{2.1.6}$$

When we refer to a mass throughout this work, we mean that the corresponding field satisfies an equation of motion like (2.1.6). To determine the masses of each field in a particular theory, we will use CFT methods to obtain spacetime equations of motion, and read off the mass terms from there. As we will see below, CFT computations yield a value for $(\partial - V)^2$; to determine the mass of the state it is essential how is each particular field related to the lagrangian.

To compute the spectrum of a theory in a linear dilaton background, it is useful to observe that the linear dilaton term in (2.1.3) is

$$S_{\text{Linear dilaton}} = -\frac{V_\mu}{2\pi}\int d^2\sigma\delta(0,0)X^\mu(\sigma,\tau) = -V_\mu X^\mu(0,0). \tag{2.1.7}$$

This has the form [6] of the dilaton vertex operator associated in the worldsheet to a “state” with momentum $iV^\mu$. This momentum is imaginary and so cannot correspond to a physical state. The best way to think of this is as a formal manipulation showing that the linear dilaton theory behaves as the ordinary Polyakov string with momenta shifted by $iV$. The spectrum can of course be computed also in a rigorous manner in covariant quantization, imposing the Virasoro constraints on physical states and quotienting out null states [4].

This is all we need to compute the spectrum. In covariant quantization, one decomposes the theory into left and right moving sectors. Physical state conditions are that the states are annihilated by both the left and right-moving Hamiltonians $H_L$ and $H_R$. We have

$$H_L = \frac{2}{\alpha'}(p+iV)^2 + N_B + N_F + A, \quad H_R = \frac{2}{\alpha'}(p+iV)^2 + \bar{N}_B + \bar{N}_F + \bar{A}, \tag{2.1.8}$$

where the $N_B$ and $N_F$ are the number operators of bosons and fermions, respectively, and $A$, $\bar{A}$ are zero-point energies which are determined above.

### 2.1.3 Bosonic supercritical string theory

This theory is the proptotypical example of a supercritical string theory, and has a long history [6, 7]. As discussed above, the worldsheet field content is $D$ embedding scalars $X^\mu$ (plus ghosts), together with a linear dilaton background with

$$V_\mu V^\mu = \frac{26-D}{6\alpha'}. \tag{2.1.9}$$

For supercritical string theories, $D > 26$ and so the linear dilaton is timelike. We may compute its lowest-lying excitations using the mass formula

$$-(p+iV)^2 = \frac{4}{\alpha'}\left[\sum_{n>0}\alpha^\mu_{-n}\alpha^\mu_n - \frac{D-2}{24}\right] = \frac{4}{\alpha'}\left[\sum_{n>0}\bar{\alpha}^\mu_{-n}\bar{\alpha}^\mu_n - \frac{D-2}{24}\right], \tag{2.1.10}$$

where the second equality is due to the level matching constraint. We find the following states:

- $|0\rangle$: Acting with no oscillators results in a scalar field $\mathcal{T}$ in spacetime, whose equation of motion is

$$(\partial_\mu + iV_\mu)\mathcal{T} = \frac{D-2}{6\alpha'}\mathcal{T}, \quad \text{or} \quad \partial^2\mathcal{T} - 2V_\mu\partial^\mu\mathcal{T} = -\frac{4}{\alpha'}. \tag{2.1.11}$$

  This equation of motion comes from a scalar field with a mass term $m^2 = -\frac{4}{\alpha'}$, that is, a tachyon. Its mass is independent of $D$, and therefore it is still there at the critical dimension $D = 26$ even if $V = 0$; it is the familiar tachyon of closed bosonic string theory [6, 11, 10]. We will come to condensation of this tachyon later in the chapter.

- $\alpha^\mu_{-1}\bar{\alpha}^\nu_{-1}|0\rangle$: These correspond to a symmetric rank two tensor (the metric), an antisymmetric rank 2 tensor (the $B$ field) and a scalar (the dilaton itself, $\Phi$). The mass formula now reads

$$p^2 + 2iV\cdot p = 0, \tag{2.1.12}$$

  which means that the corresponding fields are massless.

In all, the qualitative features of the low energy spectrum of the bosonic theory are not affected by the inclusion of extra dimensions.

### 2.1.4 Heterotic $HO^{(+n)/}$ theory

This theory was introduced by Hellerman in [4] and dubbed $HO^{(+n)/}$ theory. It is a higher dimensional analog of the $SO(32)$ heterotic string theory, being related to it via tachyon condensation as we will show explicitly later on.

The hallmark of heterotic string theories is the presence of $(0,1)$ supersymmetry in the worldsheet [12]. There are two multiplets of $(0,1)$ supersymmetry in two dimensions,

- The scalar multiplet composed of a boson and a right-moving fermion (and corresponding superfield $\mathcal{X} \equiv X + \theta_+\lambda$).

- A fermion multiplet composed of a left-moving fermion and an auxiliary field $F$ with no dynamics (with superfield $\lambda \equiv \lambda + \theta_+ F$). This multiplet is also called the Fermi multiplet.

The worldsheet content in the fermionic formulation [10] thus consists of $D = 10+n$ embedding coordinates $X^\mu$, and their right-moving superpartners $\psi^\mu$, plus a set of $32+n$ left-moving fermions $\lambda^a$. It also has a linear dilaton gradient

$$V_\mu V^\mu = \frac{10-D}{4\alpha'} = -\frac{n}{4\alpha'} \tag{2.1.13}$$

to ensure invariance under infinitesimal Weyl transformations. As it stands, however, the theory is not consistent as a string theory: it is not modular invariant, i.e. it is not invariant under large diffeomorphisms of the worldsheet [6, 10]. This property is essential for unitarity of perturbative string scattering amplitudes. There may be several different ways to achieve modular invariance with the same worldsheet field content [6] which will result in different spacetime theories[1]. In the $HO^{(+n)/}$, modular invariance is achieved by gauging the worldsheet R-parity $g_1$ which acts on the fields of the theory as shown in table 2.1. This results in a diagonal partition function which is automatically modular invariant [13]. Gauging a worldsheet symmetry allows new boundary conditions for the fields, such as

$$X(0,\tau) = g_1 X(\ell,\tau). \tag{2.1.14}$$

These are realized as new sectors in the theory, called twisted sectors. Modular invariance also requires us to sum over all these sectors. In particular, spin fields may be odd (NS boundary conditions) or periodic (R boundary conditions) as we follow a closed loop on the worldsheet, and both have to be taken into account. Often all left and right-moving fields have the same periodicity and thus is customary to talk about RR sectors, R-NS sectors, and so on. In this particular case, $g_1$ forces all spin fields to be either periodic or antiperiodic at the same time, so we will only have RR and NS NS sectors.

Gauging a discrete symmetry also constrains the states of the theory to be invariant under the action of the symmetry. Therefore, some of the states of the parent theory are projected out [10]. The spectrum of the theory may be obtained as before from the mass formulae, which will be different for each twisted sector. We have

- Untwisted (NSNS) sector: In this sector mass formulae are

$$\begin{aligned} -(p+iV)^2 &= \frac{4}{\alpha'}\left[\sum_{n>0} n\alpha^\mu_{-n}\alpha^\mu_n + \sum_{a=1}^{32+n}\left(n+\frac{1}{2}\right)\lambda^a_{-n-\frac{1}{2}}\lambda^a_{n+\frac{1}{2}} - 1 - \frac{n}{16}\right] \\ &= \frac{4}{\alpha'}\left[\sum_{n>0} n\bar{\alpha}^\mu_{-n}\bar{\alpha}^\mu_n + \left(n+\frac{1}{2}\right)\bar{\psi}^a_{-n-\frac{1}{2}}\bar{\psi}_{n+\frac{1}{2}} - \frac{1}{2} - \frac{n}{16}\right]. \end{aligned} \tag{2.1.15}$$

The lowest-lying excitations in this sector are:

[1] For instance, the $HO^{(+n)}$ which will be studied next has the same worldsheet content as the $HO^{(+n)/}$ and only differs in the way modular invariance is achieved.

– $\lambda^a_{-1/2}|0\rangle$, with mass term $m^2 = -2/\alpha'$. These fields constitute a tachyon $\mathcal{T}^a$ with an $SO(32+n)$ perturbative symmetry, which becomes a gauge symmetry in spacetime.
– $\lambda^a_{-1/2}\lambda^b_{-1/2}\bar{\psi}^\mu_{-1/2}|0\rangle$, whose mass term vanishes. These are $SO(32+n)$ gauge fields $A_\mu$.
– $\alpha^\mu_{-1}\bar{\psi}^\nu_{-1/2}|0\rangle$: From these massless states we get the usual NS-NS fields, namely the metric $g_{\mu\nu}$, $B$ field, and dilaton $\Phi$.

- Twisted (RR) sector: In this sector mass formulae are

$$\begin{aligned}-(p+iV)^2 &= \frac{4}{\alpha'}\left[\sum_{n>0} n\alpha^\mu_{-n}\alpha^\mu_n + \sum_{a=1}^{32+n}\left(n+\frac{1}{2}\right)\lambda^a_{-n-\frac{1}{2}}\lambda^a_{n+\frac{1}{2}} + 1\right] \\ &= \frac{4}{\alpha'}\left[\sum_{n>0} n\bar{\alpha}^\mu_{-n}\bar{\alpha}^\mu_n + \left(n+\frac{1}{2}\right)\bar{\psi}^a_{-n-\frac{1}{2}}\bar{\psi}_{n+\frac{1}{2}}\right]. \end{aligned} \tag{2.1.16}$$

Every state in this sector has a mass squared of at least $4(1+n/16)/\alpha'$, and therefore only contains massive states.

As it stands, the theory does not contain spacetime fermions. These arise from mixed NS-R and R-NS sectors, which are absent in the theory. However, there is a particular $\mathbb{Z}_2$ orbifold of the $HO^{(+n)/}$ theory which contains fermions living in a 10d submanifold in spacetime. "Orbifolding" here means that the worldsheet discrete gauge group is enlarged, in this case to include an action $g_2$ which is described in table 2.1. Since this breaks the Lorentz symmetry down to $SO(1,9)\times SO(n)$, we use indices $\mu = 0,\ldots,9$, $m = 10,\ldots,n$. Notice that this Lorentz symmetry was already broken by the linear dilaton anyway.

| | $HO^{+(n)/}$ | $HO^{+(n)/}/\mathbb{Z}_2$ |
|---|---|---|
| Field | $g_1$ | $g_2$ |
| $X^\mu$ | $+$ | $+$ |
| $X^m$ | $+$ | $-$ |
| $\tilde{\psi}^\mu$ | $-$ | $+$ |
| $\tilde{\psi}^m$ | $-$ | $-$ |
| $\lambda^a$ | $-$ | $-$ |

**Table 2.1:** Charge assignments for the worldsheet discreete gauge group for the $HO^{(+n)/}$ theory and its orbifold version. For convenience we use the indices $\mu = 0,\ldots,9$, and $m = 9,\ldots D-1$.

Gauging $g_2$ has the effect of introducing $g_2$-twisted sectors, as well as forcing the original $HO^{(+n)/}$ states to be orbifold-invariant. In other words, states must be invariant under the combination $X^m \to -X^m$, $\psi^m \to -\psi^m$, $\lambda^a \to -\lambda^a$. If a particular field does not vanish at the fixed locus $X^m = 0$ of the orbifold, its normal derivative will since it carries an extra index affected by the orbifold. This forces either Dirichlet or Neumann boundary conditions for the untwisted sector fields at $X^m = 0$, which are listed in table 2.2.

| Field | Boundary condition |
|---|---|
| $\mathcal{T}^a$ | Dirichlet |
| $g_{\mu\nu}$, $B_{\mu\nu}$, $\Phi$ | Neumann |
| $g_{\mu m}$, $B_{\mu m}$ | Dirichlet |
| $g_{nm}$, $B_{nm}$ | Neumann |
| $A^{ab}_{\mu}$ | Neumann |
| $A^{ab}_{m}$ | Dirichlet |

**Table 2.2:** Boundary conditions for the bulk fields of the $HO^{(+n)/}$ as a result of the $g_2$ orbifold [4].

The boundary conditions in the twisted sectors force the center-of-mass term in the $X^m$ oscillator expansion to vanish. As a result, the twisted fields are localized to the $X^m = 0$ plane and do not propagate in the bulk. We may think of the $HO^{+(n)/}/\mathbb{Z}_2$ as the regular $HO^{+(n)/}$ theory containing a 10-dimensional "brane" which carries localized degrees of freedom. There are two new sectors, whose field content is:

- $g_2$-twisted sector: This sector contains NS fields $(\bar{\psi}^\mu)$, as well as R fields $\lambda^a$, $\bar{\psi}^m$. It is therefore a R-NS sector of a sort. Mass formulae are

$$-(p+iV)^2 = \frac{4}{\alpha'}\left[\sum_{n>0} n\alpha^\mu_{-n}\alpha^\mu_n + \sum_{a=1}^{32+n}\left(n+\frac{1}{2}\right)\lambda^a_{-n-\frac{1}{2}}\lambda^a_{n+\frac{1}{2}} + 1 + \frac{n}{16}\right]$$
$$= \frac{4}{\alpha'}\left[\sum_{n>0} n\bar{\alpha}^\mu_{-n}\bar{\alpha}^\mu_n + \left(n+\frac{1}{2}\right)\bar{\psi}^a_{-n-\frac{1}{2}}\bar{\psi}_{n+\frac{1}{2}} - \frac{1}{2} + \frac{n}{16}\right]. \quad (2.1.17)$$

  States in this sector have a mass of at least $4(1+n/16)/\alpha'$.

- $g_1g_2$-twisted sector: This sector is an NS-R sector. Zero point energies are $-1$ for left movers and 0 for right movers. We also have a set of Ramond zero modes, spanned by $\bar{\psi}^\mu_0$. As usual [6, 11, 10], this gives rise to spacetime spinors. In our case, we only have Ramond zero modes for the coordinates $\mu = 0, 1, \ldots 9$, which span the orbifold fixed locus. Fields in this sector will be 10d spinors of this fixed locus. Lowest-lying states are:
  - $\lambda^a_{-1/2}\lambda^b_{-1/2}\{\bar{\psi}^\mu_0\}|0\rangle$, where the notation $\{\bar{\psi}^\mu_0\}|0\rangle$ means that any number of Ramond zero modes may be acting on the vacuum. We must impose invariance under $g_1$ and $g_2$, which constrains the number of Ramond zero modes to be odd. This fixes a particular chirality (say left-handed) for the 10d spinors [10]. The resulting state is a spinor charged in the adjoint of the bulk gauge group.
  - $\lambda^a_{-1/2}\bar{\alpha}^m_{-1/2}\{\bar{\psi}^\mu_0\}|0\rangle$. Imposing orbifold invariance results in a spin field of right-handed chirality. This spin field is in the vector of the gauge group, and also in the vector representation of the $SO(n)$ acting on the normal coordinates $m = 10, 11, \ldots 9+n$. We may say that this field is a section of

the tensor product of the gauge and the tangent bundle to the orbifolded coordinates.

- $\alpha^{\mu}_{-1}\{\bar{\psi}^{\mu}_{0}\}|0\rangle$. This state is a 10d vector spinor, of left- handed chirality, which may be decomposed into a gravitino[2] (pure spin 3/2) and a spinor of right-handed chirality.
- $\alpha^{m}_{-1/2}\bar{\alpha}^{n}_{-1/2}\{\bar{\psi}^{\mu}_{0}\}|0\rangle$. This state is another spinor, of left-handed chirality, living in a two-index symmetric representation of the normal $SO(n)$ gauge group.

The above matter content is very similar to that of type I theory in the presence of $n$ extra $D9-\overline{D9}$ brane pairs. This observation was used in [4] as grounds for proposing an $S$-duality between these two theories, generalizing the well-known duality between type I and heterotic $SO(32)$ theory. A detailed analysis of the spectrum of type I theory can be found in [14, 15]. The duality is not straightforward; for instance, heterotic bulk fields propagate in $10+n$ dimensions, whereas all type I fields propagate in 10 dimensions. Thus if the duality is to hold it has to involve the restriction of bulk heterotic fields to the orbifold fixed locus.

One may also see [4] that the low-energy effective action in the heterotic side is consistent with the proposal. This $S$-duality proposal will be one of the main motivations in Chapter 4 of the construction of the supercritical NS5-brane as a closed-string tachyon soliton.

### 2.1.5 Type $HO^{(+n)}$ heterotic theory

The $HO^{(+n)/}$ theory described in the previous section has gauge group $SO(32+n)$ and as such is a natural generalization of the heterotic $SO(32)$ theory. However, modular invariance prevents us from making a similar construction for the $E_8\times E_8$ string. To obtain this gauge group, one would have to partition the 32 left-moving fermions into two sets of 16 and specify a different action of the worldsheet gauge group on them; a diagonal modular invariant theory like the $HO^{(+n)/}$ can never achieve this.

In [4], another modular-invariant heterotic theory was introduced, the so called $HO^{(+n)}$ theory. Its worldsheet field content and linear dilaton background is exactly the same as for the $HO^{(+n)/}$ theory, and only differs on the choice of discrete worldsheet gauge group. This new action distinguishes between some of the left movers and accordingly we will relabel the left-moving fermions as $\lambda^a$, $a=1,\ldots 32$, and $\chi^b$, $b=1,\ldots n$.

The discrete worldsheet gauge group is no longer $\mathbb{Z}_2$, as it was in the bulk $HO^{(+n)/}$ theory, but rather $\mathbb{Z}_2\times\mathbb{Z}_2$, with two generators $g_1$ and $g_2$ whose action

[2] Strictly speaking there can be no gravitino in this theory, which lacks supersymmetry. However, as we will see later on, this theory is connected via tachyon condensation to the $SO(32)$ heterotic string theory. The heterotic gravitino descends from this supercritical spin 3/2 field, hence the name.

| Field | $g_1$ | $g_2$ |
|---|---|---|
| $X^\mu$ | + | + |
| $\bar{\psi}^\mu$ | + | − |
| $\lambda^a$ | − | + |
| $\chi^b$ | + | − |

**Table 2.3:** Charge assignments for the worldsheet discrete gauge group for the $HO^{(+n)}$ theory. For convenience we use the indices $a = 1, \ldots 32$, and $\chi^b$, $b = 1, \ldots n$.

on worldsheet fields can be found in table 2.3. In accordance with the discussion in the previous section, the theory has four twisted sectors, which must be analyzed separately:

- Untwisted sector: The field content and boundary conditions in this sector are the same as in the untwisted sector of the $HO^{(+n)/}$ theory discussed above, so the zero point energies are the same: $E_0 = -1 - n/16$, $\bar{E}_0 = -\frac{1}{2} - n/16$. The different worldsheet discrete gauge group however selects a different set of physical states:
  - $\chi^b_{-1/2}|0\rangle$: This is a tachyon living in the vector of $SO(n)$, in contrast with the tachyon of the $HO^{(+n)/}$ theory which lived in the vector of $SO(32+n)$.
  - $\lambda^a_{-1/2}\lambda^b_{-1/2}\bar{\psi}^\mu|0\rangle$: These massless states funish the adjoint representation of $SO(32)$ and correspond to gauge bosons.
  - $\alpha^\mu_{-1}\bar{\psi}^\mu|0\rangle$: These states are also massless and give us the usual dilaton+$B$-field+metric.
- $g_1$-twisted: In this sector, the $\lambda^a$'s are periodic, and so the left-moving zero point energy is $E_0 = 1 - n/16$. Taking into account the linear dilaton contribution, this means that all states in this sector have a mass of at least $4/\alpha'$ and are therefore very massive.
- $g_2$-twisted: In this sector, $E_0 = -1$, and $\bar{E}_0 = 0$. Since in this sector there are periodic Ramond fields, it may contain massless fermions. The orbifold projection selects the following states:
  - $\alpha^\mu_{-1}\{\chi^b_0\}\{\bar{\psi}^\mu_0\}|0\rangle$, where the total number of R zero modes is odd. Notice that left and right Ramond zero modes together furnish the Clifford algebra of $SO(1, 8 + n)$, and thus this is a vector-spinor, which may be further decomposed into spin 3/2 and spin 1/2 fields.
  - $\chi^b_{-1}\{\chi^b_0\}\{\bar{\psi}^\mu_0\}|0\rangle$, where the total number of R zero modes is even, since the $\chi^b_{-1}$ field is also seen by the $g_2$ projector. This states are spacetime spinors of opposite chirality to those discussed above.
  - $\lambda^a_{-1/2}\lambda^b_{-1/2}\{\chi^b_0\}\{\bar{\psi}^\mu_0\}|0\rangle$, where the total number of R zero modes is odd again. These would-be "gauginos" are sections of the adjoint representation of the $SO(32)$ bundle, and also spacetime spinors.

- $g_1g_2$-twisted: Similarly to the $g_1$-twisted sector in the $HO^{(+n)/}$ theory, all left-moving fermions are now periodic. Hence $E_0 = 1 + n/16$ and all states in this sector are very massive.

Unlike its cousin $HO^{(+n)/}$, this theory already contains spacetime fermions without adding an orbifold with a geometric action.

Finally, we note that all of the above construction may be repeated step by step to yield a supercritical version of the heterotic $E_8 \times E_8$ string. To do this [10], we substitute the $g_1$ action $\lambda^a \to -\lambda^a$ by two independent actions $\lambda^a \to -\lambda^a$, $\lambda^{a'} \to -\lambda^{a'}$ where we have partitioned the $\lambda^a$ into two sets of 16 left-moving fermions. The discrete worldsheet gauge group is now $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, but one can check that the resulting spectrum is exactly the same as above except for the fact that the gauge group changes from $SO(32)$ to $E_8 \times E_8$. This flexibility of the $HO^{(+n)}$ theory will allow us to discuss together the two different gauge groups simultaneously.

### 2.1.6 Type 0/type II theory

One can also describe [5] a supercritical string theory related to type II strings. However, the construction is much less straightforward than in the bosonic and heterotic cases, because type II GSO projections (necessary to achieve a modular invariant theory) do not give a modular invariant theory unless the spacetime dimension $D$ satisfies $D \equiv 2 \mod 8$. This may be seen from the partition function of type II strings (this is a one-loop vacuum amplitude evaluated on a torus, on which modular transformations act as the Möbius group on the complex structure $\tau$ of the torus) which can be constructed from the quantity [10]

$$\begin{aligned} Z_{\pm}(\tau) =& \frac{1}{2}(4\pi^2\alpha'\tau_2)^{-\frac{D-2}{2}}\eta^{-D-2}\eta^{-\frac{D-2}{2}} \cdot \\ & \left(\vartheta\begin{bmatrix}0\\0\end{bmatrix}^{-\frac{D-2}{2}} - \vartheta\begin{bmatrix}0\\1/2\end{bmatrix}^{-\frac{D-2}{2}} - \vartheta\begin{bmatrix}1/2\\0\end{bmatrix}^{-\frac{D-2}{2}} \pm \vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}^{-\frac{D-2}{2}}\right). \end{aligned} \tag{2.1.18}$$

Type II A strings have the partition function $Z_+\bar{Z}_-$, while IIB strings have $Z_+\bar{Z}_+$ as their partition function. Under modular transformations, the theta function factors acquire phases $e^{-\frac{i\pi}{4}}$ which must be a sign for the theory to be modular invariant.

There is another kind of GSO projection with the worldsheet spectrum of type II which yields a modular invariant theory in an arbitrary number of dimensions, that of type 0 theory. The theory can be made conformal invariant in an arbitrary number of dimensions by adding a linear dilaton with

$$V_\mu V^\mu = \frac{10 - D}{4\alpha'} = -\frac{n}{2\alpha'}. \tag{2.1.19}$$

Much like in our supercritical heterotic string, the GSO projection of type 0 theory can be described as the gauging of a global worldsheet symmetry, namely fermion number $(-1)^F$. As a result, we again have only NSNS and RR sectors. The spectrum in $10 + 2n$ dimensions may be computed by now familiar techniques:

- In the NS-NS sector, zero point energies are $E_0 = \bar{E}_0 = -\frac{8+2n}{16}$. There are only two kinds of low-energy states satisfying the GSO projection: $|0\rangle$, which as usual yields a tachyon with $m^2 = -\frac{2}{\alpha'}$, and $\psi^\mu_{-1/2}\bar{\psi}^\nu_{-1/2}|0\rangle$, which corresponds to the usual massless NS-NS fields $\Phi$, $g_{\mu\nu}$, $B_{\mu\nu}$.

- In the RR sector, zero point energies vanish. We also have two sets of Ramond zero modes, which can be combined to yield a generic bispinor state $\{\psi^\mu_0\}\{\bar{\psi}^\mu_0\}|0\rangle$. There is a sign ambiguity in the action of the GSO projection (an elementary example of discrete torsion [6, 16]), which may be taken as $\frac{1}{2}(1-(-1)^F)$, resulting in the so-called type 0A theory, or we may take $\frac{1}{2}(1+(-1)^F)$ to yield 0B theory. The only difference is that in 0A the bispinor is constrained to have an odd number of fermion operators, whereas in 0B it must have an even number. In either case, a bispinor field is a reducible representation of the Lorentz group. It can be shown [12] that a bispinor $\chi_1\chi_2$ in even spacetime dimension (henche the notational change from $10+n$ to $10+2n$) whose spinors $\chi_1, \chi_2$ have the same chirality decomposes into a tower of even rank $p$ forms, from 0 to $D-2/2$. For $p = (D-2)/2$ the $p$-form is (anti)self-dual, depending on the chirality of the spinors. If the spinors have opposite chiralty, the bispinor decomposes into a tower of odd-rank $p$-forms from $p = 1$ to $p = (D-4)/2$. Thus, the RR spectrum of each theory is:
  - For 0A, the states surviving the orbifold projection comprise two bispinors with $\chi_1, \chi_2$ having opposite chiralities. Hence we get a pair of $p$-forms, from $p = 1$ to $p = (D-4)/2$.
  - For 0B, we get two bispinors with $\chi_1, \chi_2$ of the same chirality, one positive and another negative. Thus we get pairs of $p$-forms from $p = 0$ to $p = (D-6)/2$, and also one self-dual and one anti-selfdual $(D-2)/2$-forms, which may be combined in a single nonchiral $(D-2)/2$-form.

  We can recognize the RR sector of type 0 theories as a doubled version of the corresponding type II RR sector [17].

  There is a last point which requires further explanation. The fields in the RR sector are bosonic, obeying second-order equations of motion. From the discussion in section 2.1.2 one would think that they would have a mass term $-V_\mu V^\mu = n/(2\alpha')$, and hence be massive. The spacetime action for a massive $p$-form is generically of the form

$$S = \int e^{-2\Phi}[F \wedge *F - V_\mu V^\mu C \wedge *C]. \tag{2.1.20}$$

  This is not an acceptable action coming from string tree-level amplitudes, since [12, 18] fundamental strings do not carry RR charge and hence must couple only to RR fieldstrengths $F$, with no terms in the RR potential being allowed.

| Field | $g_1$ | $g_2$ |
|---|---|---|
| $X^\mu$ | $+$ | $+$ |
| $X^m$ | $+$ | $-$ |
| $\psi^\mu$ | $-$ | $-$ |
| $\psi^m$ | $-$ | $+$ |
| $\bar{\psi}^\mu$ | $-$ | $+$ |
| $\bar{\psi}^m$ | $-$ | $-$ |

**Table 2.4:** Charge assignments for the worldsheet discreete gauge group for the type 0 theory and its orbifold version. For convenience we use the indices $\mu = 0, \ldots, 8+n$, and $m = 9+n, \ldots D-1$.

However, only for the particular mass (2.1.20), one may redefine $C' = e^{-\Phi}C$, in terms of which $F = e^{\Phi}[F' + d\Phi \wedge C']$ and so

$$S = \int F' \wedge *F' + [*(d\Phi \wedge *d\Phi) - V_\mu V^\mu] C' \wedge *C' \\ + F' \wedge *(d\Phi \wedge C') + (d\Phi \wedge C') \wedge *F'. \tag{2.1.21}$$

Now, the last two terms may be rewritten as $2 * d\Phi \wedge d(C' \wedge *C')$, and since $d * d\Phi = 0$ for the linear dilaton background, the last terms are a total derivative. Since $*(d\Phi \wedge *d\Phi) = V_\mu V^\mu$, we see that the field $C'$ is actually massless, and does not couple to the dilaton at tree level. This agrees with computations coming directly from worldsheet amplitudes [18, 19], which also provide a geometrical understanding of this peculiar behaviour: RR vertex operators are built cutting a hole in the worldsheet and specifying that fields must be antiperiodic around it. Each such hole lowers the Euler characteristic of the Riemann surface by 1, and since the kinetic term of the gauge fields is quadratic, a tree-level computation requires two such holes in the sphere, which gives an effective Euler characteristic of 0. Since the dilaton power appearing in front of the action is $e^{-\chi\Phi}$, where $\chi$ is the Euler characteristic of the surface on which the amplitude is evaluated, kinetic terms for RR $p$-forms do not have the ordinary dilaton dependence.

Type 0 theories do not have spacetime fermions. Much as in the heterotic case, this can be traced to the choice of diagonal GSO projection, which ensures that left-moving worldsheet spinor zero modes (responsible for spacetime fermions [10]) always show up together with right-moving zero modes. Fermions may be introduced, as in the heterotic case, by further orbifolding the theory in a way which breaks target spacetime Lorentz invariance. We will orbifold by the worldsheet action $g_2$, whose action on different worldsheet fields is given in table 2.4.

The geometric action of the orbifold consists in flipping $n$ of the coordinates. As in the heterotic case, this acts on the bulk fields of the theory, imposing various boundary conditions, which are summarized in table 2.5.

Having discussed the bulk fields, we now turn to the localized fermions which live in the $10+n$ fixed locus of the orbifold. As before, the only real issue is to

| Field | Boundary condition at $X^m = 0$ |
|---|---|
| $\mathcal{T}^a$ | Dirichlet |
| $g_{\mu\nu}$, $B_{\mu\nu}$, $\Phi$ | Neumann |
| $g_{\mu m}$, $B_{\mu m}$ | Dirichlet |
| $g_{nm}$, $B_{nm}$ | Neumann |
| $C_{\mu\nu\ldots mn\ldots}$, $\#m, n, \ldots$ =even | Neumann |
| $C_{\mu\nu\ldots mn\ldots}$, $\#m, n, \ldots$ =odd | Dirichlet |

**Table 2.5:** Boundary conditions for the bulk fields of the type 0 as a result of the $g_2$ orbifold. Last lines refer to RR fields.

compute the zero point energies of the different sectors. In this case details are more subtle, so we compute zero point energies in detail:

- $g_2$-twisted: In this twisted sector

$$E_0 = \underbrace{-\frac{1}{24}\cdot(8+n)}_{\text{Untw. bosons}} + \underbrace{\left(-\frac{1}{24}+\frac{1}{16}\right)\cdot n}_{\text{Tw. bosons}} + \underbrace{\left(\frac{1}{24}\right)\cdot(8+n)}_{\text{Tw.fermions}} + \underbrace{\left(\frac{1}{24}-\frac{1}{16}\right)\cdot n}_{\text{Tw.fermions}} = 0,$$
$$\bar{E}_0 = \underbrace{-\frac{1}{24}\cdot(8+n)}_{\text{Untw. bosons}} + \underbrace{\left(-\frac{1}{24}+\frac{1}{16}\right)\cdot n}_{\text{Tw. bosons}} + \underbrace{\left(\frac{1}{24}-\frac{1}{16}\right)\cdot(8+n)}_{\text{Unt.fermions}} + \underbrace{\left(\frac{1}{24}\right)\cdot n}_{\text{Tw.fermions}} = -\frac{1}{2}. \tag{2.1.22}$$

We have $SO(8+n)$ fermion zero modes from the left-movers, and $SO(n)$ fermion zero modes from the right movers. Let us study the $GSO$ projection in more detail. To this effect, let $g^R_X$ be the operator which swaps the sign of right-moving $X^\mu$ fermions, $g^L_Y$ the one which does the same for left-handed $\psi^m$ fermions, and so on. Let $\alpha^R_X$ be similarly defined for bosons. Then the GSO+orbifold projection is

$$P_{GSO+orb.} = \frac{1}{4}(1 + g^R_Y g^L_Y g^R_X g^L_X)(1 - g^L_X \alpha^R_Y \alpha^L_Y g^R_Y). \tag{2.1.23}$$

If we restrict to massless states, it is clear that $\alpha^R_Y$ and $g^R_Y$ act trivially. Furthermore level-matching implies $\alpha^L_Y g^X_L = -1$. Substituting above, the GSO for the massless sector is simply

$$P_{GSO+orb.} = \frac{1}{2}(1 + g^L_Y g^R_X g^L_X). \tag{2.1.24}$$

The states surviving this projection are bispinors of $SO(8+n)\times SO(n)$, with overall chirality determined by the eigenvalue of $g^L_X$. That is, we will have one vector-bispinor of overal $SO(8+n)\times SO(n)$ negative chirality, and another bispinor of $SO(8+n)\times SO(n)$, of opposite chirality, and charged as a vector of $SO(n)$.

- $g_1 g_2$-twisted:

$$E_0 = \underbrace{-\frac{1}{24}\cdot(8+n)}_{\text{Untw. bosons}} + \underbrace{\left(-\frac{1}{24}+\frac{1}{16}\right)\cdot n}_{\text{Tw. bosons}} + \underbrace{\left(\frac{1}{24}-\frac{1}{16}\right)\cdot(8+n)}_{\text{Tw.fermions}} + \underbrace{\left(\frac{1}{24}\right)\cdot n}_{\text{Tw.fermions}} = -\frac{1}{2},$$
$$\bar{E}_0 = \underbrace{-\frac{1}{24}\cdot(8+n)}_{\text{Untw. bosons}} + \underbrace{\left(-\frac{1}{24}+\frac{1}{16}\right)\cdot n}_{\text{Tw. bosons}} + \underbrace{\frac{1}{24}\cdot(8+n)}_{\text{Unt.fermions}} + \underbrace{\left(\frac{1}{24}-\frac{1}{16}\right)\cdot n}_{\text{Tw.fermions}} = 0. \tag{2.1.25}$$

Upon detailed analysis, this sector turns out to give the same spectrum as the $g_2$-twisted. The sign of the $g_2$ projection ensuring invariance of the states under the orbifold action is related via modular invariance to the discrete torsion specified above when discussing 0A and 0B theories.

This completes the survey of the supercritical string theories which will be employed in this thesis. For more details, the reader can consult [3, 4, 5, 20].

## 2.2 Tachyon condensation

All of the supercritical string theories described in the previous section share two outstanding features: They are not Lorentz-invariant due to the presence of a linear dilaton background, and have a tachyon in their perturbative spectrum. In this section we will see that the two are closely related, by following the tachyon condensation procedure to its lowest energy state, which will turn out to be a (possibly different) string theory in a different number of spacetime dimensions. First we will briefly describe open tachyon condensation in D-brane systems. We then proceed to analyze in detail the closed string tachyon condensation taking place in the different supercritical string theories.

### 2.2.1 Open string tachyon condensation

As is well-known, type II and type I string theories admit sectors with open strings [12, 10]. The spacetime loci where these open strings may end are known as D-branes. These extended states are one of the outstanding features of string theory and one of our main handles on its nonperturbative physics.

It turns out [12, 10] that these D-branes are dynamical objects of their own. They turn out to be BPS states [12] and carry charges under the RR fields of the closed-string sector of the theory. As such, CPT invariance guarantees the existence of anti D-branes, which we will call $\overline{Dp}$-branes. These branes have opposite charges with respect to RR fields and preserve exactly those supercharges not preserved by ordinary D-branes. As a result, a $Dp - \overline{Dp}$ system carries the conserved charges of the vacuum, and in principle nothing prevents the system to decay into this state (or rather, to a bunch of perturbative excitations over the vacuum).

Interactions of a system of two D-branes are described in terms of the open strings stretching between them, and this is also true of a $Dp-\overline{Dp}$ system. The only difference between the two is that the sign of the $GSO$ projection is reversed. This can be ascertained by using open-closed duality, in which an open string diagram is regarded as a closed string one, and using the different RR charges of the two branes [21]. For a stack of D-branes, the action of the Chan-Paton projection on the boundary labels is also modified [14, 15]. Another way of showing all of this is by considering branes at an angle; standard quantization gives a non-supersymmetric spectrum with a tachyon. At $\theta = \pi$, the orientation of one of the branes is reversed and the brane becomes an antibrane [12]. It may seem weird that one can turn an object into an anti-object via ordinary rotations; this is not exclusive of string theory and will happen for any theory with extended objects. The charges of these objects are integrals of their worldvolume, and by swapping the sign of one of their worldvolume coordinates one can change the sign of their charges. This is of course a formal operation; one cannot rotate objects that extend to infinity for it would cost infinite energy.

The result in any case is the appearance of some fermions and, crucially, a tachyon with $m^2 = -1/\alpha'$. This tachyon is charged in a bifundamental representation of the gauge bundles of the branes: In type II it lives in the bifundamental of $U(n) \times U(m)$, whereas in type I it is a section of the bivector of $SO(32+n) \times SO(m)$.

This open string tachyon is very similar to the examples studied above: it represents an instability of the $Dp - \overline{Dp}$ system. Although the whole process is very similar to the Higgs mechanism, there are some differences. In a field-theoretic Higgs mechanism, the scalar getting the vev only gives mass to fields under which it is charged. In the present case (take type II and $m = n = 1$ for simplicity), the tachyon is only charged under the antidiagonal $U(1)$, while its $U(1)_{\text{diag.}}$ charge vanishes. So $U(1)_{\text{diag.}}$ cannot be Higgsed and yet it disappears from the spectrum after tachyon condensation. It has been suggested that some sort of confinement [22] is responsible for making $U(1)_{\text{diag.}}$ massive.

One of the first, simplest arguments to find the endpoint of tachyon condensation was described by A. Sen in [23]. The idea is to use well-established dualities in string theory to relate the $Dp - \overline{Dp}$ system to a certain gauge theory, on which the endpoint of the tachyon condensation can be found using field-theoretic techniques. The endpoint of the condensation process was the string theory vacuum; the brane and antibrane annihilate, leaving behind nothing but closed-string radiation. Relying on these results, Sen conjectured that, for any brane-antibrane system [24],

- The endpoint of homogeneous tachyon condensation on the world-volume of the D-brane is the closed string vacuum.
- the condensation of inhomogeneous modes of the tachyon field leads to lower dimensional D-branes.

The second conjecture is justified by the fact that the open string tachyon admits solutions with nontrivial topological charges which must be preserved by the condensation process. Thus, tachyon condensation of these configurations can never lead to the vacuum. As an example, the tachyon effective action for type II brane-antibrane systems must be invariant under $\mathcal{T} \to e^{i\theta}\mathcal{T}$, since this is just a global gauge transformation. As a result, if the vacuum expectation value of the tachyon is $\mathcal{T}_{min}$, a profile which winds around some $\mathbb{R}^2$ at infinity $n$ times, i.e. is of the form

$$\mathcal{T}(r \to \infty) = e^{in\theta}\mathcal{T}_{min} \tag{2.2.1}$$

where $r$ is a radial coordinate in two dimensions, will asymptote to a configuration which is the vacuum everywhere except at the vicinity of $r = 0$; in this region a nontrivial field profile will carry the winding charge. Since the real codimension of this object in the brane-antibrane system is 2, it has the right dimensionality to describe a $D(p-2)$ brane; this is further supported by the fact that the profile (2.2.1) requires some flux turned on in the antidiagonal $U(1)$ of the brane-antibrane system, in order to have a finite energy configuration. This worldvolume flux will couple to a RR $(p-2)$-form field, via the Chern-Simons couplings on the brane [12]

$$\int_{Dp} F \wedge C_{p-2}. \tag{2.2.2}$$

If the field configuration sources $C_{p-2}$ before tachyon condensation, it must continue to do so afterwards. As a result, this codimension 2 object carries precisely the RR charge of a $D(p-2)$ brane. If it smells like a $D(p-2)$ brane and tastes like a $D(p-2)$ brane, it probably is a $D(p-2)$ brane.

It is difficult to provide solid evidence for these conjectures, since they involve nonperturbative physics. String theory is only defined (if we exclude dualities for the time being) in a perturbative manner, as a set of rules for computing scattering amplitudes. To study tachyon condensation, we will need to study off-shell configurations, in a non-perturbative way. This may achieved via the use of string field theory, a formalism which second-quantizes open strings to yield a field theory with a well-defined action and in which nonperturbative questions may be addressed.

For a long time, evidence for the first Sen conjecture was based on numerical approximations [24, 25]. However, in 2005 Schnabl [26] found an exact tachyon profile corresponding to the energy minimum, proving analitically Sen's first conjecture. [27] proposes a dynamical solution describing explicitly the condensation process. Numerical verification of Sen's second conjecture is also possible, see e.g. [28], and also in type II superstrings [29, 30].

### 2.2.2 Closed-string tachyon condensation

As discussed above rigorous treatment of open-string tachyon condensation relies heavily on open string field theory. The problem of closed-string tachyon condensation is much more difficult due to the fact that no simple closed string field theory is known (in fact any covariant closed string field theory would have an infinite number of interaction vertices [31]) and also because the spacetime picture here is not clear at all. Since closed-string tachyons propagate in the bulk of spacetime, closed-string tachyon condensation should be somehow related to the decay of spacetime itself, although there are some examples [32] in which closed-string tachyons living in a fixed locus in spacetime have much milder behaviour.

In [4, 5], Hellerman and Swanson showed how, in supercritical string theories, closed-string tachyon condensation in supercritical string theories can be studied even in the absence of closed string field theory. Their ideas are the main tool employed in the first part of this thesis.

### 2.2.3 Tachyon condensation in bosonic string theory

The simplest example to show the mechanism of closed-string tachyon condensation is the bosonic string theory studied in section 2.1.3. We saw that the spectrum of the supercritical string theory included a tachyon $\mathcal{T}$ of mass squared $-2/\alpha'$. We will study tachyon condensation by turning on a nontrivial tachyon profile, which also satisfies the equations of motion, and analyzing its effects on a probe worldsheet.

We first need to discuss how the tachyon background couples to the worldsheet. A general recipe is that a background for any field can be coupled to the string in a unique way, by means of the state-operator mapping. The background is a semiclassical, i.e. coherent state and will therefore correspond to a certain operator in the worldsheet. The standard prescription is that [12], whenever we have a classical background for a field $\phi_b(x)$, with vertex operator $V(X)$, the corresponding worldsheet interaction is (up to some constants)

$$\mathcal{S}_{\text{int.}} =: \exp(\phi_b(X)V(X)) : \tag{2.2.3}$$

so the vertex operator determines uniquely in which way can any background couple to the string. For closed-string tachyons, the vertex operator is the identity [6], and so a nontrivial tachyon background is seen in the CFT as an extra term

$$\Delta\mathcal{L} = -\frac{1}{2\pi} : \mathcal{T}(X) : \tag{2.2.4}$$

The fact that $\mathcal{T}(X)$ obeys the spacetime equations of motion ensures that the perturbation (2.2.4) is marginal and the resulting theory is still a CFT. We will be interested in the case of lightlike[3] tachyon condensation with nontrivial dependence along some coordinate. For instance, a profile of the form

$$\mathcal{T} = \mu_0^2 \exp(\beta X^+) - \mu_k^2 \cos(kX^2) \exp(\beta_k X^+) \tag{2.2.5}$$

is a solution of the tachyon equation of motion (2.1.11) if

$$\beta = -\sqrt{8}/(V_0\alpha'), \quad \beta_k = -\frac{\sqrt{2}}{V_0}\left(2/\alpha' - \frac{1}{2}k^2\right). \tag{2.2.6}$$

Here, $X^+ = (X^1 + X^0)/\sqrt{2}$, $X^- = (X_1 - X_0)/\sqrt{2}$ are a pair of light-like coordinates. The timelike linear dilaton gradient is just $(-V_0, 0, \dots, 0)$.

[3] We focus on lightlike instead of timelike tachyon condensation because the former is dissipative in nature [27] and thus we do not have to worry about "reheating" of the tachyon condensate.

The dynamics of a string propagating in the background (2.2.4) is easy to study, at least at the classical level in the limit $k^2 \ll 1/\alpha$. It is precisely in this limit where we expect (2.1.11) to be a good approximation to the exact equations of motion. In this case, $\beta_k \approx \beta$ and the tachyon potential simplifies to

$$\mathcal{T} = \frac{\mu^2}{2\alpha'} \exp(\beta X^+) X^2 - \frac{\mu^2 X^+}{\alpha' V_\mu V^\mu \sqrt{2}} \exp(\beta X^+) + (\mu_0^2 - \mu_k^2) \exp(\beta X^+). \qquad (2.2.7)$$

The first term is just a mass for $X^2$ which increases exponentially with $X^+$. The resulting physics is rather simple: At $X^+ \to -\infty$, the string is free to propagate in all $D-1$ spatial dimensions. At $X^+ \sim 0$, it starts to feel a potential which confines it to the $X^2 = 0$ slice from then on. Thus, at least for the perturbative degrees of freedom which may be described in terms of closed strings, the tachyon has the effect of removing one dimension; it is indeed causing a partial 'decay of spacetime'. The two regions are separated by a wall which moves at the speed of light along the $X^1$ direction.

The other terms in (2.2.7) can be dealt with in different ways: We will fine-tune the last term to 0 (even after quantum corrections are taken into account). The middle term cannot be fine-tuned to zero, but will cancel against a one-loop correction in the worldsheet (as it must, since it is not conformal invariant at the classical level).

A detailed analysis of the solutions of the equations of motion coming from (2.2.7) shows [5] that any state with $X^2 \neq 0$ will, upon meeting the wall, accelerate along with it and will never reach the interior, since by the time the wall reaches it has a very large potential energy which is converted into kinetic. If we are interested in the dynamics within the wall, in the region where $X^2$ is very massive, we should integrate this field out. In general, the coupling (2.2.4) defines an interacting 2d quantum field theory in which integrating out some field is a rather difficult task. However, for the particular profile (2.2.7) the classical theory is exactly solvable [5], which implies that the Feynman diagram expansion terminates at the one-loop level. Alternatively, we may expand the field $X^2$ in a basis of solutions to the equations of motion. After doing this, the path integral is gaussian. In any case, integrating out $X^2$ results in the following changes to the effective action:

- There is a correction of the form $X^+ \exp(\beta X^+)$ which precisely cancels the middle term of (2.2.7).
- There is a shift in the linear dilaton term,

$$\Phi \to \Phi + \frac{\beta}{12} X^+. \qquad (2.2.8)$$

- The metric also changes, as

$$g_{\mu\nu} \to g_{\mu\nu} + \frac{\alpha' \beta^2}{24} \delta_\mu^+ \delta_\nu^+. \qquad (2.2.9)$$

- The tachyon profile in the final state is $\mathcal{T} = 0$ since we have fine-tuned it away.

We end up with a theory in one spacetime dimension less, with modified metric and linear dilaton term. The norm of the new dilaton gradient is

$$g'_{\mu\nu}(V')^\nu(V')^\mu = V_\mu V^\mu + V_- \frac{\beta}{6} V_- + \frac{\alpha'\beta^2}{24}(V_+ + \frac{\beta}{12})^2 = V_\mu V^\mu + \frac{1}{6\alpha'}, \tag{2.2.10}$$

where we have used that $V_+ = -V_0/\sqrt{2}$, $V_- = -V_0/\sqrt{2}$. The new central charge associated to the linear dilaton is

$$c'_{\text{dilaton}} = -6\alpha' g'_{\mu\nu}(V')^\nu(V')^\mu = c_{\text{dilaton}} - 1. \tag{2.2.11}$$

This mechanism, which was dubbed 'central charge transfer' in [4], ensures that the theory is conformal at the quantum level after the condensation procedure, since we have one less boson contributing to the central charge.

To sum up, we now know the endpoint of closed-string tachyon condensation in supercritical bosonic string theory - it is another bosonic string theory with one spacetime dimension less. The mechanism, which has been dubbed "dimension quenching" [20], does not depend on $V^\mu$ being timelike, and so it can be applied also to critical and subcritical string theories - providing an interpretation of the usual bosonic string tachyon (albeit only in a very special background in which the dilaton has a lightlike constant gradient).

An obvious question now is when does all this stop? As shown in section 2.1.3 there is always a tachyon in bosonic string theories, so can we "condense" all spacetime dimensions? The answer is that we can, until we reach $D = 2$. In two dimensions, the tachyon is stabilized due to its modified dispersion relation in the background of the linear dilaton [7]. For this value of $D$, $-V_\mu V^\mu = (D-26)/6\alpha') = -\frac{4}{\alpha'}$, so the dilaton gradient is equal and opposite to the mass of the tachyon. We may repeat the trick used for RR fields in section (2.1.6) to show that the field $e^\Phi \mathcal{T}$ is uncoupled to $\Phi$ and massless. The resulting theory is the well-known Liouville field theory, as mentioned earlier.

### 2.2.4 Tachyon condensation in $HO^{(+n)/}$ heterotic theory

Having studied tachyon condensation in the bosonic theory, we now turn to the more interesting case of the superstring. We also expect tachyon condensation to remove some spacetime dimensions and leave us with a lower dimensional heterotic theory, but there are several important differences from the start. The tachyon $\mathcal{T}^a$ is now charged under the gauge group, so tachyon condensation will break it. This makes sense as the $SO(32+n)$ gauge group of the heterotic theory is dimension-dependent. Another difference is worldsheet $(0,1)$ supersymmetry, which changes the way the tachyon background couples to the worldsheet. The only possibility which respects both supersymmetry and gauge invariance is

$$\Delta\mathcal{L} = -\frac{1}{2\pi}\int d\theta_+ \sum_a \lambda^a \mathcal{T}^a. \tag{2.2.12}$$

We use the superspace formalism [33] to simplify the task of obtaining supersymmetric actions. To compute the action, we expand the superfields as $\mathcal{T}^a = \mathcal{T}^a + \theta_+ \partial_\mu \mathcal{T} \lambda^a$, $\lambda^a = \lambda^a + \theta_+ F^a$, yielding

$$\Delta\mathcal{L} = -\frac{1}{2\pi}\sum_a \left(F^a \mathcal{T}^a(X) - i\sqrt{\alpha'}2 : \partial_\mu \mathcal{T}^a(X)\lambda^a \psi^\mu\right). \tag{2.2.13}$$

As usual in supersymmetric field theories, the auxiliary fields $F^a$ appear only algebraically in the action, without their derivatives. They can therefore be integrated out to obtain an effective potential

$$\Delta\mathcal{L} = -\frac{1}{2\pi}\sum_a \left(\mathcal{T}^a \mathcal{T}^a(X) - i\sqrt{\alpha'}2 : \partial_\mu \mathcal{T}^a(X)\lambda^a \psi^\mu\right). \tag{2.2.14}$$

The tachyon profile must obey the equations of motion for the tachyon,

$$\partial_\mu \partial^\mu \mathcal{T}^a - 2V^\mu \partial_\mu \mathcal{T}^a + \frac{2}{\alpha'}\mathcal{T}^a = 0. \tag{2.2.15}$$

By analogy with the bosonic case, we will be interested in a profile

$$\mathcal{T}^a(X) = \sum_\nu \mu^a_\nu \exp(\beta X^+)\,[X^\nu] + O(kX^2), \tag{2.2.16}$$

where as before we will work in the long wavelength $k \to 0$ limit and drop the quadratic terms in the tachyon potential. The corresponding bosonic lagrangian is

$$\mathcal{L}_{\text{int.}} = -\frac{1}{4\pi}\left(\sum_a \mu^a_\mu \mu^a_\nu\right)\exp(2\beta X^+)X^\mu X^\nu + \frac{i\mu^a_\mu}{2\pi}\exp(\beta X^+)\lambda^a(\psi^a + \beta X^+ \mu^a_\nu X^\nu \psi^+). \tag{2.2.17}$$

The first term is a mass term for the bosonic coordinates, with mass matrix $\mu^2$. Depending on the rank of this matrix, several bosonic coordinates will become massive and will be integrated out, as we did in the bosonic case. The second terms are Yukawa couplings pairing left and right-moving fermions, thus giving a mass to the supersymmetric partners of the $X^\mu$ and some of the $\lambda^a$. Thanks to worldsheet supersymmetry, the masses of bosons and fermions are the same. The last term is just a coupling between bosons and fermions.

As in the bosonic case above, the classical theory is exactly solvable [5], which means that one can integrate-out the massive fields from the quantum theory in an exact manner. We will end up with a theory with $10 + n - \text{rank}(\mu)$ dimensions, as $\text{rank}(\mu)$ bosons and right-moving fermions will become massive. The process will also break the gauge group to $SO(32+n-\text{rank}(\mu))$ for the same reasons. Integrating out the massive fields now produces a shift in metric and dilaton gradient, given by

$$\Delta V_+ = +\frac{\text{rank}(\mu)\beta}{4}, \qquad \Delta G_{++} = +\frac{\text{rank}(\mu)\beta^2\alpha'}{4}, \qquad \Delta G_{--} = -\frac{\text{rank}(\mu)\beta^2\alpha'}{4}. \tag{2.2.18}$$

An interesting thing is that this time we do not have to fine-tune several backgrounds, as we did in the bosonic string, in order to have a vanishing potential in the leftover theory. Supersymmetry has done this for us, since we are integrating out complete supermultiplets and their contributions to zero-point energies must vanish.

Finally, computing the new dilaton central charge, it turns out to be

$$c^{\text{dilaton}} = \frac{3}{2}[\text{rank}(\mu) - (D-10)], \tag{2.2.19}$$

while it originally was $\frac{3}{2}(D-10)$. Integrating out the massive fields has resulted in a transfer of 3/2 rank($\mu$) units of central charge from the dynamical fields to the dilaton.

As in the bosonic theory, we can choose rank($\mu$) to be any integer from 1 to $D-2$. We always end up at $X^+ = +\infty$ with a consistent string theory in $10+n-\text{rank}(\mu)$ dimensions with total central charge equal to $(26, 15)$. If $\text{rank}(\mu) < n$, the final theory is another supercritical theory. If $\text{rank}(\mu) > n$, we end up with a subcritical theory with a spacelike linear dilaton. In both cases there is a tachyon in the fundamental representation of the unbroken gauge group. If $\text{rank}(\mu) = n$, our final theory is a critical, unstable heterotic string theory with gauge group $SO(32)$ and a lightlike linear dilaton rolling to weak coupling in the future. In particular, the tachyons from the NS sector still survive.

This picture of tachyon condensation does not significantly change in the orbifolded $HO^{(+n)}/\mathbb{Z}_2$ theory: We still have the worldsheet superpotential (2.2.12) and the structure of fields becoming massive and central charge transfer is the same. There is however one important difference: As discussed in section 2.1.4, the orbifold forces the tachyon to vanish at the orbifold fixed locus. Before the orbifold, we were free to specify where the tachyon would condense -at any specific locus of our choosing, or nowhere at all-. The orbifold ensures that the 10d fixed locus where the fermions live is always preserved by tachyon condensation.

This has dramatic consequences for the particular case $\text{rank}(\mu) = n$. In this case, all the coordinates on which the orbifold acts geometrically are gone, its only effect in the worldsheet being swapping the sign of the $\lambda^a$. The worldsheet discrete gauge group is generated by the actions $(-1)^F$, global fermion number, and $(-1)^{F_L}$, left-moving fermion number. But this is precisely the spacetime-supersymmetric GSO projection, and thus our theory is none other than (a linear dilaton background of) the standard $SO(32)$ heterotic.

All this structure may be seen in spacetime terms, by using the rule of thumb that any field with an index along the supercritical directions (which are eaten up by the tachyon) will not be there after condensation. Physically, this means that any such field will have arbitrarily high energy levels due to its interactions with the tachyon condensate. Only fields which live entirely on the 10d locus, where the tachyon vanishes thanks to the orbifold projection, will be protected. The same holds in the gauge sector: Only fields lying in the $SO(32)$ left invariant by the tachyon vev will remain massless.

This means, at the massless level, that the tachyon and mixed components of the metric and $B$-field are gone, as are gauge field components along the supercritical directions and those which do not lie in $SO(32)$. The surviving fields comprise the bosonic sector of the $SO(32)$ heterotic. As for the fermions living in 10d, recall that they were actually bispinors of the tangent and normal bundle to the orbifold fixed locus: The tachyon condensate will couple to normal bundle spinors via couplings of the form $\bar{\chi}_n \Gamma^\mu \chi_n \partial_\mu \mathcal{T}$, so only fermions which are not charged under the normal bundle will remain massless. This selects the vector-spinor (which, as advertised in section 2.1.4, becomes the gravitino+dilatino) and the $SO(32)$ adjoint fermions (gauginos).

Remarkably, introducing the orbifold has the effect of removing the tachyons from the critical theory, and the tachyon condensation procedure reaches a stable endpoint.

### 2.2.5 Tachyon condensation in the $HO^{(+n)}$ theory

All of the above discussion carries over in a very similar fashion to the other supercritical version of heterotic string theory, the $HO^{(+n)}$ theory discussed in section 2.1.5. In this case, the tachyon couples only to the $\chi^b$ superfields,

$$\Delta \mathcal{L} = -\frac{1}{2\pi} \int d\theta_+ \sum_b \chi^b \mathcal{T}^b, \tag{2.2.20}$$

so the gauge group (be it $SO(32)$ or $E_8 \times E_8$) is manifestly unbroken. A tachyon profile with nontrivial dependence on some coordinates may be approximated, near a minimum, by a profile of the form $\mu^b_\rho X^\rho e^{\beta X^+}$. Upon computing the potential exactly as we did after (2.2.12), we see that $k = \text{rank}(\mu^b_\rho)$ coordinates get a mass, as do $k$ $\chi$ fermions which get Yukawa couplings with the supertparners of the massive coordinates. At large times, we may integrate out all these fields, going from the $HO^{(+n)}$ theory to the $HO^{(+(n-k))}$ theory with the appropriate central charge shift. Unlike in the previous examples, this process can only take us as far as the critical dimension $n = 0$, since after that we run out of $\chi$ fermions.

### 2.2.6 Tachyon condensation in type 0 theory

Tachyon condensation can also take place in type 0 theories. The main idea is identical to the bosonic and heterotic examples, although there are significant technical differences. The tachyon couples to the worldsheet as a $(1,1)$ superpotential

$$\Delta \mathcal{L} = \frac{i}{2\pi} \int d\theta^+ d\theta^- \mathcal{T}(X) \tag{2.2.21}$$

where $\mathcal{T}(X)$ depends on the $(1,1)$ superfields $X^\mu + i\theta^- \psi^\mu + i\theta^+ \bar{\psi}^\mu + i\theta^- \theta^+ F^\mu$ (see e.g. [34]). Upon integrating out the auxiliary fields $F^\mu$, the terms in components describe a worldsheet potential and Yukawa couplings

$$\mathcal{L}_{\text{pot}} \sim \partial^\mu \mathcal{T}(X) \partial_\mu \mathcal{T}(X) \quad , \quad \mathcal{L}_{\text{Yuk}} \sim \partial_\mu \partial_\nu \mathcal{T}(X)\, \bar{\psi}^\mu \psi^\nu \tag{2.2.22}$$

As in the heterotic example above, condensation of this tachyon can produce dimension quenching in type 0 theories, but cannot connect down to 10d type II theories. A perhaps surprising result is [35] that a tachyon profile of the form $\exp(\beta X^+)$ does not change spacetime dimension, but instead changes the type of string theory, connecting type 0 superstrings to (supercritical) bosonic strings.

The equation of motion for the type 0 tachyon is the same as for the type II tachyon, so again we will focus on solutions of the form (2.2.16). From (2.2.22), we see that a linear tachyon profile would only contribute as a worldsheet vacuum energy. Interesting physics starts to show up at quadratic order, and hence we will take a tachyon profile of the form $\mathcal{T} \sim \mu_{\mu\nu} X^\mu X^\nu$. This will yield a mass matrix $\mu_{\mu\nu}$ both for bosons and fermions, which can be integrated out in a by now familiar fashion. The resulting endpoint of the tachyon condensation procedure is again type 0 theory with a lower number of dimensions.

However, as in the heterotic case, things change drastically if we consider the orbifolded cousin of type 0 theory presented in section 2.1.6. Recall that in this case the orbifold fixed locus has dimension $10 + n$. The tachyon profile has to be an odd function of the spacetime coordinates normal to this fixed locus, which means that a tachyon profile quadratic in these coordinates is forbidden. Thus, one cannot annihilate the normal coordinates leaving the $10 + n$-dimensional theory as we did in the heterotic case[4]. However, we may pair a supercritical coordinate with coordinates of the fixed locus, via a term $\mathcal{T} \sim XYe^{\beta X^+}$ (here, $Y$ is a supercritical coordinate and $X$ lives along the critical slice). This generates a mass term of the form $(X^2+Y^2)e^{\beta X^+}$ for both bosons. The Yukawa couplings in (2.2.22) ensure that the superpartners get the right masses.

At late times, all these fields become very massive and should be integrated out. The dilaton and metric shifts per supermultiplet are the same as in the heterotic theory. If we integrate out all the supercritical coordinates, we are left with a 10d theory without tachyons, since the $10 + 2n$ tachyon vanishes at the fixed locus as in the heterotic case. The surviving orbifold action is just $(-1)^{F_L}$, overall left-moving fermion number. Together with the type 0 $(-1)^F$ orbifold, these constitute the worldsheet GSO projection of type II strings. Thus, tachyon condensation in orbifolded type 0A (B) theory leads to (a lightlike dilaton background of) type IIA (B) theory.

### 2.2.7 Final comments

We conclude the discussion with some generic remarks: Dimension quenching is drastically different from Kaluza-Klein dimensional reduction. From the spacetime viewpoint, dimension quenching causes the extra dimensions to completely disappear from the theory. In particular, there remain no towers of extra-dimensional momentum modes. Even at the level of massless modes, its effect on spacetime

[4] There is a good reason why this should not be possible: The fermion spectrum in orbifolded 0B is chiral, meaning that the resulting theory suffers from gravitational anomalies [36] and is therefore inconsistent .

bosons differs from a truncation to the zero mode sector; for instance, mixed components $G_{\mu m}$ disappear completely (whereas they can survive in KK dimensional reduction).

The natural order parameter measuring the tachyon condensation process is whatever couples to the spacetime fermions, namely $\partial_m \mathcal{T}^a$ in heterotic, and $\partial_m \partial_n \mathcal{T}$ in type 0/II. The symmetry breaking pattern is mostly encoded in the quantum numbers of this quantity. For instance, in section 2.2.4 the background $\partial_m \mathcal{T}^a$ breaks the $SO(32+n) \times SO(n)$ gauge and rotational symmetries down to $SO(32)$ (times a diagonal $SO(n)_{\text{diag}}$). Similarly, in section 2.2.6 the background $\partial_m \partial_n \mathcal{T}$ breaks the $SO(n) \times SO(n)$ rotational symmetries (seemingly with a left-over diagonal $SO(n)_{diag}$). Although the diagonal symmetries $SO(n)_{diag}$ would seem unbroken from a Higgsing perspective in spacetime, the microscopic worldsheet computation shows that they actually disappear. This is reminiscent of the similar situation in open string tachyon condensation described in section 2.2.1. If we take the S-duality described in section 2.1.4 seriously, as well as the proposal of [22] for the type I phenomenon, we would conclude that $SO(n)_{diag}$ is removed from the spectrum thanks to the condensation of unstable heterotic 7-branes described in [37] and which will be studied in detail in Chapter 4.

Finally, the interested reader can find the partition functions of the supercritical theories discussed in the text in Appendix A.

# 3

# Discrete gauge symmetries from closed-string tachyon condensation

In the context of string theory model building, one is often faced with seemingly global discrete symmetries of the effective field theory. On general grounds, one would expect these symmetries to be gauged, but it is difficult to argue that this is the case in general. In this Chapter we will employ supercritical string theories to embed any discrete symmetry we may be interested in into a continuous one in the supercritical theory. We will first discuss the rationale behind the dislike for global symmetries in consistent quantum theories of gravity. After that, we will describe the main strategy to embed discrete symmetries into continuous ones while discussing a variety of examples.

Most of the results of this Chapter are published in [13].

## 3.1 Global symmetries in quantum gravity

There is a standard argument [38, 39, 40, 41, 42, 43] against the existence of continuous symmetries in any quantum theory of gravity. This is implemented by a group action on the theory which leaves the Hamiltonian and every transition amplitude invariant.

Suppose there is one such symmetry and take one object charged under the nontrivial representation $\mathbf{r}$. A collection of $n$ such objects will live in an irreducible representation $\mathbf{R}_n \subset \mathbf{r} \otimes \ldots \otimes \mathbf{r}$. Since the group is continuous, for each $n$ we get at least one new irreducible representation. In other words, we have infinitely many inequivalent irreducible representations as well as the means to build objects charged under them.

By collapsing these objects into black holes, we can achieve black holes charged in the irreducible representation $\mathbf{R}_n$ and of arbitrary mass up to $M_p$. If we want to increase or decrease the mass we may simply add neutral matter or wait until the BH evaporates enough mass[1].

[1]This argument relies on Hawking's computation [44] to allow the black hole to evaporate. This

Now, take any object, living in a rep. $\mathbf{R}_o$ of the symmetry group. Generically, we can expect a nonzero amplitude for the object to tunnel into a pair of BH's into reps. $\mathbf{R}_n$, $\mathbf{R}_{n'}$ such that $\mathbf{R}_n \otimes \mathbf{R}_{n'} = \mathbf{R}_o$, since the process is classically allowed. Since the two final state BH's are far apart and by hypothesis the symmetry is global, i.e. does not affect the dynamics, the physics must be invariant under the action of the symmetry group on each of the black holes separately. In other words, the amplitude is the same no matter the values $n$, $n'$, as long as $\mathbf{R}_n \otimes \mathbf{R}_{n'} = \mathbf{R}_o$.

To compute the total decay width to a pair of black holes, we should sum over all possible $n$, $n'$. It can be shown that a continous group has infinitely many irreducible representations[2], and so the number of possible final states for this process is unbounded, which means that the decay width to a pair of black holes is not bounded. In other words, if one believes this argument, an exact continuous symmetry means that everything should have decayed to black holes long ago. The requirement for the symmetry to be continuous is crucial in this argument: Discrete symmetries only have a finite number of inequivalent irreducible representations, and so there is only a finite number of BH final states.

There is a refinement of the argument, proposed by Susskind [43], which does not rely on arguments about black hole decay rates we cannot compute anyway. As discussed above, in a theory with an exact global symmetry we can build an arbitrarily number of black hole states, charged under arbitrarily high representations of the symmetry. If all of these large black holes evaporate following Hawking's calculation all the way to the Planck scale, we will have a large (formally infinite) number of stable remnants of Planckian mass. On the other hand, Unruh [48] showed that a uniformly accelerated observer in the vacuum state sees a thermal bath at the Unruh temperature $T = \frac{a}{2\pi}$. If there are infinitely many remnants almost degenerate in mass their contribution will dominate the partition function of the Unruh thermal bath, no matter how small the temperature. This is clearly a theory out of theoretical control.

Admittedly, the above arguments are not watertight. In particular they rely on the validity Hawking's calculation all the way to the Planck scale, or at least to some scale independent of the black hole mass. If black holes with arbitrarily high charges under the global symmetry leave the semiclassical approximation at arbitrarily high masses, then the above conclusions can be avoided. But then one has to explain the existence of very large black holes which are nevertheless outside the semiclassical approximation.

calculation was done in the context of quantum field theory in curved classical backgrounds. The argument will not hold in any model in which it is invalid. Most recently, the proposals in [45, 46] allow for global symmetries in quantum gravity by having even macroscopic black holes behave as quantum objects.

[2]The order of a group is the sum of the squares of the characters of its irreducible representations. If this order is infinite, as is the case for a continuous group, there must be infinitely many inequivalent irreducible representations. For a nice introduction to index theory, see [47].

### 3.1.1 No global symmetries in the worldsheet

String theory is a quantum theory of gravity, in which Hawking's calculation seems to hold for large black holes. Hence, it should be devoid of continuous global symmetries. Indeed, this can be proven explicitly, at least in perturbative setups which admit a worldsheet description [12]. In string perturbation theory, amplitudes are computed as correlators of the 2-dimensional CFT describing the worldsheet of the particular string theory we are concerned with. Any continuous symmetry of the theory therefore gives rise to a continuous symmetry of the worldsheet effective action and, by Noether's theorem, to a conserved charge, which may be written as

$$Q = \frac{1}{2\pi i} \int (j_z dz - j_{\bar{z}} d\bar{z}) \tag{3.1.1}$$

in the bosonic string (an analogous expression holds for the superstring). $Q$ must be conformally invariant, since it takes physical states (represented by conformally invariant vertex operators) to physical states. This implies that $j_z$ is a $(1,0)$ tensor, and we can construct the vertex operators

$$j_z \bar{\partial} X^\mu e^{ikX}, \quad \partial X^\mu j_{\bar{z}} e^{ikX} \tag{3.1.2}$$

which correspond to spin 1 particles charged under the continuous symmetry under consideration, the gauge bosons. These considerations extend to open strings [49], in which gauge bosons may be constructed via vertex operators

$$\lambda_{ab} \partial X^\mu e^{ikX}, \tag{3.1.3}$$

where $\lambda_{ab}$ is a Chan-Paton label (telling us on which D-branes the open string starts and ends).

As mentioned above, this argument has a loophole in that the charge $Q$ must be nonvanishing, i.e. strings must be charged under the symmetry. There are several examples of symmetries in string theory which do not satisfy this. For instance, type IIB theory contains an axion $C_0$ in its Ramond-Ramond sector. There is a perturbative $U(1)$ Peccei-Quinn symmetry which acts by shifting the axion field. This does not contradict the above argument since strings do not carry Ramond-Ramond charges and therefore do not see the above symmetry. Nonetheless, the symmetry is broken at the nonperturbative level via $D$-instanton effects.

### 3.1.2 Discrete symmetries versus discrete gauge symmetries

Discrete symmetries, unlike continuous ones, are not forbidden by principle in a quantum theory of gravity. In fact, there are several examples of these in string theory: The nonperturbative $SL(2,\mathbb{Z})$ S-duality of type IIB strings, ordinary $T$-duality for strings, the $\mathbb{Z}_2$ symmetry in heterotic string theory which acts by $-1$ on the massive spinor states of the theory. Consider also diffeomorphisms and gauge transformations of the $B$ or any Ramond-Ramond field which do not vanish at infinity.

For some of these symmetries, especially the last ones, it is apparent that they are elements of the global part of a gauge symmetry. Thus, configurations related by a gauge transformation are identified and the physical objects of the theory are the resulting equivalence classes. To be specific, a gauge transformation is a map $\mathcal{M} \to G$ from spacetime to the gauge group which vanishes at infinity. Transformations which do not vanish at infinity constitute the global part of the gauge group and act as a symmetry relating physically inequivalent configurations. These configurations may be distinguished from each other by the gauge flux they create at infinity [50]. For instance, for a $U(1)$ gauge theory, states fall into representations of the global part of the gauge group. This global part is precisely $U(1)$, and the integer indexing the representation is electric charge.

The global part of the gauge group may be broken by our choice of background. In general, the global gauge group will be broken to a discrete subgroup. We will call such a group a discrete gauge symmetry, to remember its origin as the global part of a gauge group. Continuing with the $U(1)$ example discussed above, now couple the gauge field to an axion $\phi \sim \phi + 2\pi$ with the usual Stuckelberg coupling:

$$\mathcal{L} = -\frac{1}{2g^2} F \wedge *F + v^2 (d\phi - pA) \wedge *(d\phi - pA). \tag{3.1.4}$$

The gauge symmetry in this theory acts by $A \to A + d\lambda$, $\phi \to \phi + p\lambda$. It has a set of degenerate vacua with $A = 0$, $\phi =$ constant, which we may parametrize by the value of $\phi$. Different vacua of this sort are related by global gauge transformations $\lambda = \lambda_0$ which are constant throughout spacetime, since they take $\phi_0$ to $\phi_0 + p\lambda_0$. The vacum will be preserved whenever $p\lambda_0$ is a multiple of $2\pi$, by virtue of the periodicity of $\phi$. This will happen for $\lambda_0 = 2\pi \frac{n}{p}$ for some integer $n$. Now, $\lambda_0$ is a map from spacetime to $U(1)$, which means it is also an angle and defined only modulo $2\pi$. If $p = 1$, no global gauge transformations other than the trivial one preserve the vacuum. But for $p > 1$, there are $p$ global gauge transformations which do. These constitute the discrete gauge symmetry of our configuration.

Given a discrete symmetry, it is not easy to tell if it is gauged or not. It will be if we can find a continuous gauge symmetry whose global part is partially broken by the choice of background to our discrete symmetry. For instance, $T$-duality at the self-dual radius $R = \sqrt{\alpha'}$ is enhanced to an $SU(2)$ gauge symmetry. We may then say that $T$-duality is a $\mathbb{Z}_2$ remnant of the self-dual radius $SU(2)$.

This strategy however fails for the IIB $SL(2,\mathbb{Z})$ symmetry. There is no simple way to embed this in some continuous gauge symmetry, as can be seen from a more geometric perspective: The $SL(2,\mathbb{Z})$ of type IIB acts on the axio-dilation $\tau \equiv C_0 + \frac{i}{g_{IIB}}$ as a Möbius transformation, exactly as large diffeomorphisms would on the complex structure of a torus. By compactifying IIB on a circle and $T$-dualizing to the IIA theory, after which we take the M theory limit[51], we arrive at M theory on $T^2$, with $\tau$ indeed being its complex structure parameter. The $SL(2,\mathbb{Z})$ of IIB is in this way geometrized to the group of large diffeomorphisms of the M-theory torus. As diffeomorphisms are gauged in a quantum theory of gravity, this shows that the the IIB $SL(2,\mathbb{Z})$ is indeed a discrete gauge symmetry.

For the heterotic $\mathbb{Z}_2$ swapping the sign of massive spinors, there is no (known) chain of dualities that would allow us to embed the symmetry in a continuous group, or at least in a group which is manifestly gauged. We will use supercritical theories to show that this is indeed the case for this and other symmetries which often arise in string model building.

## 3.2 Discrete gauge symmetries as quenched rotations

In general, discrete symmetries cannot be regarded in an evident way as a discrete subgroup of a continuous group action on the theory. This happens for instance for discrete large isometries in compactifications, namely discrete isometries associated to large diffeomorphisms of the compactification manifold (diffeomorphisms which cannot be continuously deformed to the identity). In compactifications, discrete isometries of the internal space become discrete gauge symmetries of the lower-dimensional theory, much like continuous isometries become gauge symmetries à la Kaluza-Klein [52]. Hence, discrete gauge symmetries from large isometries cannot in principle be regarded as unbroken remnants of some continuous gauge symmetry of the effective field theory.

We will now show that such discrete symmetries can be embedded into continuous groups, which however act on extra dimensions in a supercritical extension of the theory. The continuous orbits come out of the critical spacetime slice, and are quenched by a closed string tachyon condensation process, which thus breaks the continuous symmetry, yet preserves the discrete one (see Figure 3.1). This process is similar to a Higgs mechanism, but with some important differences. The appearance of extra dimensions is very intuitive in extending discrete isometries to a continuous action, since these symmetries are related to properties of the metric; still our constructions apply to fairly general discrete symmetries.

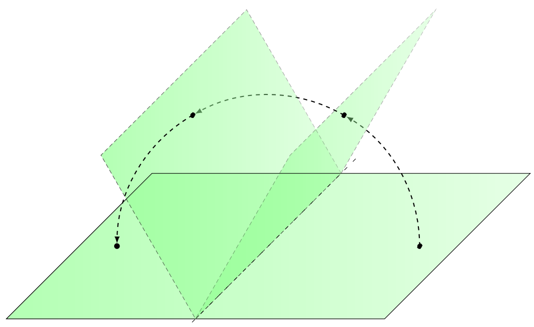

**Figure 3.1:** Schematic description of the quenched rotation construction. The original theory lives in the horizontal plane, and has a discrete reflection symmetry which cannot be connected to the identity. By adding an extra, supercritical dimension, we are able to embed this discrete action into the enlarged rotation group.

### 3.2.1 Spacetime parity

A prototypical example of large diffeomorphisms it that of orientation reversing actions, for instance, and most prominently, spacetime parity. Consider the action

$$(X^0, X^1, \ldots, X^{2n-1}) \to (X^0, -X^1, \ldots, -X^{2n-1}) \tag{3.2.1}$$

in $2n$-dimensional Minkowski spacetime. It has determinant $-1$, so it lies in a disconnected component of the Lorentz group. For convenience, we equivalently focus on the action $X^{2n-1} \to -X^{2n-1}$ (with other coordinates invariant), which lies in the same component disconnected from the identity.

It is easy to embed this symmetry into a continuous one, by adding one extra dimension $X^{2n}$, and considering the $SO(2)$ rotation in the 2-plane $(X^{2n-1}, X^{2n})$. Since this changes the number of dimensions, any string theory implementation will require us to work in a supercritical theory. We now show this for the supercritical bosonic string (since superstrings are actually not parity-invariant, due to chiral fermions in type IIB or Chern-Simons couplings in the quantum effective action for type IIA; still, parity can be combined with other actions to yield symmetries of such theories, see the CP example in Section 3.3.3).

Consider the supercritical bosonic string theory with one extra dimension, denoted by $X^{26}$, with the appropriate timelike linear dilaton. The theory is invariant under continuous $SO(2)$ rotations in the 2-plane $(X^{25}, X^{26})$. As in Section 2.2.3 we can connect it with the critical 26d bosonic theory by a closed string tachyon profile

$$\mathcal{T}(X^+, X^{26}) \sim \frac{\mu^2}{2\alpha'} \exp(\beta X^+)\,(X^{26})^2. \tag{3.2.2}$$

At $X^+ \to -\infty$ the tachyon vanishes and we have a 27d theory with a continuous $SO(2)$ rotational invariance in the 2-plane $(X^{25}, X^{26})$. At $X^+ \to \infty$, the onset of the tachyon truncates the dynamics to the slice $X^{26} = 0$, breaking the $SO(2)$ symmetry to the $\mathbb{Z}_2$ subgroup $X^{25} \to -X^{25}$. Hence the $\mathbb{Z}_2$ parity symmetry can be regarded as a discrete subgroup of a continuous higher dimensional rotation group, broken by the tachyon condensation removing the extra dimension.

The analogy of this breaking with a Higgs mechanism can be emphasized by using polar coordinates, $W = X^{25} + iX^{26} = |W|e^{i\theta}$. Then $X^{26} \sim W - \bar{W}$ and

$$\mathcal{T} \sim (X^{26})^2 \sim W^2 - 2W\bar{W} + \bar{W}^2 \sim e^{2i\theta} - 2 + e^{-2i\theta}. \tag{3.2.3}$$

In these coordinates, the $\theta$ angle behaves as an axion coupled to the rotational $U(1)$ symmetry. The tachyon sets a potential for this axion which is invariant under $\theta \to \theta + \pi$, hence a $\mathbb{Z}_2$ symmetry remains. This description will be useful for the construction of $\mathbb{Z}_2$ charged defects in Section 3.2.3.

Finally, note that although this construction embeds the discrete group into a continuous one, there is no actual 26d $SO(2)$ gauge boson, since space is noncompact. This will be achieved in a slightly different construction in Section 3.3.

### 3.2.2 A heterotic $\mathbb{Z}_2$ from closed tachyon condensation

In the 10d $SO(32)$ heterotic string, the gauge group is actually $Spin(32)/\mathbb{Z}_2$, and there is a $\mathbb{Z}_2$ symmetry under which the (massive) spinor states are odd, while fields in the adjoint are even[3][10]. Namely, in the fermionic formulation, this symmetry acts by swapping the sign of all the fields coming from a Ramond left-moving worldsheet sector. We now propose a realization of this $\mathbb{Z}_2$ symmetry as a discrete remnant of a continuous $U(1)$ symmetry, exploiting the supercritical heterotic strings introduced in Section 2.1.4.

For that purpose, it suffices to focus on the case of $D = 12$, i.e. two extra dimensions, denoted $X^{10}$, $X^{11}$. We complexify the extra worldsheet fields into a complex scalar $Z \equiv X^{10} + iX^{11}$, and a complex fermion $\Lambda = \lambda^{33} + i\lambda^{34}$. We consider the $U(1)$ action

$$Z \to e^{i\alpha} Z \quad , \quad \Lambda \to e^{-i\alpha}\Lambda \tag{3.2.4}$$

namely, the anti-diagonal $SO(2)_{\text{anti}} \subset SO(2)_{\text{rot}} \times SO(2)_{\text{gauge}} \subset SO(2)_{\text{rot}} \times SO(2{+}32)$. The idea here is that as we rotate along the extra dimensions the fermions pick the right phase so as to recover the original $\mathbb{Z}_2$ symmetry when a full turn is completed. Consider now the tachyon background which condenses to the $SO(32)$ heterotic, namely

$$\mathcal{T}^{33}(X) \sim e^{\beta X^+} X^{10} \quad , \quad \mathcal{T}^{34}(X) \sim e^{\beta X^+} X^{11} \quad \to \quad \mathcal{T}(X) \sim e^{\beta X^+} Z, \tag{3.2.5}$$

which we have recast in terms of a holomorphic complex tachyon $\mathcal{T} \equiv \mathcal{T}^{33} + i\mathcal{T}^{34}$. The order parameter $\partial_Z \mathcal{T}$ transforms in the bifundamental of $SO(2)_{\text{rot}} \times SO(2)_{\text{gauge}}$, and breaks the $SO(2)_{\text{rot}} \times SO(34)$ symmetry down to $SO(32)$ (times a diagonal factor which 'disappears'). The anti-diagonal $SO(2)_{\text{anti}}$, generated by $Q_{\text{anti}} = Q_{SO(2)_{\text{gauge}}} - Q_{SO(2)_{\text{rot}}}$, is broken by a charge $+2$ the tachyon background. To show that there is an unbroken $\mathbb{Z}_2$ acting as $-1$ on the $SO(32)$ spinors, it suffices to show that they descend from states with $SO(2)_{\text{anti}}$ charge $\pm 1$. Indeed, they descend from the massive groundstates in the $g_2$ twisted Ramond sector (see Section 2.1.4), which has $\lambda^a$, $\psi^m$ fermion zero modes paired together; the states transform as a $SO(34) \times SO(2)_{\text{rot}}$ bi-spinor, hence have $SO(2)_{\text{anti}}$ charge $\pm 1$, and descend to $\mathbb{Z}_2$ odd $SO(32)$ spinors.

A slightly unsatisfactory aspect of the construction is that there are actually no gauge bosons associated to the $SO(n)_{\text{rot}}$ group because the extra dimensions are noncompact, as in the bosonic case. We could achieve this by curving the geometry of the extra dimensions in the radial direction, with the angular coordinates asymptoting to a finite size $S^{n-1}$. Its $SO(n)$ isometry group would then produce an $SO(n)$ gauge symmetry (in 11d). Although the geometric curvature will render the worldsheet theory non-solvable, we expect basic intuitions of the flat space case to extend to the curved situation, in what concerns the relevant topology of symmetry breaking. In any event, we will eventually turn to a more general construction, with physical gauge bosons, in Section 3.3.

[3] Recall that the $\mathbb{Z}_2$ by which $Spin(32)$ is quotiented prevents the presence of states in the vector representation; also notice that this is not the $\mathbb{Z}_2$ discrete gauge symmetry we are interested in.

### 3.2.3 Topological defects from closed tachyon condensation

In the embedding of a discrete symmetry into a continuous one acting on extra dimensions, all relevant degrees of freedom are eventually removed by the closed string tachyon condensation. We may therefore ask what we gain by such construction. The answer is that, in analogy with open string tachyon condensation, certain branes of the final theory can be constructed as solitons of the tachyon field. Indeed, discrete gauge symmetries are often identified via the procedure of finding objects charged under them. Since our focus is on aspects related to $\mathbb{Z}_n$ discrete gauge symmetries, in this Section we describe the tachyon profiles corresponding to the codimension-2 $\mathbb{Z}_n$ charged defects (real codimension-2 objects around which the theory is transformed by a discrete $\mathbb{Z}_n$ holonomy, e.g. in 4d we would have $\mathbb{Z}_n$ strings). The interesting question of constructing other possible defects, and the possible mathematical structures underlying their classification is the focus of Chapters 4 and 5.

Focusing back on the construction of $\mathbb{Z}_n$ defects, let us consider again the analysis of parity in the bosonic theory above. This $\mathbb{Z}_2$ symmetry is embedded into a $U(1)$ rotating the 2-plane $(X^{25}, X^{26})$. The critical vacuum is recovered by a tachyon background $\mathcal{T} \sim (\text{Im}\, W)^2$ c.f. (3.2.3), with $W \equiv X^{25} + iX^{26} = |W|e^{i\theta}$, whose zero cuts out the slice $\sin\theta = 0$. In order to describe a $\mathbb{Z}_2$ defect transverse to another 2-plane, e.g. $(X^{23}, X^{24})$, we write $Z \equiv X^{23} + iX^{24} = |Z|e^{i\varphi}$ and consider a closed string tachyon background vanishing at the locus $\sin\theta' = 0$, with $\theta' = \theta - \frac{1}{2}\varphi$. To put it another way, the condition $\sin\theta' = 0$ is equivalent to

$$\sqrt{Z}\, W = \sqrt{\overline{Z}}\, \overline{W} \tag{3.2.6}$$

so we are just choosing a tachyon profile with nontrivial dependence along all the coordinates. The critical slice defined by (3.2.6) is non-flat, and if we turn by an angle $\delta\varphi$ in the $(X^{23}, X^{24})$-plane, to remain at the tachyon minimum locus we must also turn on $\theta$. As we make a full turn on $\varphi$, we only make half a turn in $\theta$, so in terms of these coordinates a full turn on the $(X^{23}, X^{24})$-plane results in a parity action. In other words, any object sensitve to parity transformations (such as e.g. the volume form of spacetime) will pick up a phase upon following a closed path encircling the $Z = 0$ locus, where (3.2.6) degenerates. This is one of the characterizations of a charged object (namely, that the dual object sees the relevant holonomy).

Other similar examples will be discussed in Section 3.3.2.

## 3.3 Discrete gauge symmetries as quenched translations

The set of discrete symmetries amenable to the quenched rotation construction in the previous Section is limited, since we need actions which are geometric in nature. Additionally, the gauge nature of these symetries was unclear as there was no gauge boson associated to the continuous symmetry. We now present a far more universal embedding of discrete symmetries into continuous ones, which are realized as continuous translations in an extra (supercritical) $S^1$ dimension. The extra $S^1$ is subsequently eaten up by closed string tachyon condensation, which breaks the KK $U(1)$ symmetry down to the discrete subgroup. The basic strategy is to use periodic tachyon profiles to yield condensation processes which are well-defined on $S^1$. In doing this it will be impossible to write down exactly marginal tachyon profiles as we did in Section 2.2, since the spacetime equations of motion will become convoluted. We do not mind giving up exact solvability of the worldsheet CFT, and rely on the main lesson that condensation truncates dynamics to the vanishing locus of the 2d potential energy[4]. This can be justified in a nonrigorous manner by the observation that, locally, the tachyon condensation process happens in a flat target spacetime, and therefore our previous analysis applies.

### 3.3.1 The mapping torus

The basic ingredient in the construction is the mapping torus, whose construction we illustrate in fairly general terms. Although we apply it in the string theory setup, most of the construction can be carried out in the quantum field theory framework[5]; the ultimate removal of the extra dimensions by tachyon condensation is however more genuinely stringy.

The mapping torus may be defined as follows: Consider an $N$-dimensional theory on a spacetime $X_N$ (which includes gravity), and let $\Theta$ be the generator of an abelian discrete gauge symmetry $\mathbb{Z}_n$. We consider extending the theory to $X_N \times \mathbf{I}$, where $\mathbf{I}$ is a one-dimensional interval[6] parametrized by a coordinate $0 \leq y \leq 2\pi R$. We subsequently glue the theories at $y = 0$ and $y = 2\pi R$, but up to the action of $\Theta$, as illustrated in Figure 3.2. The final configuration is the theory on $X_N$ non-trivially fibered over $S^1$. The fibration is locally $X_N \times \mathbb{R}$, but there is a non-trivial discrete holonomy implementing the action of $\Theta$. For example, if $\Theta$ is a discrete large isometry, the (purely geometric) glueing is

$$(x, y = 0) \sim (\Theta(x), y = 2\pi R). \tag{3.3.1}$$

Returning to the general situation, the compact extra dimension produces a KK $U(1)$ gauge boson from the $N$-dimensional viewpoint. The orbit of the associated translational vector field $\partial_y$ clearly contains the discrete $\mathbb{Z}_n$ transformations, which are thus embedded as a discrete subgroup $\mathbb{Z}_N \subset U(1)$. Since our discrete symmetry has been embedded into a KK $U(1)$, states of the theory $X_N$ transforming nontrivially under the discrete symmetry generator $\Theta$ must carry a nontrivial

[4]This is analogous to the applications of open string tachyon condensation in annihilation processes, even if they are not exactly solvable BCFTs.

[5]Incidentally, the mapping torus is widely used in the study of anomalies [53, 54, 55].

[6]This is easily implemented in the supercritical bosonic and heterotic $HO^{+(n)}$ theories; for the supercritical type 0 orbifolds decaying to type II we must add another extra dimension, which can be kept non-compact; for the $HO^{+(n)}/$ theory, the orbifolding breaks the translational invariance and there is no actual continuous KK gauge symmetry, so we do not consider it here.

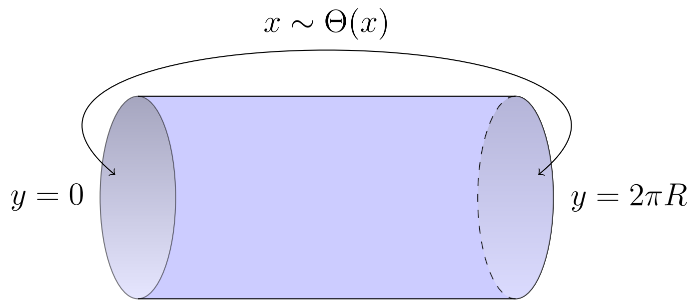


**Figure 3.2:** Mapping torus construction. The theory at $y = 0$ is glued to the theory at $y = 2\pi R$ modulo the action of the discrete symmetry $\Theta(x)$.

KK charge, i.e. momentum. In fact, states transforming with phase $e^{2\pi i\, p/n}$ under the $\mathbb{Z}_n$ symmetry extend as states with fractional momentum $p/nR$ (mod $\mathbb{Z}$) along $S^1$, since a state with momentum $P$ picks up a phase $e^{2\pi i PR}$ upon a $2\pi R$ translation. We note that the minimal $U(1)$ charge unit is $1/nR$, since upon transversing the circle $n$ times all states must be periodic; hence $nRP$ must be an integer.

In the supercritical string theory construction, the extra dimension is removed by a tachyon profile with periodicity $2\pi R$, which truncates the theory to the slice $y = 0$ (mod $2\pi R$). For instance, in the bosonic string theory, we sketchily write[7]

$$\mathcal{T} \sim \mu^2 \left[ 1 - \cos\left(\tfrac{y}{R}\right) \right] \; = \; 2\mu^2 \sin^2\left(\tfrac{y}{2R}\right) \tag{3.3.2}$$

and similarly in other supercritical string theories. Concretely, we use the heterotic $HO^{+(n)}$ theory (since the $HO^{+(n)/}$ breaks translational invariance in the extra dimensions) and take $\mathcal{T} \sim \sin(\frac{y}{2R})$; for type II extended as supercritical type 0 orbifold, we take $\mathcal{T} \sim \sin(\frac{y}{2R})X'$. Here we assume that $R \gg \alpha'$, namely that the variation of the tachyon profile is slow enough so that the low-energy equations of motion involved in the construction of these profiles can be trusted. Of course, all of them are dressed by a $e^{\beta X^+}$ factor ensuring the onset of the dynamical condensation process.

The Higgsing picture can also be recovered in this construction. Since the tachyon profile has a periodicity of $2\pi R$, it only excites components of integer KK momentum. We showed above that the theory had states with fractional momentum; these will survive the condensation process, in which the $U(1)$ symmetry is broken only by fields of integer charge. Normalizing the minimal charge to $+1$, the breaking is implemented by fields of charge $n$. Hence, the continuous symmetry is broken to a discrete $\mathbb{Z}_n$ symmetry in the slice $y = 0$, i.e. to the $\mathbb{Z}_n$ of the original theory at $X_N$.

### 3.3.2 Topological $\mathbb{Z}_n$ defects and quenched fluxbranes

A basic property of theories with discrete $\mathbb{Z}_n$ gauge symmetries with particles charged under them is the existence of $\mathbb{Z}_n$ charged defects, real codimension-2 objects

[7]Note that the radius $R$ can be kept arbitrary; for instance, the operator can be made marginal by turning on a lightlike dependence and adjusting the $\beta$ coefficient appropriately.

around which the theory is transformed by a discrete $\mathbb{Z}_n$ holonomy. This non-trivial behaviour of the theory around the $S^1$ surrounding the $\mathbb{Z}_n$ defect, is identical to the fibration over $S^1$ in the mapping torus in the previous sections; both of them pick an action of the discrete symmetry generator when going full circle. This may be regarded as an underlying reason for the universality of the mapping torus construction, which applies to fairly general discrete symmetries (as opposed to those in Section 3.2). In this Section we use this relation to construct tachyon condensation profiles which produce the $\mathbb{Z}_n$ charged defects of the theory, generalizing Section 3.2.3 to the more universal mapping torus setup.

The construction is very reminiscent of the fluxbranes[8] in [57, 58, 59, 60], as we now explain. In compactifications on an $S^1$, parametrized by a coordinate $y$ with period $2\pi R$, a fluxbrane is a brane-like configuration with two transverse real coordinates constructed as follows. Pick a 2-plane, denoted by $(X^8, X^9)$ for concreteness, and use polar coordinates $Z \equiv X^8 + iX^9 \equiv |Z|e^{i\varphi}$. The fluxbrane configuration is obtained by performing dimensional reduction not along the original Killing vector $\partial_y$, but rather along $\partial_{y'}$, where $y' = y - \varphi R$ parametrizes a combined $S^1$. In the resulting solution, the orthogonal combination $\varphi' = \varphi + y/R$ plays the role of angular coordinates in a 2-plane transverse to the fluxbrane, and there is a non-trivial magnetic flux for the KK gauge boson $A_\mu \sim G_{\mu y'}$ (hence the name). In other words, circumventing the fluxbrane by rotating $\varphi'$ secretly rotates in $\varphi$ and translates along $y$, such that the total holonomy is a whole shift in the original $S^1$ parametrized by $y$.

The $\mathbb{Z}_n$ defects can be constructed by using the same strategy in dimensional quenching, rather than in dimensional reduction, as described in Figure 3.3. We start with the mapping torus of $X_N$ fibered over a $S^1$ parametrized by $y$, as described in Section 3.3.1. We choose a 2-plane $(X^8, X^9)$, or $Z \equiv X^8 + iX^9 \equiv re^{i\varphi}$. Finally, we turn on a closed string tachyon profile (3.3.2), but now depending on $y' = y - \varphi R$, to remove one extra $S^1$ dimension. The resulting configuration contains a $\mathbb{Z}_n$ defect at the origin of the 2-plane, since a rotation in $\varphi'$ results in a $\mathbb{Z}_n$ holonomy. Note that the disappearance the KK gauge bosons in dimensional quenching (as compared with dimensional reduction), implies that there is no actual magnetic flux on the 2-plane, yet there is a non-trivial holonomy, as required to describe a $\mathbb{Z}_n$ charged object. This is precisely the nontrivial holonomy specified in the construction of the mapping torus. We refer to these $\mathbb{Z}_n$ defects as 'quenched fluxbranes'.

Note that, in contrast with actual fluxbranes, we do not require quenched fluxbranes to solve the equations of motion of the spacetime effective theory. We use the construction to characterize the relevant topology describing $\mathbb{Z}_n$ defects. That is, tachyon condensation has provided us with a new topological sector of the theory, characterized by some charges. Solving the equations of motion in this sector will yield the actual on-shell configuration which realizes these charges. These solutions may be interesting on their own as they could be unstable (it could be energetically

[8]Here, 'fluxbrane' means 'extended solution with non-trivial (compactly supported) magnetic field in the transverse dimensions'. This is not to be confused with the more recent use of the term as branes carrying worldvolume magnetic flux [56].

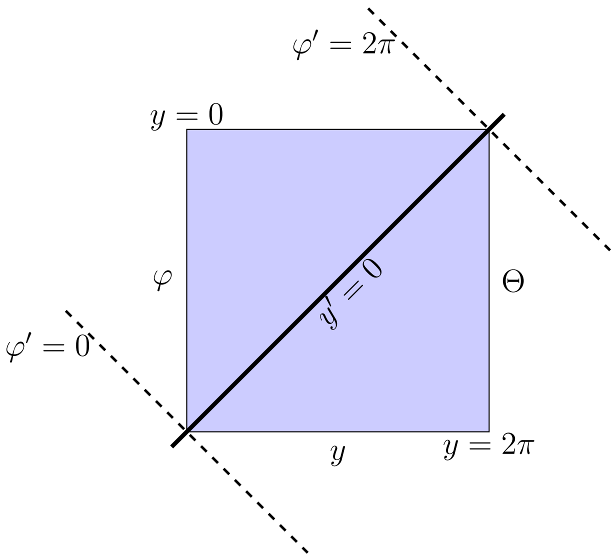


**Figure 3.3:** A schematic representation of the quenched fluxbrane construction. $\varphi$ is the polar angle in the $(X^8, X^9)$ plane, while $y$ is the coordinate of the mapping torus. The theory at $y = 0$ is glued to the theory at $y = 2\pi$ modulo an action of the discrete symmetry generator $\Theta$. Ordinary tachyon condensation would eliminate the critical dimension and constrain the theory to live in the critical slice $y = 0$. The fluxbrane is constructed by choosing a different tachyon profile depending on a tilted coordinate $y'$, such that the condensation leaves the $y' = 0$ slice instead. By choosing a new coordinate $\varphi'$, orthogonal to $y'$, we see that upon a full turn on $\phi'$ the theory we reach the right side of the rectangle and therefore pick a $\Theta$ action. Since this construction is valid everywhere except at the origin in the $(X^8, X^9)$ plane, it describes a string-like defect located at the origin in this plane, such that by circling it we pick the action of the discrete symmetry $\Theta$.

favourable for the string to become thicker and thicker).

The construction makes manifest that $\mathbb{Z}_n$ defects are conserved modulo $n$. Indeed, $n$ $\mathbb{Z}_n$ defects correspond to a fluxbrane with trivial monodromy: going around it once implies moving $n$ times around $S^1$ in the mapping torus in Figure 3.3. The configuration can be trivialized by a coordinate reparametrization.

### 3.3.3 Examples

We now provide some specific examples of the mapping torus construction.

#### Spacetime parity revisited

Consider the realization of a spacetime $\mathbb{Z}_2$ parity in e.g. the bosonic theory. Differently from Section 3.2.1, the 27d supercritical geometry has the dimension $X^{25}$ fibered non-trivially along an extra $S^1$ parametrized by $y$, forming a Möbius strip. In this non-orientable geometry, spacetime parity is a $\mathbb{Z}_2$ subgroup of a continuous KK $U(1)$. The symmetry breaking is triggered by closed string tachyon condensation. The $\mathbb{Z}_2$ defects of the 26d theory, regions around which spacetime parity flips, can be constructed as quenched fluxbranes with a tachyon condensate (3.3.2), with the replacement $y \to y' = y - R\varphi$, where $\varphi$ is the angle in the 2-plane transverse to the defect.

### $\mathbb{Z}_2$ symmetries of heterotic theories

We can easily implement the mapping torus construction to realize continuous versions of certain discrete symmetries of 10d heterotic theories. In order to have an extra (translational invariant) $S^1$, rather than an orbifold, we exploit the $HO^{+(1)}$ theory of Section 2.1.5. For instance, we can propose a different $U(1)$ embedding[9] of the $\mathbb{Z}_2$ symmetry of the 10d $SO(32)$ theory of Section 3.2.2, as follows. To reproduce the $\mathbb{Z}_2$ holonomy along the $S^1$, the supercritical $HO^{+(1)}$ theory should have an $\mathbb{Z}_2$ Wilson line (a constant gauge field background which however is not pure gauge and has effects on the physics) along the mapping torus $S^1$, introducing a $-1$ phase on $SO(32)$ spinors, e.g.

$$A = \frac{1}{R}\,\mathrm{diag}\,(\,i\sigma_2, 0, \ldots, 0) \quad \to \quad \exp\left(\frac{1}{2}\int_{S^1} i\sigma_2\, dy\right) = -\mathbf{1} \tag{3.3.3}$$

where the factor $\frac{1}{2}$ corresponds to the charge of spinors.

By constructing the quenched fluxbrane associated to this symmetry, we only take the tachyon profile to depend on $y'$ instead of $y$, as usual. The resulting 7-brane is a gauge defect around which spinors pick a $-1$ phase. This defect can be analyzed in familiar gauge theory terms, as is done in [37]. It is a singular representative of the class of gauge bundles associated to the nontrivial element of the homotopy group $\pi_1(SO(32)) = \mathbb{Z}_2$. This solution is unstable, and tends to decay into gauge flux.

Consider a second example, given by the $\mathbb{Z}_2$ symmetry exchanging the two $E_8$'s in the 10d $E_8 \times E_8$ heterotic, i.e. an outer automorphism. The gauge nature of this symmetry, argued in [61], can be made manifest using the mapping torus construction in the supercritical $E_8 \times E_8$ theory mentioned in Section 2.1.5. In this case, we must introduce a permutation Wilson line along the $S^1$, similar to those in CHL strings in lower dimensional compactifications [62, 63] (see also [64, 65, 66] for permutation Wilson lines in toroidal orbifolds).

### Discrete isometries of $T^2$

We now give new examples based on discrete isometries of $T^2$. To prevent a notational clash, we use $\beta \simeq \beta + 2\pi$ to parametrize the supercritical $S^1$.

The construction of the mapping torus for $T^2$ is basically an orbifold of $T^3 = T^2 \times S^1$, by a discrete isometry in $T^2$ and a simultaneous shift in $S^1$. These kind of models are described in string theory by free worldsheet CFTs and are familiar in orbifold constructions (in critical strings), see e.g. [67]. Instead, we recast the

[9] Incidentally, the same discrete symmetry may have different supercritical embeddings into continuous symmetries. This is similar to embedding the same $\mathbb{Z}_n$ symmetry into different continuous $U(1)$ groups, in fixed dimension.

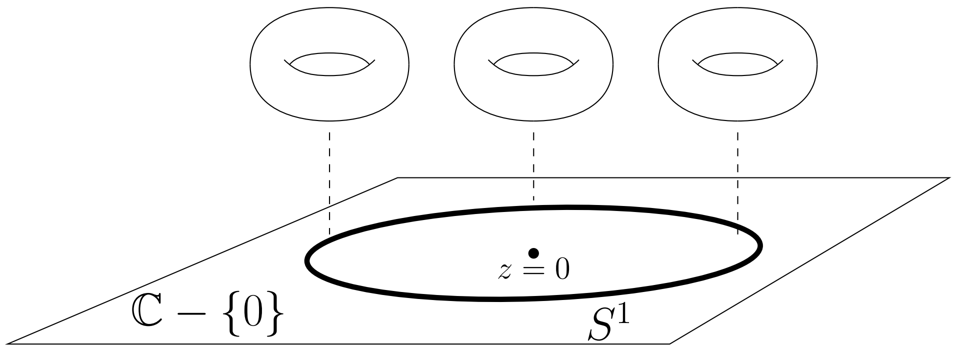


**Figure 3.4:** The mapping torus construction for $T^2$ can be regarded as a nontrivial torus fibration over $S^1$. We may embed this $S^1$ as the $|w| = 1$ locus of the complex plane, and extend the fibration to $\mathbb{C}^*$ to yield a torus (in fact, elliptic for our purposes) fibration in the complex plane.

construction in a language which will admit an easy generalization to Calabi-Yau spaces in projective spaces[10]. In order to exploit the power of complex geometry, let us extend $S^1$ to $\mathbb{C}^* \equiv \mathbb{C} - \{0\}$, by introducing a variable $w = |w|e^{i\beta}$ (which can eventually be fixed to $|w| = 1$ to retract onto $S^1$). The mapping torus is now associated to an elliptic fibration over $\mathbb{C}^*$, with constant $\tau$ parameter on the fiber, and suitable $SL(2, \mathbb{Z})$ monodromies around the origin (see Figure 3.4). This allows us to harness the power of algebraic geometry and in particular all the technology of elliptic fibrations common in F-theory [68]. It turns out that holomorphic fibrations suffice for our purposes. Indeed, the constant $\tau$ holomorphic fibrations discussed in the context of F-theory [69, 70], can be readily adapt to the present (non-compact) setup.

Consider a Weierstrass fibration over a complex plane $w$

$$y^2 = x^3 + f(w)x + g(w). \tag{3.3.4}$$

This defines a torus fibration as follows. The coordinates $x, y, z$ are defined on a weighed projective space $\mathbb{P}_{[3,2,1]}$, such that points $(x, y, z)$ and $(\lambda^3 x, \lambda^2 y, \lambda z)$ for any complex $\lambda \neq 0$ are identified For each $w$, the equation (3.3.4) defines a torus embedded in $\mathbb{P}_{[3,2,1]}$, see for instance [71]. From a constant $\tau$ fibration is achieved by [69]

$$f(w) = \alpha\, \phi(w)^2 \quad , \quad g(w) = \phi(w)^3 \tag{3.3.5}$$

with $\phi(z)$ some polynomial. The value of $\tau$ is encoded in $\alpha$.

In order to describe the $\mathbb{Z}_2$ $x \to -x, \quad y \to -y$, which exists for generic values of $\tau$, we simply choose $\phi(w) = w$, and have

$$y^2 = x^3 + \alpha\, w^2\, x \,+\, w^3 \tag{3.3.6}$$

Moving along $S^1$ (namely, $w \to e^{i\delta\beta} w$), the coordinates transform as $x \to e^{i\delta\beta} x$, $y \to e^{3i\delta\beta/2} y$. The holonomy along $S^1$ is $x \to x$, $y \to -y$, precisely the desired $\mathbb{Z}_2$ action.

[10] Recall [10] that Calabi-Yau spaces are of phenomenological interest since they preserve some supersymmetry in the low-energy effective field theory.

The construction of $\mathbb{Z}_2$ defects is now straightforward. We simply introduce a complex coordinate $z$ for the two real transverse dimensions, and consider the configuration obtained from (3.3.6) by the replacement $w \to w + z$ (the tachyon profile will be a radial extension of the profile around $S^1$ and hence will force $|z| = |w|$. The condition $\Im(z + w) = 0$ then enforces $y' = \varphi - y/R = 0$, as usual) . Note that two $\mathbb{Z}_2$ strings are described by a fibration with $\phi = w^2$ (since we must pick up twice the phase as we move around the string), which can be made trivial by a reparametrization,

$$y^2 = x^3 + \alpha\, w^4\, x \,+\, w^6 \;\longrightarrow\; y'^2 = x'^3 + \alpha x' + 1 \quad \text{with} \;\; y = w^3 y',\;\; x = w^2 x'. \tag{3.3.7}$$

This represents the physical fact that two $\mathbb{Z}_2$ string can annihilate.

A similar discussion can be carried out for the other holomorphic $\mathbb{Z}_n$ actions of the torus. Skipping further details, we simply quote the relevant fibrations:

$$\begin{aligned} \mathbb{Z}_4 \;&: \qquad y^2 = x^3 + w\,x \\ \mathbb{Z}_6 \;&: \qquad y^2 = x^3 + w \\ \mathbb{Z}_3 \;&: \qquad y^2 = x^3 + w^2 \end{aligned} \tag{3.3.8}$$

These fibrations differ slightly from those in [70], because the latter describe crystallographic actions on global geometries. We note that the local behavior of their fibrations around fixed points on the base is equivalent to ours, modulo reparametrizations describing creation/annihilation of $n$ $\mathbb{Z}_n$ defects.

### Discrete isometries in Calabi-Yaus: the quintic

The above strategy generalizes easily to more general Calabi-Yau spaces, as we illustrate for the quintic $X_6 = \mathbb{P}_5[5]$. This is complex dimension three Calabi-Yau space defined by a rank-five polynomial on five-dimensional projective space $\mathbb{P}_5$ in which $(z_1, z_2, z_3, z_4, z_5) \sim \lambda^5(z_1, z_2, z_3, z_4, z_5)$. Its expression at the so-called Fermat point is

$$z_1^5 + z_2^5 + z_3^5 + z_4^5 + z_5^5 = 0. \tag{3.3.9}$$

This polynomial guarantees that the resulting space is Calabi-Yau [72]. Compact Calabi-Yau spaces do not have continuous isometries[11], but can have discrete ones. This fact is apparent from (3.3.9); the equation is invariant under several diffeomorphisms such as for instance permutation of two coordinates. We simply focus on the $\mathbb{Z}_5$ generated by $z_1 \to e^{2\pi i/5} z_1$, with $z_2, \ldots, z_5$ invariant. The fibration associated to the mapping torus can be written

$$w\, z_1^5 + z_2^5 + z_3^5 + z_4^5 + z_5^5 = 0 \tag{3.3.10}$$

[11] If they did, lowering the index of the Killing vector via the metric tensor would yield a holomorphic $(1, 0)$-form, but $h^{1,0} = 0$ for Calabi-Yau manifolds.

Setting $w = e^{i\delta y}$, moving along the $S^1$ gives $w \to e^{i\delta y} w$ and $z_1 \to e^{-i\delta y/5} z_1$, so that completing the circle implements the desired $\mathbb{Z}_5$ monodromy. The construction of $\mathbb{Z}_5$ strings in the 4d theory amounts to a reinterpretation of $w$ in terms of the transverse coordinates, as in the previous section.

Note the important point that motion in $w$ does not correspond to changing the moduli of $X_6$. Complex structure moduli are described by deformations of the defining equation corresponding to monomials $\prod_i (z_i)^{n_i}$'s with $n_i < 4$[12]. This means that as one moves around the string there is no physical scalar which is shifting. This is fine because the monodromy is not part of a continuous gauge symmetry acting on any scalar of the 4d theory (as it disappears from the theory in the tachyon condensation).

These discrete isometries are phenomenologically relevant, since they often correspond to discrete R-symmetries of the 4d effective theory. The discussion of possible applications of our tools to phenomenologically interesting discrete R-symmetries will not be addressed in this work.

#### Antiholomorphic $\mathbb{Z}_2$ and CP as a gauge symmetry

A final class of discrete isometries of Calabi-Yau compactifications are given by antiholomorphic $\mathbb{Z}_2$ actions, e.g. $z_i \to \bar{z}_i$, which are large isometries of the Calabi-Yau spaces with defining equations with real coefficients. These are orientation-reversing, and hence are not symmetries of the superstrings, but can be actual symmetries if combined with an extra action. For instance, their combination with 4d parity gives a discrete symmetry, which in heterotic compactifications corresponds to a CP transformation [73]. Applying the mapping torus construction to this $\mathbb{Z}_2$ symmetry results in a description of CP as a discrete gauge symmetry explicitly embedded in a (supercritical) $U(1)$ symmetry. This is a new twist in the history of realizing CP as a gauge symmetry, see e.g. [61, 74]. Of course, CP is not a symmetry of 4d physics, so if it is a gauge symmetry it must be explicitly broken by some field.

#### $\mathbb{Z}_n$ symmetries already in $U(1)$ groups

Although we have focused on discrete symmetries from (large) isometries, the constructions can be applied to general $\mathbb{Z}_n$ discrete symmetries, even those embedded in continuous $U(1)$ factors already in the critical string theory (see [75, 76, 77, 78, 79, 80, 81] for such symmetries in string setups). Focusing on the 4d setup for concreteness, recall [42] (also Section 3.1.2) that the key ingredient is a $U(1)$ group with potential $A$, acting on a real periodic scalar $\phi \simeq \phi + 1$ as

$$A \to A + d\lambda \quad , \quad \phi \to \phi + n\lambda \tag{3.3.11}$$

[12] These deform the projective surface defined by (3.3.9); changing the embedding of the tangent space at each point. This changes the complex structure of the manifold, which is $PJP$ where $P$ is projection onto the tangent space to the manifold and $J$ is the complex structure of the ambient space.

There is an unbroken $\mathbb{Z}_n$, preserved even by non-perturbative effects, since it corresponds to a global gauge transformation. For instance, the axion field will couple in a natural way to instanton fields via a coupling of the form $\phi F \wedge F$ in the lagrangian. The coupling is anomalous, and in order to have gauge-invariant amplitudes, the amplitude in the presence of an instanton at a point $P$ must be dressed as

$$e^{-i\phi} \exp(2\pi i n \int_L A) \tag{3.3.12}$$

which describes the emission of electrically charged particles of total charge $n$ (i.e. preserving the $\mathbb{Z}_n$) along semi-infinite worldlines $L$ starting at $P$.

Let us now consider embedding this $\mathbb{Z}_n$ symmetry as a mapping torus construction in a supercritical extension of the theory. Along the extra $S^1$ there is a non-trivial $U(1)$ transformation (integrating to the $\mathbb{Z}_n$ generator) and a corresponding shift $\phi \to \phi + 1$. In other words there is one unit of flux for the field strength 1-form $F_1 = d\phi$

$$\int_{S^1} F_1 = 1 \tag{3.3.13}$$

Notice that the mapping torus of length $2\pi R$ and the $n$-cover circle of length $2\pi n R$ provide a physical realization of the two $S^1$'s in [77], associated to the periodicity of $\phi$ and of the $U(1)$.

Fields with charge $q$ under the $\mathbb{Z}_n$ have $S^1$ boundary conditions twisted by $e^{2\pi i\, q/n}$, hence have KK momenta $k + q/n$, with $k \in \mathbb{Z}$, and so carry charge under the KK $U(1)$. This piece allows to recover (3.3.12) in this picture, as follows. The operator $e^{-2\pi i \phi}$ at a point in the $S^1$ (and at point $P$ in the critical spacetime) picks up phase rotations under translation, i.e. under a KK $U(1)$ transformation, which must be cancelled by those insertions of KK modes, with total KK momentum 1, e.g. $n$ states of minimal $\mathbb{Z}_n$ charge.

The discussion in the previous paragraph has a nice string theory realization in the context of $\mathbb{Z}_n$ symmetries arising from the $U(1)$ gauge groups on D-branes. In particular we focus on D6-brane models c.f. [75], where $A$ is the gauge field on D6-branes, $\phi$ is the integral of the RR 3-form $C_3$ over some 3-cycle $\Sigma_3$, and the instanton is an euclidean D2-brane on a 3-cycle $\Sigma_3$. The non-trivial shift of $\phi$, namely the flux (3.3.13), corresponds to a 4-form field strength flux

$$\int_{\Sigma_3 \times S^1} F_4 = 1 \tag{3.3.14}$$

The D2-brane instanton on $\Sigma_3$ is not consistent by itself, but must emit particles with one unit of total KK momentum. This is more clear in the T-dual picture, which contains a D3-brane on $\Sigma_3 \times S^1$ with one unit of $F_3$ flux over $\Sigma_3$, which must emit a fundamental string with one unit of winding charge [82]. Another way to see this inconsistency is directly from the Chern-Simons couplings in the $D6$-brane worldvolume. An euclidean instanton can be constructed as a gauge field configuration with nontrivial $F \wedge F$ in the 4 noncompact dimensions. Then, the

Chern-Simons coupling

$$S_{CS} \supset \int_{D6} C_3(F \wedge F) = \int_{\mathbb{R}^4} \left( \int_{\Sigma_3} C_3 \right) F \wedge F = \int_{\mathbb{R}^4} F \wedge F, \tag{3.3.15}$$

which is precisely the kind of axion-instanton coupling discussed above.

### 3.3.4 Dual versions

All the previous examples involve discrete symmetries which may be realized in geometric terms, as gauge field or metric configurations. The mapping torus construction is however much more general, and can also accomodate genuinely stringy discrete symmetries. In this Section we focus on this class of discrete large symmetries, which involve stringy non-geometric transformations. The corresponding mapping torus constructions can be regarded as containing non-geometric fluxes, in analogy with the duality twists in e.g. [67].

#### Quantum symmetry in orbifolds

Consider an orbifold compactification to four dimensions with parent space $Y_6$ and orbifold group $\mathbb{Z}_n$ with generator $\Theta$ (which may also act on the gauge bundle in heterotic models). A general result from the structure of worldsheet amplitudes is the existence of a ‘quantum’[13] $\mathbb{Z}_n$ discrete symmetry (different from the $\mathbb{Z}_n$ generated by $\Theta$). The generator $g$ acts on fields in the $\Theta^k$-twisted sector with a phase $e^{2\pi i\, k/n}$. Invariance under this $\mathbb{Z}_n$ leads to restrictions in amplitudes usually known as ‘point group selection rules’, and have been discussed in the context of heterotic compactifications (see [83, 10] for reviews, and e.g. [84, 85] for early and recent applications). These symmetries have not been embedded in a continuous group.

For the present purposes, we may focus on the simple non-compact setup of $\mathbb{C}^3/\mathbb{Z}_n$, and take $\mathbb{C}^3/\mathbb{Z}_3$ as illustrative example. In fact, the quantum symmetry in this case is part of the $\Delta_{27}$ discrete symmetry analyzed in [86] (see [87] for generalizations involving $\mathbb{Z}_n$). It is straightforward to apply the mapping torus construction in Section 3.3 to derive this symmetry from a continuous $U(1)$ acting on an extra $S^1$. The construction is however particularly interesting because it involves a non-geometric $\mathbb{Z}_n$ action.

It is worthwhile to point out that the quantum symmetry can be geometrized by application of mirror symmetry [88]. For instance, focusing on $\mathbb{C}^3/\mathbb{Z}_3$, the theory is dual to a mirror geometry described in e.g. [89] (in the setup of D3-branes at singularities), and can be recast as the Calabi-Yau manifold described by

$$z = uv, \quad y^2 = x^3 + zx + 1. \tag{3.3.16}$$

[13]It is quantum in the sense of the worldsheet $\alpha'$ expansion. The symmetry is not visible in the large volume limit, because blow-up modes (which control the volume of the compactification manifold) are in twisted sectors and thus $\mathbb{Z}_n$-charged, so their vevs break the symmetry.

This is a Weierstrass fibration, similar to the ones studied in the previous section. The geometry is a double fibration over the complex plane $z$, with one $\mathbb{C}^*$ fiber parametrized by $u, v$ (degenerating over $z = 0$), and the second fiber being a $T^2$, degenerating over three points on the $z$-plane.

Since mirror symmetry is a duality transformation of the full string theoy, the original $\mathbb{Z}_3$ symmetry must be also present in some way. Indeed, its mirror is generated by

$$x \to e^{2\pi i/3} x \quad , \quad z \to e^{-2\pi i/3} z \quad , \quad u \to e^{-2\pi i/3} u \tag{3.3.17}$$

This preserves the holomorphic 3-form $\Omega = \frac{dx}{y} \wedge \frac{du}{u} \wedge dz$. Note that the action restricted to $z = 0$ is given by the $\mathbb{Z}_3$ isometry of $T^2$ generated by $x \to e^{2\pi i/3} x$. Therefore, mirror symmetry renders the embedding of this $\mathbb{Z}_3$ symmetry similar to the examples analyzed in Section 3.3.3.

### T-duality in type II $T^2$ compactifications

T-duality is a fundamental duality, i.e. a discrete transformation implying the equivalence of different string models, or different points in the moduli space of a single string model. In systems mapped to themselves by a T-duality transformation, the latter should manifest as a discrete symmetry of the configuration. For instance, the 26d closed bosonic string theory on an $S^1$ of radius $R = \sqrt{\alpha'}$ is self-T-dual. In this case, the $\mathbb{Z}_2$ symmetry is actually part of an enhanced $SU(2)^2$ gauge symmetry at the critical radius [6, 90].

For 10d type II theories we can obtain self-T-dual configurations by considering an square $T^2$ with equal radii $R = \sqrt{\alpha'}$, and vanishing B-field, i.e. $T = B + iJ = i$. This configuration is fixed by a $\mathbb{Z}_4$ subgroup of the $SL(2, \mathbb{Z})$ duality group, generated by

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{3.3.18}$$

However, there is no enhanced continuous gauge symmetry at this point, and the gauge nature of the $\mathbb{Z}_4$ symmetry is not as manifest as in the bosonic case.

This is however easily achieved in terms of a mapping torus configuration in a supercritical extension of the theory. The $\mathbb{Z}_4$ symmetry is thereby embedded in a continuous KK $U(1)$, making its gauge nature manifest.

Notice that the above matrix (3.3.18) corresponds to the $\mathbb{Z}_4$ isometry of $T^2$ given by $x \to -x$; $y \to iy$. Rather than a coincidence, this is because both systems are related by application of one-dimensional mirror symmetry (T-duality along one dimension), which exchanges the $T^2$ Kähler and complex structure parameters $T \leftrightarrow \tau$. Therefore the mapping torus construction in this Section is mirror dual to one of the torus isometries considered in Section 3.3.3.

We conclude with a final remark regarding the mapping torus construction for the $\mathbb{Z}_4$ self-T-duality. It provides an example of non-geometric fluxes in extra supercritical dimensions. Namely, as the theory moves along the extra $S^1$ it suffers a non-geometric transformation, which is not a diffeomorphism, but is a symmetry of string theory. This is a particular instance of the non-geometric (but locally geometric) $Q$-fluxes in [91], which have been concretely considered in orbifold language in e.g. [67]. Clearly the appearance of non-geometric fluxes along the supercritical $S^1$ will occur in the mapping torus construction of any discrete symmetry described by a non-geometric symmetry on the underlying compactification space.

#### Type IIB $SL(2,\mathbb{Z})$ $S$-duality and F-theory

Type IIB supergravity has a continuous $SL(2,\mathbb{R})$ symmetry group, which shifts the sign of the dilaton field. Since the dilaton is related to the string coupling as $g_s = \exp(\Phi)$, this is a weak-to strong coupling duality, or $S$-duality. In the full string theory, only a $SL(2,\mathbb{Z})$ lattice survives and constitutes the quantum $S$-duality group of type IIB string theory. This is related to the existence of Dirac quantization conditions which are absent in the classical theory.

This $S$-duality group acts on the complexified axio-dilaton $\tau = C_0 + i\frac{e^{-\Phi}}{2\pi}$ via Möbius transformations. It has an abelian subgroup consisting of $C_0$ shifts $\tau \to \tau + 1$, which may be embedded via a standard mapping torus construction. In fact, as discussed in the introduction, the axio-dilaton may be regarded as the complex structure parameter of a torus and so the symmetry may be embedded via the Weierstrass fibrations discussed as an example in the previous section[14]. Specifically, the monodromy $\tau \to \tau + 1$ corresponds to the Weierstrass fibration [68]

$$y^2 = x^2 + xz + w^3 + \frac{w}{4} \tag{3.3.19}$$

Notice that, unlike in previous cases, the monodromy cannot be readily seen from the Weierstras equation as some action on $x$ and $y$ as we turn around. This is so because the uniformization mapping [71] which takes us from the torus coordinates to the Weierstrass form automatically respects this periodicity. A way to see that it indeed carries the right monodromy is to compute the $j$-invariant,

$$j(\tau) = \frac{4(24f^3)}{27g^2 + 4f^3} = \frac{96w^3}{27w^2 + \frac{35}{2}w^3 + \frac{27}{16}w^4}. \tag{3.3.20}$$

This function has a pole at $w \sim 0$ of order one. Since $j(\tau)$ only diverges near $\Im(\tau) \to \pm\infty$, we may use the asymptotic expansion $j(\tau) \sim e^{2\pi i \tau}$. From this expression, $\tau \sim 2\pi i w$, and we see that near the origin $\tau$ picks a monodromy $\tau \to \tau + 1$, as expected[15].

[14]Notice that supercritical type 0 theory forces us to include two extra dimensions to achieve this. The extra two dimensions are, however, not related in any way to the 12-dimensional description of $F$-theory.

[15]The discriminant of the elliptic fibration (3.3.19) also vanishes at other loci far from the origin. We consider our fibration to be defined on a disk with $|w|$ small enough so that these points are outside of it. Recall that the mapping torus in this case is defined by retraction to an $S^1$ within the complex plane, so there is no problem in doing this if the $S^1$ is close to the origin.

The charged codimension 2 object associated to this symmetry are $D7$-branes (since around these objects the RR 0-form shifts by a period). Notice that these charged objects do not have fractional momentum along the compact direction as the states in the previous Section did. This is so because the index of the abelian discrete symmetry group is not finite, and the covering torus of radius $nR$ described in the previous Section is actually the real line; there is no requirement that the wavefunctions be periodic after some number of turns around the mapping torus. Physically, upon an axion field shift all observables acquire a trivial phase. Monodromies may be detected however by looking at fieldstrengths; the force between a stack of $D7$-branes and a $\overline{D7}$ brane is a direct measure of the axio-dilaton monodromy around the first.

Similarly, one can try to embed the other generator of $SL(2,\mathbb{Z})$, a $\mathbb{Z}_2$ which acts on the axio-dilaton as $\tau \to -1/\tau$. This actually corresponds to the $\mathbb{Z}_4$ symmetry discussed in the previous section. This $\mathbb{Z}_4$ symmetry is realized as a $\mathbb{Z}_2$ in F-theory because the modular group is more nearly $PSL(2,\mathbb{Z})$ than $SL(2,\mathbb{Z})$ [71]. This means that modular transformations that differ by the nontrivial element of the center of $SL(2,\mathbb{Z})$, namely the action $x \to -x$, $y \to -y$ are physically equivalent.

The codimension 2 object associated to this discrete symmetry is a $\mathbb{Z}_2$ string around which there is no well defined weak-coupling region; the monodromy takes a weak coupling region to a strong coupling one. In fact, the particular constant $\tau$ fibration discussed in Section 3.3.1 is not very interesting, since the axio-dilaton $\tau$ is fixed to $\tau = i$ (hence physics is everywhere nonperturbative). The most general fibration with this monodromy has $\tau$ of the form

$$\tau = ie^{\sqrt{f(w,\bar{w})}}, \tag{3.3.21}$$

where $f(w,\bar{w})$ has a zero of order 1 at the origin and vanishes nowhere else on the subset of $\mathbb{C}-\{0\}$ where the fibration is defined. Taking for instance $f(w,\bar{w}) = w$ one can achieve regions where physics is locally perturbative (large imaginary part of $\tau$), although globally regions of strong coupling are unavoidable.

Two of these 7-branes can ne nucleated form the vacuum, as can be seen as follows: Consider the elliptic fibration with $\tau = ie^{w}$, with smooth $f$ and a asymptotically flat metric. This is a trivial fibration describing some nontrivial dilaton profile over otherwise flat space. Now, parametrize $w = \sqrt{z^2}$. We have done nothing as the function $\sqrt{z^2} = z$ throughtout the complex plane. However we can smoothly deform the fibration to $w = \sqrt{z^2 - a^2} = \sqrt{(z+a)(z-a)}$. Near $z \approx a$, the fibration has the form (3.3.21). Notice that the coordinate $w$ for which the metric is asymptotically locally flat suffers a monodromy $w \to -w$. By making $a$ very large, one can get an asymptotically locally flat metric in $w$ around $z \approx a$ to arbitrarily good accuracy. Therefore, these $\mathbb{Z}_2$ 7-branes may be regarded as a particular $\mathbb{C}/\mathbb{Z}_2$ orbifolds. Finally, we note that this object has the same monodromy as a configuration of two $(1,0)$ 7-branes on top of a $(1,-1)$-brane, and the same monodromy as the $\mathbb{Z}_8$ orbifold supporting tensionless strings described in [92]; the difference in spacetime curvature of both defects is due to the fact that the defects studied in that the elliptic fibration is required to be Calabi-Yau in that reference.

# 4

# Heterotic NS5-branes from closed-string tachyon condensation

In this Chapter we continue the study of the condensation of tachyon profiles other than the vacuum in a more geometrical setup than in the previous chapter. We will focus on the $HO^{(+n)/}$ model, which is especially interesting because via the S-duality proposal of [4] (see Section 2.1.4), we can identify defects constructed in the $HO^{(+n)/}$ with some brane in type I string theory.

The classification of possible topological defects in type I theory was tackled by Witten in [37], where he showed that all these defects are classified by real K-theory. We will see that a similar classification holds for the heterotic theory, in a manner consistent with the S-duality. We will start by providing an exact CFT description of the heterotic NS5 in the supercritical theory, giving several arguments based on anomaly cancellation both from spacetime and worldsheet perspectives. We will then describe other solitons which can be obtained from tachyon condensation, corresponding to different K-theory classes.

The results of this Chapter are published in [93]. Appendix B contains a short review on K-theory and its applications to tachyon condensation.

## 4.1 NS5-branes from tachyon condensation

Every string theory has among its massless fields an antisymmetric tensor $B$, which couples directly to the string worldsheet $\Sigma$ as $\int_\Sigma B$. It is believed [42] that in a consistent quantum theory of gravity, for any field there are both electric and magnetic sources. For the $B$-field, whose gauge-invariant fieldstrength is called $H$, this would be an $H$-monopole, a codimension 4 object which is commonly referred to as NS5-brane, since the $B$ field comes from a NS-NS sector in the worldsheet.

Most topological defects in string theory do not have a simple perturbative description or other tractable tools to construct them, except as supergravity solutions. This includes the heterotic NS5-brane, whose near-core region is inherently non-perturbative due to a blowing-up coupling, even if the asymptotic coupling con-

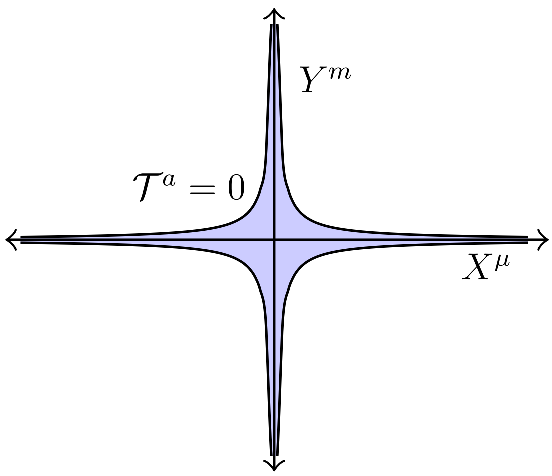


**Figure 4.1:** Schematic representation of tachyon condensation. The region where the tachyon vanishes is shaded in blue. In the homogeneous process described in Chapter 2, the tachyon only vanishes along the critical slice $Y^m = 0$. However, with a non-homogeneous profile, the tachyon has nonvanishing rank also at $X^\mu = 0$; formally, the extra dimensions "survive" there. The low energy theory after tachyon condensation describes this as soliton.

stant is kept perturbative [94]. Because of this, no general worldsheet description of a single heterotic NS5-brane has been found (see, however, [95] for a CFT description of an heterotic background containing NS5-branes in a double scaling limit). This is not true in the type II context, where the CHS worldsheet CFT [96] describes a stack of $N \geq 2$ NS5 branes. These can also be related in an indirect way to tachyon condensation in non-supersymmetric backrounds [32]. However, such a CFT cannot exist for heterotic NS5 branes. An argument due to Witten [97] starts with the observation that a heterotic NS5 is essentially a supersymmetric zero-size instanton, and one can guess part of the matter content of the theory by demanding that the moduli space includes that of a pointlike instanton. The resulting matter would have a far too large central charge to be described by a CFT.

To evade the above argument, one must give up either the central charge requirement or supersymmetry. Both things happen in the context of supercritical string theories. We will show that NS5-branes are associated to non-trivial instanton backgrounds in these bundles (all the other possible cases will be discussed in Section 4.2). One can motivate this by noticing that S-duality turns $D5$-branes into NS5-branes. If, presumably, the rotational and gauge bundles of the supercritical theory correspond to the $D9$ and $\overline{D9}$ gauge bundles, it seems natural to assume that a type I $D9$ worldvolume gauge instanton, which describes a $D5$, should describe an NS5 in the heterotic theory.

In Chapter 2, we described how by turning on the appropriate tachyon profile one may recover the $SO(32)$ heterotic vacuum from the $HO^{(+n)}$ theory. The matrix $M^a_\mu = \partial_\mu \mathcal{T}^a$, responsible for giving masses to the supercritical worldsheet fields, was taken to be position-independent. We now generalize the analysis to $X$-dependent tachyon background matrices, with rank jumping at particular loci. The idea is illustrated in figure 4.1

In principle, one could just write down tachyon profiles with rank jumps anywhere. However, there is no reason to expect that there is an actual marginal perturbation with these properties; the rank jumps may disappear once we flow to the IR to get an exactly marginal profile. A simple way around this difficulty is to consider cases in which the tachyon matrix zeroes are topologically protected. This is the case if we allow for non-trivial bundles $V$, $V'$ for the $SO(n) \subset SO(32+n)$ gauge group and the geometric $SO(n)_{\text{rot}}$ rotation group.

The matrix $M = \partial T$ is a section of $V \otimes V'^{*}$, and necessarily has zeroes over particular loci in the 10d slice, as discussed in AppendixB. Tachyon condensation leads to the appearance of topological brane defects located at those loci, allowing for the description of heterotic branes as closed string tachyon solitons.

The above argument assumes that the tachyon profile will be nonsingular throughout space, at least at early times before the onset of the condensation. We can argue that this is the case via an adiabatic argument: Consider an instanton of very large size. Classical equations of motion, which imply a nonsingular tachyon profile, are then valid to an arbitrarily good accuracy.

To build the tachyon profile, we need to know the topological charges of the NS5-brane. The Poincaré dual 4-form to the NS5-brane worldvolume is

$$\delta_4(\text{NS5}) = \text{tr}\, F^2_{\text{gauge}} - \text{tr}\, R^2_{\text{rot}}. \tag{4.1.1}$$

The tachyon profile can be specified by the Atiyah-Bott-Shapiro construction [98], described in detail in AppendixC, which we summarize briefly now. Take a 4d subspace of the 10d slice, parametrized by coordinates $\vec{X}$, in the theory with $n = 4$ extra dimensions, and regard the $SO(4)$ bundles $V$, $V'$ as chiral spinor bundles $S^{\pm}$. Take the (derivative of the) tachyon as the map $S^{+} \to S^{-}$ (determined by Dirac matrices $\Gamma$)

$$\partial_m \mathcal{T}^a \sim \vec{\Gamma}_{ma} \cdot \vec{X}. \tag{4.1.2}$$

This tachyon soliton carries the appropriate topological charge, and upon tachyon condensation leaves a codimension 4 defect localized at the vanishing locus $\vec{X} = 0$.

In order to show that this defect is actually an NS5-brane, the difference of instanton numbers in the bundles must be a magnetic source for the 2-form $B_2$, i.e. it must obey an anomalous Bianchi identity

$$\frac{1}{\alpha'} dH_3 = \text{tr}\, F^2_{SO(32+n)} - \text{tr}\, R^2_{10d} - \text{tr}\, R^2_{\text{rot}} \tag{4.1.3}$$

The first two terms are a generalization of the coupling in the critical 10d heterotic theory; the extension to $SO(32+n)$, required by gauge invariance and already noted in [4], reproduces the first term in (4.1.1). The second term is expected from invariance under diffeomorphisms mixing the 10d slice and the supercritical dimensions in the theory before the orbifold projection. We will show how its presence is required by anomaly cancellation from a number of different viewpoints. In addition, we also give a worldsheet argument that the physics at $\vec{X} = 0$ indeed corresponds to the NS5-brane singular CFT [94].

### 4.1.1 Anomalies in 2d worldsheet

As discussed in Section 2.1.1 for the conformal symmetry, some worldsheet symmetries are often anomalous. In the case of a continuous global symmetry in the worldsheet, it was argued in Section 3.1.2 that these correspond to gauge anomalies in spacetime. Hence, we must demand cancellation of anomalies of continuous global symmetries in the worldsheet. The $HO^{(+n)/}$ theory has a global $SO(32+n) \times SO(n) \times SO(1,9)$ global symmetry, where the first factor corresponds to gauge transformations and the second to (normal) target space diffeomorphisms. The Bianchi identity (4.1.3) is in fact essential in the cancellation of anomalies of these two factors, as we now show. Gauge anomalies in $2n$ dimensions can be easily studied via the so-called anomaly polynomial [99], a $2n+2$ form which will vanish in a non-anomalous theory. Using the worldsheet content, the relevant anomaly polynomial is

$$I_4 = \alpha'(\mathrm{tr}\, F^2_{SO(32+n)} - \mathrm{tr}\, R^2 - \mathrm{tr}\, R^2_{\mathrm{rot}}) \equiv dQ_3, \tag{4.1.4}$$

where we have introduced the gravitational, normal bundle and gauge Chern-Simons 3-forms (defined locally by $\mathrm{tr}\, F^2 = d\omega_{\mathrm{YM}}$, etc), via

$$Q_3 \equiv \omega_{3\,\mathrm{YM}} - \omega_{3\,\mathrm{L}} - \omega_{3\,\mathrm{rot}}. \tag{4.1.5}$$

The 2d anomaly $\delta_\Lambda S_{2d} \equiv \Lambda I_2$ is obtained from the descent relations $\delta_\Lambda Q_3 = \Lambda dI_2$, with $\Lambda$ a transformation parameter. This is canceled by an anomaly inflow as follows. The Bianchi identity (4.1.3) implies that

$$H_3 = dB_2 + Q_3 \tag{4.1.6}$$

with the same Chern-Simons form $Q_3$. The kinetic term produces interactions (in conventions of [6])

$$S_{B_2} = \int_{10+n} |H_3|^3 \to \int_{10+n} *H_3 \wedge Q_3 \tag{4.1.7}$$

Since the string worldsheet $\Sigma_2$ is electrically charged under $B_2$, its Poincaré dual $\delta(\Sigma)$ acts as a source

$$d * H_3 = \delta(\Sigma_2) \tag{4.1.8}$$

Target diffeomorphisms and gauge transformations produce a variation of (4.1.7) given by

$$\int_{10+n} *H_3 \wedge \delta_\Lambda Q_3 = \Lambda \int_{10+n} *H_3 \wedge dI_2 = -\int_{\Sigma_2} I_2 \tag{4.1.9}$$

where in the last step we integrated by parts and used (4.1.8). This cancels the anomaly from (4.1.4).

### 4.1.2 The 10d anomaly argument

The anomalous Bianchi identity for $H_3$ is a crucial ingredient in the Green-Schwarz anomaly cancellation mechanism in the critical 10d heterotic theory. The extension (4.1.3) can therefore be obtained by extending the anomaly cancellation analysis to supercritical theory in the presence of a non-trivial normal bundle. Using the fermion spectrum of the $HO^{(+n)}/\mathbb{Z}_2$ theory described in Section 2.1.4, the anomaly polynomial can be shown to factorize as $I_{12} = Y_4 Y_8$, where $Y_8$ is a degree-8 polynomial in curvatures and field-strengths, and

$$Y_4 = \alpha'(\text{tr}\, R_{10d}^2 - \text{tr}\, F^2 + \text{tr}\, R_{\text{rot}}^2) \tag{4.1.10}$$

Incidentally, the computation is isomorphic to that in [14, 100] in a different system [1]. Anomaly cancellation is achieved through a Green-Schwarz mechanism which involves a coupling $\int B_2 Y_8$ and the Bianchi identity (4.1.3).

As a further consistency check, the integrated version of the Bianchi identity can be probed by compactifying the 10d slice into a 4-manifold $X_4$, for instance K3[101] (as is familiar from supersymmetric compactifications). The K3 compactification of the gravitinos produces an anomaly exactly as in the critical case, see e.g. [102]. A direct test of eq. (4.1.3) is the computation of the pure gravitational anomaly, whose cancellation requires a net multiplicity of 244 positive chirality fermions in the theory. The spectrum of 6d massless fermions is determined by the index theorem in terms of characteristic classes of the tangent bundle, the gauge bundle, and the normal $SO(n)_{\text{rot}}$ bundle.

Consider the theory with $n = 4$ supercritical dimensions, and introduce arbitrary instanton numbers $k, k'$ in $SU(2) \subset SO(4)$ factors of the gauge and normal bundles. The decomposition of the charged fermion spectrum of the $HO^{(+n)/}/\mathbb{Z}_2$ under the $SO(32)$ and instanton $SU(2)$ groups is

$$\begin{array}{lll} & SO(36) \times SO(4)_{\text{rot}} & \supset SO(32) \times SU(2) \times SU(2)_{\text{rot}} \\ \text{Ferm}_+ & (\begin{smallmatrix}\square\\ \square\end{smallmatrix}, 1) & \to (\begin{smallmatrix}\square\\ \square\end{smallmatrix}, 1, 1) + 2(\square, 2, 1) + \\ & & +(1, 3, 1) + 3(1, 1, 1) \\ \text{Ferm}_- & (\square, \square) & \to \quad 2(\square, 1, 2) + 4(1, 2, 2) \\ \text{Ferm}_+ & (1, \square\!\square) & \to \quad 3(1, 1, 3) + (1, 1, 1) \end{array}$$

The net multiplicities we get for the different $SO(32)$ representations are:

$$\#(\begin{smallmatrix}\square\\ \square\end{smallmatrix}) = -1 \ , \ \#(\square) = k - k' \ , \ \#(\mathbf{1}) = -2(k - k') \tag{4.1.11}$$

Cancellation of the gravitational anomaly requires that $k - k' = 24$. This is the familiar statement that heterotic K3 compactifications requires instanton/NS5-brane number 24. In our setup, this number includes contributions from the extra gauge and normal bundles, which thus carry instanton/NS5-brane charge. Since the 10d

[1] Actually, this is because the fermion quantum numbers of the $HO^{(+n)}/\mathbb{Z}_2$ theory are identical to those arising in type I with additional $D9 - \overline{D9}$ brane pairs [100], as noticed in [4].

gauge group $SO(32)$ is unbroken, configurations with different instanton numbers in the bundles $V$, $V'$ flow upon tachyon condensation to the critical heterotic theory with explicit NS5-branes (rather than finite size instantons).

### 4.1.3 The worldsheet argument

What the above arguments really show is that the codimension 4 leftover of tachyon condensation carries the right NS5-charge, but this is not the same as actually showing it is an NS5. To show this, let us return to the case of flat non-compact 10d space, with $n = 4$ supercritical dimensions, and non-trivial gauge and/or normal bundles. We will analyze the worldsheet dynamics near the vanishing locus of the (derivative of the) tachyon. As above, we take the instanton to be nontrivially fibered over a $\mathbb{R}^4 \subset \mathbb{R}^{9,1}$, parametrized by $X^\mu$ coordinates, and take the gauge bundle to be nontrivial only for a $SO(4) \subset SO(36)$ subgroup. The normal coordinates are still denoted by $Y^m$. In the ABS ansatz (4.1.2) (see also AppendixC), $\partial_n \mathcal{T}^a$ vanishes at $X = 0$, so near the vanishing locus the $Y$ coordinates remain massless and seemingly parametrize an extra branch of the CFT. In other words, tachyon condensation is unable to "kill" the extra dimensions precisely at its vanishing locus.

As show in AppendixC, the tachyon with the right charges is

$$\mathcal{T}^a = X^\mu (\bar{\eta_\mu})^{am} Y^m, \tag{4.1.12}$$

where the $\bar{\eta}_\tau$ are the 't Hooft matrices, defined in AppendixC. This conclusion is independent of which bundle has been twisted, be it the normal or the gauge bundle. This tachyon profile couples to the worldsheet fields through (2.2.12). In this case, the only fields with non-trivial coupling to the tachyon profile are the $(4+4)$ bosons $X^\mu$, $Y^m$, their right-moving fermion superpartners, and the 4 left-moving fermions $\lambda^a$ of $SO(4) \subset SO(36)$. This matter content can be arranged in $(0,4)$ multiplets (containing either four real right-moving fermions and four real scalars, or one left-handed fermion). The theory actually has $(0,4)$ susy at the level of interactions as well, since the Yukawa couplings and the F-term potential turn out to obey the ADHM equations, as in the analysis in [103] (not surprisingly, since our system describes instanton backgrounds). Moreover, the precise sigma model arising from (4.1.12) can be shown to correspond to the singular CFT arising in the zero instanton size limit in [94] (times the decoupled extra bosons and fermions, which are not relevant for the discussion).

This can be shown as follows: First, we can use the identity

$$(\bar{\eta})^\mu_{am} = \mathrm{Tr}[\epsilon\overline{\sigma}^a \epsilon\sigma^\mu \epsilon\sigma^n], \tag{4.1.13}$$

where $\epsilon$ is the totally antisymmetric $2\times 2$ matrix. Substituting this in (4.1.12), we may write

$$\begin{aligned}\lambda^a\mathcal{T}^a &= \frac{1}{2}\mathrm{Tr}\epsilon[\bar{\sigma}^a\epsilon\sigma^\mu\epsilon\sigma^m]\lambda^a X^\mu Y^m = \frac{1}{2}(\lambda^a\sigma^a_{YY'})\epsilon^{Y'Z'}(X^\mu\sigma^\mu_{BZ})\epsilon^{YZ}(Y^m\sigma^m_{B'Z'})\epsilon^{BB'}\\ &= \bar{\lambda}_{YY'}X^Y_B Y^{Y'}_{B'}(\frac{1}{2}\epsilon^{BB'}),\end{aligned} \tag{4.1.14}$$

where uppercase letters are $SU(2)$ indices. This peculiar notation is used to facilitate comparison with eq. (3.2) of [103], with $c^{BB'} = \frac{1}{2}\epsilon^{BB'}$. This explicitly shows that the above choice of tachyon profile results in the zero-size instanton nonlinear sigma model of [103], whose Yukawa couplings obey the ADHM equations ensuring $(0,4)$ SUSY.

Thus, the infrared behavior of the system near the vanishing locus of the tachyon profile (4.1.12) is exactly the CFT in a zero size instanton (by definition, NS5-brane) background. Deformations of this tachyon profile lead to marginal deformations in the IR CFT. As argued in [103], the Coulomb branch parametrized by the $Y$ is decoupled from the physical (in our setup) Higgs branch parametrized by the $X$ coordinates. Rather than physical spacetime coordinates sticking out from $X = 0$, the Coulomb branch signals that the worldsheet theory is the singular CFT describing the presence of a background NS5-brane (small instanton as in [97]) at $X = 0$.

In the next Chapter we will se a similar example laid out in more detail in the context of type 0 supercritical theories.

Notice that, even though we have seen the NS5 manifest itself as a singular CFT, this is due to the quadratic approximation (4.1.12) for the tachyon profile, which breaks conformal invariance. A dynamical tachyon profile will be an exactly marginal perturbation. Therefore supercritical string theory provides an exact, nonsingular worldsheet description of strings propagating in the background of a NS5-brane, even though we are not able to provide an explicit example.

## 4.2 K-theory and other heterotic topological defects

Closed string tachyon condensation allows for the nucleation/annihilation of pairs of bundles associated to the supercritical geometry and the gauge group. Specifically, the pair of bundles $(V, V')$ and $(V \oplus H, V' \oplus H)$ for any real bundle $H$ of rank $k$ will yield the same state after tachyon condensation, since one can first do tachyon condensation in the $HO^{+(n+k)/}$ theory along the $H$-twisted directions to yiend the original $HO^{+(n)/}$ theory. Physically, this corresponds to the nucleation of brane-antibrane pairs from the vacuum. This process does not change the topological charges of the state and thus one should regard two configurations related by brane-antibrane annihilation as equivalent. As discussed in AppendixB, the equivalence classes of bundles $(V, V')$ over a base manifold $M$ modulo addition of real bundles of arbitrary rank conform the real K-theory group $KO(M)$. As advertised, the discussion is very similar to that of D-brane charges in type I theories, as suggested by the S-duality proposed in [4] and discussed in Chapter 2. Two signficant differences are that the K-theory class representatives in the type I case contain two gauge bundles, whereas in our case they contain a gauge bundle and a geometric one, and that different representatives of the same K-theory class actually live in spacetimes of different dimension.

The corresponding real K-theory thereby classifies the charges of topological defects realized as closed string solitons. For instance, our construction of the NS5-brane relates it to $KO(S^4) = \mathbb{Z}$, and unveils a K-theoretic structure in the anomalous Bianchi identity for $H_3$, refining its earlier interpretation [104] in de Rahm or integer cohomology — analogous to the fact that RR fluxes in type II theories are best understood in the context of K-theory [105].

The non-trivial real K-theory groups motivate the construction of other defects as closed tachyon solitons. In particular, the fundamental string can be identified as a soliton whose charge is $KO(S^8) = \mathbb{Z}$, i.e. the second Pontryagin class at the cohomological level. This approach can also be exploited to construct $\mathbb{Z}_n$ charges in the heterotic. For instance, the non-trivial element in $KO(S^2) = \mathbb{Z}_2$ corresponds to a $\mathbb{Z}_2$ charged 7-brane, built from tachyon condensation in $n = 2$ supercritical dimensions, with normal or gauge $SO(2) \equiv U(1)$ bundles with

$$\int_{\mathbb{R}^2} F_{\text{gauge}} - \int_{\mathbb{R}^2} R_{\text{rot}} = 1 \text{ mod } 2\mathbb{Z} \tag{4.2.1}$$

This charge can be detected by considering its relative holonomy with the dual torsion object, a particle classified by $KO(S^9) = \mathbb{Z}_2$. The latter is actually realized as a perturbative massive state, transforming as a bi-spinor of the $SO(32+2)$ gauge group and the normal $SO(2)$, which indeed picks up a $-1$ holonomy when moved around the defect associated to bundles satisfying (4.2.1). It is also possible to show that the 7-brane $\mathbb{Z}_2$ charge is detected as a global gauge anomaly [53] on an NS5-brane probe of the system, in the spirit of [106]. The tachyon soliton realization of this 7-brane was proposed in [13]. Similarly, the groups $KO(S^1) = KO(S^{10}) = \mathbb{Z}_2$ are associated to closed tachyon solitons describing an 8-brane and an euclidean instanton.

The $\mathbb{Z}_2$ charged objects can all be related to properties of the gauge group which are manifest in perturbative heterotic string theory. The 7-brane describes a gauge bundle with a nontrivial holonomy for fermions, which as discussed above implies that the dual object is a fermion and the gauge group must actually be (a quotient of) a spin cover of $SO(32)$. The 8-brane actually describes an $O(32)$ bundle, and the fact that there is a dual object (the instanton) means that the $O(32)$ is broken to $SO(32)$.

# 5

# Closed tachyon solitons in type II string theory

In the previous Chapter we used tachyon condensation to construct interesting solitons of $SO(32)$ heterotic string theory, but as discussed in Chapter 2 we also have a supercritical version of type 0 superstring theories. Furthermore, a particular orbifold of the supercritical type 0 is also related to type II theories.

In this Chapter we will explore supercritical string theories, their rich anomaly structure, and the branes constructed as solitons of the tachyon. Similarly to the heterotic example, the charges of these branes are classified by real K-theory. We also apply these ideas to reinterpret the worldsheet gauged linear sigma model (GLSM), regarded as a supercritical theory on the ambient toric space with closed tachyon condensation onto the CY hypersurface. We also suggest the possible applications of supercritical strings to the physical interpretation of the matrix factorization description of F-theory on singular spaces.

## 5.1 Codimension four solitons

As for the heterotic tachyons of Chapter 4, nontrivial tachyon profiles in type 0 or type II theories can be associated to different kinds of solitons. In order to connect with the 10d type II theories, we consider the orbifolded version of supercritical type 0 theory described in Section 2.1.6. We split the coordinate fields as the 'critical' $X^\mu$, $\mu = 0, \ldots, 9$, and the 'supercritical' $x_m$, $y_m$, with $m = 1, \ldots, k$. The orbifold action flips $k$ extra coordinates $y^m \to -y^m$, while leaving the others invariant $x^m \to x^m$.

As discussed in Chapter 2 around equations (2.2.21) and (2.2.22), the type 0 tachyon gives a potential term

$$V = \frac{\alpha'}{16\pi} G^{MN} \partial_M \mathcal{T} \partial_N \mathcal{T} \tag{5.1.1}$$

One can now study tachyon profiles which lead to dimension change within type 0 theories. This is achieved by pairing the extra coordinates, with mass terms

$$W = \mathcal{T} = \mu \exp(\beta X^+) \mathscr{T}_{mn}\, x_m\, y_n. \tag{5.1.2}$$

where we take $\mathscr{T}_{mn}$ to be a matrix of rank $k$ at generic points of the critical dimensions, given in terms of the tachyon as

$$\mathscr{T}_{mn} = \partial_{x_m}\partial_{y_n}\mathcal{T} \tag{5.1.3}$$

where in this last expression we ignore the dependence in $X^+$.

As discussed in Appendix C, the different topological classes of tachyon solitons are classified by homotopy groups. Let us consider the tachyon background associated to $\Pi_3(SO) = \mathbb{Z}$, which describes supercritical bundles with non-trivial first Pontryagin class. For concreteness, we consider $4+4$ supercritical dimensions, and fiber them such that the $SO(4)_x \times SO(4)_y$ bundles have non-trivial $\text{tr}\, R_x^2 - \text{tr}\, R_y^2$ over an $\mathbb{R}^4$ in the 10d critical slice.

Using the ABS tachyon as prototype, the tachyon condensation removes the supercritical dimensions everywhere except at the origin in $\mathbb{R}^4$. Just like for open string tachyon solitons, we expect this to signal the presence of a left-over real codimension-4 topological defect. One may be tempted to propose that this is an NS5-brane, as occurred in the supercritical heterotic case of the previous chapter. In the following we show that it rather corresponds to a gravitational soliton, carrying the gravitational charges of $\mathbb{C}^2/\mathbb{Z}_2$, namely, a localized contribution to $\text{tr}\, R^2$, where here $R$ is the curvature 2-form of the tangent bundle in $\mathbb{R}^4$.

The tachyon and its condensation proceeds essentially in the same way in both type IIA and type IIB theories. We focus on the latter case, since the nature of the endpoint and the corresponding charges are easier to identify. This is because in type IIB theory ADE geometries localize chiral field content at the singularities, whereas for type IIA theory the localized modes inherit the non-chiral nature of the ambient 10d theory. The result of type IIB theory can be translated to type IIA theory by a T-duality along the worldvolume dimensions of the defect (even in the supercritical setup).

In addition, in setups with one compact $\mathbb{S}^1$ dimension transverse to the defect it is possible to T-dualize the gravitational instantons into NS5-branes. The corresponding configurations and relevant technical details are discussed in appendix C.2. This hence provides a realization of the NS5-brane as a closed tachyon soliton.

### 5.1.1 Chiral worldvolume content

Although the tachyon condensation process is highly non-trivial, one could hope to compute the worldvolume chiral spectrum by simple dimensional reduction of the 10d fields using the coupling to the physical tachyon background, as done in [4]. However, although our ABS-like profile (taking (5.1.2) of the form (4.1.2), see Appendix C) specifies the right K-theory class of the soliton, and it is expected to flow to the physical tachyon background, as it stands it does not solve the spacetime equations of motion. Therefore, even if we focus on the chiral sector, the computation of the spectrum is not safe against leaking of modes off to infinity in the supercritical dimension, as we now describe.

If we attempt the dimensional reduction of the fermions using the ABS tachyon profile, the following picture emerges. For each fermion, which in the supercritical theory lives in the $10+k$-dimensional orbifold fixed locus (we remind the reader that $k = 4$ in this section), one can construct two different sets of zero modes. One of them is localized along a 10-dimensional slice which will become the critical type II spacetime after tachyon condensation. Presumably these describe the 10d fermion content of the theory, as happens for the heterotic in [107]. There is a second set, which is instead localized along a different 10-dimensional slice, which intersects the critical slice on the worldvolume of the defect. This different set of fermions is associated with the $X^\mu = 0$ branch of the superpotential obtained from the ABS tachyon matrix (4.1.2); they may be thought of as a supercritical version of the localized worldvolume fermions of the gravitational instanton, before condensation. From the point of view of the critical slice, they are 6d localized fermions, even though they propagate in 4 extra supercritical dimensions. It is expected that the physical tachyon localizes the modes in the latter, leading to genuinely 6d chiral modes; this phenomenon is however not robust, as such modes can leak off to infinity along the 4 supercritical dimensions. From a different but related perspective, the computation of the chiral spectrum using the index theorem is ambiguous due to contributions from the $\eta$-invariant at infinity.

This picture suggests that localized fermions indeed arise from tachyon condensation with the ABS-like superpotential (4.1.2), but as stressed above, we cannot trust this tachyon profile in spacetime. Fortunately, other arguments, which we now proceed to review, show clearly that the endpoint of the condensation is a gravitational instanton.

### 5.1.2 Induced D-brane charges

The problem of tachyon condensation in supercritical type 0 theories with a codimension 4 ABS profile is in principle similar to what happened in the supercritical heterotic tachyon solitons of the previous chapter, with non-trivial first Pontryagin class in the supercritical dimensions. There, the localized defect was argued to be an NS5-brane, which actually carries the same charges as $\operatorname{tr} R^2$. What allowed to read the charges of the soliton more easily there was the existence of couplings such as $B_6 \operatorname{tr} R^2$ (where $B_6$ is the potential dual to the 2-form), which survived in the supercritical setup and included contributions such as $B_{6+n} \operatorname{tr} R_N^2$, where $R_N$ is the curvature of the normal bundle the critical slice. However, a similar argument does not apply straightforwardly to type 0 theories, since there are no topological couplings of this kind.

We can however use a different strategy to detect the physical charges carried out by supercritical dimension bundles with non-trivial first Pontryagin class. The key idea is to introduce D-brane probes, on which the topological charge tr $R^2$ will induce lower-dimensional D-brane charges.

D-branes in type 0 theories are discussed in e.g. [108, 109]; most of these results apply to the supercritical theories in a straightforward manner. Basically, there is one D$p$-brane per RR gauge potential (hence, two D-branes, dubbed 'electric' and 'magnetic' for each allowed value of $p$). On a D$p$-brane stack of fixed kind, the spectrum is given by gauge bosons and scalars associated to the transverse dimensions. For overlapping electric and magnetic D$p$-branes, the mixed sector gives rise to a Majorana-Weyl fermion of the rotation group of the full $(10+2n)$-dimensional space (we assume $n \in 4\mathbb{Z}$ here), suitably decomposed with respect to the worldvolume Poincaré group.

We now consider D-branes in the $\mathbb{Z}_2$ quotient describing the supercritical type II configuration. Specifically, we will consider only supercritical D-branes which are localized on a submanifold contained in the 10d critical slice, and which span all supercritical directions, including the orbifolded ones. In this way, tachyon condensation takes place along the worldvolume of the D-brane and we avoid the tricky question of determining what happens to other RR charges which exist in the supercritical theory but not in the type II endpoint.

Since the supercritical type II theories are constructed as $\mathbb{Z}_2$ orbifolds of supercritical type 0 theories, the D-branes of supercritical type II arise in the parent type 0 theory $\mathbb{Z}_2$ as invariant pairs of D-branes [1]. Since the $\mathbb{Z}_2$ exchanges the electric and magnetic D-branes of the parent type 0 theory, the supercritical type II D-brane is a coincident pair of electric and magnetic supercritical type 0 D-branes of the corresponding dimension. Note that we have to include electric and magnetic branes to guarantee invariance under the $\mathbb{Z}_2$ at the orbifold fixed locus; if the brane were away from it, it would be completely consistent to have a single electric or magnetic brane.

For concreteness, we focus on type 0B theory, in 10+4+4 dimensions, and consider a pair of electric and magnetic D13-branes (denoted $e$ and $m$ in the following) along the directions $x^0$, $x^1$ and the $\mathbb{R}^4$ parametrized by $(x^6,\ldots,x^9)$ in the 10d critical slice, and extending over the 4 orbifold-even $x$ coordinates as well as the 4 orbifold-odd $y$ coordinates. We are interested in computing the worlvolume Chern-Simons couplings to the curvatures, in order to read out the D-brane charges induced by the curvature of the normal supercritical dimensions. Such couplings are intimately related to the worldvolume fermion content, as required by anomaly inflow arguments [110, 111], so we determine the latter. As reviewed above, the D13$_e$-D13$_e$ and D13$_m$-D13$_m$ sectors do not contain fermions, while the mixed D13$_e$-D13$_m$ (and D13$_m$-D13$_e$) give rise to two fermions (one per sector) of fixed chirality with respect to the total $SO(1,9{+}4{+}4)$ group, decomposed under the $SO(1,5)\times SO(4)_{2345}\times SO(4)_x\times SO(4)_y$ group. This worldvolume spectrum is non-chiral in 14 dimensions but at the orbifold fixed locus only one linear combination of the fermions survives, leading to a localized contribution to the 10d anomaly.

[1]In fact, this is completely analogous to what happens for critical 10d type II theories, constructed as $\mathbb{Z}_2$ orbifolds of type 0 theories [108, 109].

To be more precise, the theory has two 14d fermions, of opposite 14d chirality. They can be decomposed in terms of 10d and 4d spinors as $|C\rangle_{10}|C\rangle_4+|S\rangle_{10}|S\rangle_4$ and $|C\rangle_{10}|S\rangle_4+|S\rangle_{10}|C\rangle_4$, where $|C\rangle$ and $|S\rangle$ denote spinors of opposite chiralities. The $\mathbb{Z}_2$ orbifold flips the sign of the second 14d spinor, which implies that one cannot regard the $\mathbb{Z}_2$ orbifold as having a simple geometric action on the 4d spinors (this would swap the sign of one of the two 4d spinors). Hence, the computation of the localized anomaly at the $y=0$ locus is subtle, but can be dealt with as follows. In fact, the action of the orbifold on the chiral fermions is a single copy of its action on the bispinor generating the bulk RR forms. Later, in Section 5.2, we will see that both anomaly cancellation in 0B theory and reduction to the anomaly polynomial in IIB after condensation requires a specific form for the anomaly polynomial of the RR forms.

Inspired by the form of this RR anomaly polynomial, a natural guess for the anomaly polynomial of the chiral fermions turns out to be

$$P_{10} = \frac{\hat{A}(R_T^{D13\cap\{y=0\}})}{\hat{A}(R_N^{D13\cap\{y=0\}})}\, e(R_y), \tag{5.1.4}$$

where $R_T^{D13\cap\{y=0\}}$ is the curvature of the tangent bundle to the intersection of the $y=0$ locus and D13-brane worldvolume, and $R_N^{D13\cap\{y=0\}}$ its normal bundle. The last factor, $e(R_y)$, is the Euler class of the $y$-bundle, which will localize anomalies in a 6-dimensional locus. This will become the D5-brane after tachyon condensation (see Section 5.2).

One can now apply standard anomaly inflow arguments [110, 111], considering two branes intersecting along the critical slice, to describe the Chern-Simons coupling on the D13-brane worldvolume in terms of curvatures of the corresponding bundles. We denote by $R$, $R_N$, $R_x$, $R_y$ the curvatures of the tangent bundle along the D13-brane, and along the normal bundles in the directions 2345 and the supercritical $x$ and $y$ directions, respectively. Since the $\hat{A}$-genus is a multiplicative class, we have

$$\hat{A}(R_T^{D13\cap\{y=0\}}) = \hat{A}(R)\hat{A}(R_x) \quad , \quad \hat{A}(R_N^{D13\cap\{y=0\}}) = \hat{A}(R_n)\hat{A}(R_y). \tag{5.1.5}$$

The resulting Chern-Simons couplings are

$$S_{\rm CS} \;=\; \int_{\rm D5} \sum_p C_p \wedge {\rm ch}(F) \left( \frac{\hat{A}(R)}{\hat{A}(R_N)}\,\frac{\hat{A}(R_x)}{\hat{A}(R_y)} \right)^{\frac{1}{2}} \tag{5.1.6}$$

(The Euler class contribution is rewritten here as the localization to the D5 slice on the critical directions.) Notice that the $x$ and $y$ bundles appear asymmetrically in this expression. This difference between the $x$ and $y$ bundles is crucial for the $K$-theory picture advocated throughout this Chapter to make sense. Namely, isomorphic bundles in the $x$ and $y$ dimensions should not generate new charges in the system, and indeed the contribution of isomorphic $x$ and $y$ bundles to induced D-brane charges cancel in (5.1.6); see below for an explicit example of this. Similar expressions can be obtained for other D$p$-branes. Note that in the critical case this reduces to the familiar expression in [112].

Consider now a D13-brane in flat space, $R = R_N = 0$, but in the presence of non-trivial bundles in the supercritical dimensions. Expanding the above we have

$$\frac{1}{48}\int_{\mathrm{D13}} C_{10} \wedge (\mathrm{tr}\, R_y^2 - \mathrm{tr}\, R_x^2). \tag{5.1.7}$$

This implies that in the presence of supercritical bundles with non-trivial first Pontryagin class, there is an induced (supercritical) D9-brane charge on the volume of the D13-brane. Upon tachyon condensation, the extra dimensions disappear everywhere except at the core of the tachyon soliton, while each brane loses 8 worldvolume directions along the $x, y$ directions[2]. The left-over defect in the endpoint critical type II theory must be such that it reproduces the induced D1-brane charge on a D5-brane probe. This nicely fits with the interpretation of the soliton as a localized defect supporting a non-trivial $\mathrm{tr}\, R^2$ along the tangent directions of critical type II theory.

### 5.1.3 Worldsheet CFT

The above arguments show that the endpoint of tachyon condensation is the critical type II theory with a localized, real codimension-4 object characterized by a nonvanishing $\mathrm{tr}\, R^2$, i.e. a gravitational instanton. We will now argue that the worldsheet theory with the ABS tachyon has the right properties for reproducing, after tachyon condensation, the singular worldsheet CFT for a $\mathbb{C}^2/\mathbb{Z}_2$ singularity.

The first piece of the argument concerns the topological charge of the gravitational instanton that we just found. As we describe in appendix C the supercritical tachyon $\mathscr{T}$ is obtained by embedding the $SU(2)$ instanton into $SO(4)$. The induced brane charge is given by the Pontryagin class of the real bundle, which is defined to be the second Chern class of its complexification, so the real embedding of the ABS instanton induces two units of topological charge, compatible with a $\mathbb{C}^2/\mathbb{Z}_2$ singularity.

A detailed analysis of the supercritical CFT gives further support for this identification, as follows. As usual, take $X^\mu$ to be the critical slice coordinates, $y^n$ the four supercritical orbifold-odd coordinates, and $x^m$ the supercritical orbifold-even coordinates. As discussed above, the matrix $\mathscr{T}$ (5.1.3) of tachyon derivatives with respect to the supercritical coordinates may be regarded a mapping between the orbifold-odd and orbifold-even supercritical tangent bundles. If we regard these as nontrivial bundles over the critical slice parametrized by the $X^\mu$, the tachyon describes a K-theory class over the critical slice, as described in Appendix B. In

[2] This follows from applying the worldsheet arguments in Chapter 2 to the open string sector describing the D-branes, i.e. the 2d fields associated to the $x, y$ directions are very massive and can be integrated out.

particular, the ABS tachyon (see Appendix C) (C.1.38)

$$\mathcal{T} = X^\tau (\overline{\eta}_\tau)_{ab} x^a y^b = \frac{1}{2}\mathrm{tr}\,(\tilde{\mathsf{x}}\mathsf{X}\mathsf{y}) \tag{5.1.8}$$

describes the K-theory class of a gravitational instanton. Notice the similarity to the heterotic tachyon (4.1.12) which yielded the heterotic NS5-brane. We have argued for this interpretation using spacetime considerations. In the following we will study the effects of this tachyon superpotential as a worldsheet perturbation.

The equations of motion coming from (5.1.8) (which, recall, couples to the worldsheet as a superpotential) give a branch structure. Consider for example $y \neq 0$. By a combined $SO(4)^3$ rotation we can put $y = (y_0, 0, 0, 0)$ with $y_0 \neq 0$ without changing $\mathcal{T}$. Now consider the equation of motion for $\widetilde{x}_\mu$, which is

$$\frac{\partial \mathcal{T}}{\partial \widetilde{x}_\mu} = \frac{1}{2}\mathrm{tr}\,(\sigma^\mu \sigma^\tau)\, y_0 X^\tau = \pm y_0 X^\mu = 0 \tag{5.1.9}$$

for all $\mu$. (The sign is determined by $\frac{1}{2}\mathrm{tr}\,(\sigma^\mu\sigma^\tau) = \mathrm{diag}\,(1,-1,-1,-1)$.) Since $y_0 \neq 0$, we have $X^\mu = 0$, and from the $X$ equation of motion also $\widetilde{x}^\mu = 0$. A completely analogous argument holds when $X$ or $\widetilde{x}$ are non-vanishing, so at the classical level we find a three-branch structure, with each branch being determined by which of $X, \widetilde{x}, y$ is non-vanishing, and the other two fields vanishing on that branch. The three branches meet at $X = x = y = 0$. This situation is very similar to that analyzed in [103], and for the same reasons as those given in there, we expect the branch structure to disappear quantum mechanically (we will review the details of the argument momentarily). This can occur in the full CFT once we include quantum corrections: the sigma model metric is corrected, and the branch intersection is pushed to infinite distance down an infinite throat developing at $X = 0$. Such an infinite throat behavior close to $X = 0$ is suggestive of light degrees of freedom (coming from wrapped branes) living there, again indicating a singularity in the target description. The real ABS instanton (5.1.8) is the minimal instanton we can construct starting with a $SU(2)$ instanton, so it is perhaps not surprising that the resulting configuration is the minimal CY two-fold singularity.

Finally, we will show that the sigma model given by (5.1.8) has $(4,4)$ worldsheet supersymmetry, as befits type II on $\mathbb{C}^2/\mathbb{Z}_2$. The argument is very similar to that presented in Chapter 4 for the heterotic case; however, the peculiarities of the type 0 tachyon make the comparison to known CFT's less straightforward and thus we lay out the argument in more detail.

Actually, proving $(0,4)$ will be enough; since the arguments will be completely left-right symmetric, this will also prove $(4,0)$. The fields come in supermultiplets of $(4,4)$ containing four bosons and four fermions of each chirality; we only have to show that the interactions preserve this supersymmetry. A $(1,1)$ superfield

$$\mathcal{X} \equiv X + \psi_+\theta_+ + \psi_-\theta_- + F\theta_-\theta_+ \tag{5.1.10}$$

can be decomposed as two $(0,1)$ superfields

$$\mathcal{X} \equiv (X + \psi_+\theta_+) + \theta_-\lambda_-, \quad \lambda_- \equiv \psi_- + F\theta_+. \tag{5.1.11}$$

Then one can compute the $(0,1)$ superpotential from the $(1,1)$ superpotential (5.1.8) by integrating over $\theta_-$, which yields:

$$W_{(0,4)} = \int d\theta_- \mathcal{T} = X_\mu \overline{\eta}^\mu_{mn} (\lambda^x)^m y^n + (\lambda^X)_\mu \overline{\eta}^\mu_{mn} x^m y^n + X_\mu \overline{\eta}^\mu_{mn} x^m (\lambda^n_y). \quad (5.1.12)$$

To show that the theory with superpotential $W_{(0,4)}$ actually has $(0,4)$ supersymmetry, it is enough to show that the ADHM equations are satisfied [103]. Following [93, 103], it is convenient to write down the indices of the above superfields in terms of the would-be $SU(2) \times SU(2)$ $R$-symmetry. Uppercase unprimed indices will transform as a doublet over the first factor, whereas uppercase primed indices will transform in the same way under the second. We will choose the $R$-symmetry to act on the scalar superfields in such a way that $X$ transforms as $X^{AB}$, $x$ transforms as $x^{YY'}$, and $y$ transforms as $y^{A'B'}$. The three fields transform differently under $R$-symmetries[3], and this choice will ensure the $(0,4)$ symmetry of the theory. We still have to choose how the Fermi superfields transform under the $R$-symmetry; there is no need to be consistent with the rules we set up above for the scalar superfields[4]. We will take

$$\lambda^X_{A'B'}, \quad \lambda^x_{Y'Y}, \quad \lambda^y_{AB}. \quad (5.1.13)$$

With this choice in mind, one may now use the identities

$$\begin{aligned} 2\overline{\eta}^\mu_{mn} &= \mathrm{Tr}[\epsilon \overline{\sigma}^m \epsilon \sigma^\mu \epsilon \sigma^n] = \overline{\sigma}^m_{E'A} \epsilon^{AB} \sigma^\mu_{BC} \epsilon^{CC'} \sigma^n_{C'D'} \epsilon^{D'E'} \\ &= -\overline{\sigma}^m_{CD'} \epsilon^{D'A'} \sigma^\mu_{A'B'} \epsilon^{B'E'} \sigma^n_{E'C'} \epsilon^{CC'} = -\overline{\sigma}^m_{EC'} \epsilon^{CC'} \sigma^\mu_{CA} \epsilon^{AB} \sigma^n_{BD} \epsilon^{DE}, \end{aligned} \quad (5.1.14)$$

to rewrite the superpotential (5.1.12) as

$$\begin{aligned} W_{(0,4)} &= \frac{1}{2} ((\lambda^x)^m \overline{\sigma}^m_{E'A}) \epsilon^{AB} (X^\mu \sigma^\mu_{BC}) \epsilon^{CC'} (y^n \sigma^n_{C'D'}) \epsilon^{D'E'} \\ &\quad - \frac{1}{2} (x^m \overline{\sigma}^m_{CD'}) \epsilon^{D'A'} ((\lambda^X)^\mu \sigma^\mu_{A'B'}) \epsilon^{B'E'} (y^n \sigma^n_{E'C'}) \epsilon^{CC'} \\ &\quad - \frac{1}{2} (x^m \overline{\sigma}^m_{EC'}) \epsilon^{CC'} (X^\mu \sigma^\mu_{CA}) \epsilon^{AB} ((\lambda^y)^n \sigma^n_{BD}) \epsilon^{DE} \\ &= \frac{1}{2} \epsilon^{CC'} \left[ \lambda^x_{E'A} X^A_C y^{E'}_{C'} - \lambda^X_{A'B'} x^{A'}_C y^{B'}_{C'} - \lambda^y_{BD} X^B_C x^D_{C'} \right]. \end{aligned} \quad (5.1.15)$$

Defining $c^{CC'} = \frac{1}{2} \epsilon^{CC'}$, the above may be rewritten as

$$\begin{aligned} W_{(0,4)} &= c^{CC'} \left( \sum_{a = X,x,y} C^a_{CC'} \lambda^a \right), \\ C^x_{CC'} &= X^A_C y^{D'}_{C'}, \quad C^X_{CC'} = -x^{A'}_C y^{B'}_{C'}, \quad C^y_{CC'} = -X^B_C x^D_{C'}. \end{aligned} \quad (5.1.16)$$

[3] There is another scalar multiplet of $(0,4)$ not considered in [103], namely $x^{YY'}$. The supersymmetry transformations are $\delta_\eta x^{YY'} = i\epsilon_{X'B'} \eta^{XX'} \psi^{B'Y'}_-$, $\quad \delta_\eta \psi^{B'Y'}_- = \epsilon_{YB} \eta^{BB'} \partial_- X^{YY'}$.

[4] Indeed, doing so would ruin $(0,4)$ supersymmetry.

Showing that the theory has $(0, 4)$ supersymmetry amounts to showing that the $C$'s satisfy the ADHM equations, explicitly given by (2.13) and (2.20) of [103], which is easy. The first constraints are

$$\frac{\partial C^a_{AA'}}{\partial X^{BY}} + \frac{\partial C^a_{BA'}}{\partial X^{AY}} = \frac{\partial C^a_{AA'}}{\partial x^{YY'}} + \frac{\partial C^a_{YA'}}{\partial x^{AY'}} = \frac{\partial C^a_{AA'}}{\partial y^{B'Y'}} + \frac{\partial C^a_{AB'}}{\partial y^{A'Y'}} = 0 \tag{5.1.17}$$

where $a$ runs over $X, x, y$ and we have omitted the upper indices. The three different $C$'s above satisfy the equations. Secondly, we have to show that the constraint

$$\sum_a C^a_{AA'} C^a_{BB'} + C^a_{BA'} C^a_{AB'} = 0 \tag{5.1.18}$$

is satisfied. Even though (5.1.18) involves a sum over the three different sets of fermions, for (5.1.16) the sum is identically zero for each fermion. This proves that the lagrangian under consideration actually has $(0, 4)$ supersymmetry. In the same way we may prove that it has $(4, 0)$, and since left and right-handed supercharges anticommute this shows $(4, 4)$ supersymmetry.

We can now argue as in [113], and follow what happens to the right-handed $SU(2) \times SU(2)$ $R$-symmetry as we flow to the IR (equivalently, as tachyon condensation takes place). The endpoint of the condensation must be a $(4, 4)$ nonlinear sigma model, and one of the two $R$-symmetry $SU(2)$'s must be included as part of the superconformal algebra. If we focus on the branch on which $X$ is nonzero, the $SU(2)$ factor which acts nontrivially on $X$ cannot be conserved and thus it must be the factor acting nontrivially on $x, y$ the one which becomes part of the superconformal algebra. The same argument then shows that in this sigma model $x, y$ must not show up: even if semiclassically there is a throat at $X = x = y = 0$ through which the three branches join, quantum mechanically they are split and there is no way to transition from one to the other. This is exactly what happened, in a simpler setup, in the heterotic case of the previous chapter.

To sum up, the theory perturbed by (5.1.8) flows in the IR to a superconformal $(4, 4)$ model, with the same topological charge as $\mathbb{C}^2/\mathbb{Z}_2$, and with singular behavior at the origin. It seems fairly natural to identify this theory as type II string theory on $\mathbb{C}^2/\mathbb{Z}_2$.

## 5.2 Codimension eight solitons

Let us consider the tachyon backgrounds associated to $\Pi_7(SO) = \mathbb{Z}$, which describe supercritical bundles with non-trivial second Pontryagin class. Using e.g. the profile (5.1.8) as prototype, the tachyon condensation removes the supercritical dimensions everywhere except at the origin in $\mathbb{R}^8$. We will argue that, in type IIA/0A theory, the left-over real codimension-8 topological defect actually corresponds to a fundamental string (upon T-duality on a circle, the winding string turns into a momentum mode in type IIB/0B theory). In order to show it, we derive a supercritical analog of the familiar 10d IIA 1-loop term $B_2 X_8$ [114], which shows that a configuration with nonvanishing second Pontryagin class in the supercritical dimensions carries the charge of an F1 string. The strategy to compute the 1-loop Chern-Simons term is to relate the computation, via a T-duality on a circle, with an anomaly cancellation computation in the supercritical type IIB theory (i.e. the $\mathbb{Z}_2$ orbifold of supercritical 0B).

### 5.2.1 The one-loop Chern-Simons term in 10d type IIA

Recall that 10d type IIA theory has the one-loop Chern-Simons term for its NSNS 2-form $B_2$ [114]

$$B_2\, X_8(R) \tag{5.2.1}$$

where $X_8$ is a polynomial in the curvature 2-form $R$. It can actually be shown to correspond to the anomaly polynomial of the 6d worldvolume field theory on a type IIA NS5-brane, since the above coupling cancels this anomaly by inflow [115], as we now review for completeness. The coupling induces an anomalous Bianchi identity for the field strength $H_7$ of the dual gauge potential

$$dH_7 = X_8 \quad , \quad H_7 = dB_6 + Q_7 \tag{5.2.2}$$

where $Q_7$ is the Chern-Simons form obtained from descent $X_8 = dQ_7$.

The kinetic term of $B_2$, recast using the dual potential $B_6$, and field strength $H_7$, displays an anomalous variation

$$S_{\rm kin} = \int_{10d} H_3 \wedge H_7 \quad , \quad \delta S_{\rm kin} = \int_{10d} H_3 \wedge \delta Q_7 = \int_{10d} H_3 \wedge dX_6 = -\int_{10d} dH_3 \wedge X_6 \tag{5.2.3}$$

where we have used the descent $\delta Q_7 = X_6$, and integration by parts. In the presence of an NS5-brane, $dH_3 = \delta_4(\text{NS})$, where $\delta_4$ is a bump 4-form transverse to the NS5-brane volume. Hence the variation is

$$\delta S_{\rm kin} = -\int_{10d} \delta_4(\text{NS}) \wedge \text{X}_6 = -\int_{\text{NS}} \text{X}_6 \tag{5.2.4}$$

and cancels the chiral anomaly of the NS5-brane worldvolume field theory if $X_8$ is the anomaly polynomial, as anticipated.

A simple way to compute the coupling (5.2.1) is to compactify on $\mathbb{S}^1$. In the resulting 9d theory, there is a coupling $A_1\, X_8(R)$, where $A_1$ arises from $B_2$ with one leg on $\mathbb{S}_1$. In the T-dual IIB, the $A_1$ field is a graviphoton and the corresponding coupling involves the KK gauge field coming from the metric. This coupling appears not from dimensional reduction of a one-loop 10d coupling, but rather as a genuinely 9d Chern-Simons coupling appearing from the parity anomaly [116, 117]. It can be easily obtained from the chiral type IIB field content, by integrating out the infinite towers of KK modes (which are charged under the KK gauge field) [118], as we now review.

Chiral fields harged fermions and four-forms in 9d suffer from a parity anomaly [116, 117]. The theory cannot be regulated in a gauge and parity-invariant way. The fermion determinant is not invariant under large gauge transformations and from the index theorem one gets that there is a term in the effective action of the form $\frac{1}{2}\int_{9d} I_9$, where $dI_9 = X_{10}$, the standard anomaly polynomial in 8d for Weyl fermions and self-dual 3-forms, respectively.

Within this $I_9$ term to the effective action there is an $A_1X_8$ term, coming from the $FX_8$ term in $X_{10}$. Here, $X_8$ is the 6d anomaly 8-form. The $A_1X_8$ contribution can be rewritten as a $F_2I_7$ contribution, via integration by parts.

This means that a very massive fermion or 4-form in 9d will yield the above effective action when integrated out, apart from other terms which die with inverse powers of the mass. All KK modes of chiral 10d fields contribute to the anomaly. The contribution to the effective action coming from $I_9$ is weighted by the sign of the 9d mass. The contribution to a term linear in $F$ such as the one we are concerned with is also proportional to the charge. This means that the KK reduction of a 10d chiral field yields a contribution

$$\frac{1}{2}\sum_{n=-\infty}^{\infty} n\,\mathrm{sign}(n)\int F_2 I_7 = \left(\sum_{n=1}^{\infty} n\right)\int F_2 I_7, \tag{5.2.5}$$

which is divergent. We can use zeta function regularization, which yields $-1/12$ for the sum. A way to justify this would be that Zeta-function regularization is the result of any regularization procedure which preserves diffeomorphism invariance along the $S^1$.

As a simple consistency check, the anomaly of a Weyl fermion should be the same computed in 10d or with the full KK tower in 9d. In 10d we get an anomalous variation of the action which includes a term coming from descent from $-\frac{1}{24}\int p_1(F)X_8(R)$, where we have factorized the 12d curvature in terms of a 4d curvature $F = F_0 - Ad\theta$, $d\theta$ being the P.D. to the $S^1$, and an 8d curvature. This yields a term in 9d of the form $-\frac{1}{24}\int d\Lambda X_8$, where $\Lambda$ is a $U(1)$ gauge transformation. On the other hand, the KK tower above yields an effective action which changes, upon a gauge transformation, as $\frac{S}{2}\int d\Lambda X_8$, where $S$ is the regularized sum over the KK tower. Equality of both anomalies clearly demands $S = -1/12$.

We now repeat the above procedure for all the IIB bulk fields:

- For the vector-spinor, we should get a term coming from $X_{10} = \hat{A}[ch_9(R) - 1] = \hat{A}[ch_{10}(R) - 2]$. The reason we only remove the contribution of one spinor instead of two is that one of the constraints (invariance under $\psi \to \psi + k\xi$, with $\xi$ a spin 1/2 field) is absent for a massive spinor. The KK modes are not independent due to the reality condition in 10d, so we should count only half of the KK modes, say those of positive momentum. But we got two identical vector-spinors, which cancels the preovious effect and overall we are left with the contribution

$$X_8^{v-s} = -\frac{1}{12}[\hat{A}[ch_{10}(R) - 2]]_8 \tag{5.2.6}$$

- For the 4-form, we get a tower of complex massive KK forms in 9d. The reality condition makes us sum only over positive momenta. We should also drop the factor of 1/2 in the anomaly polynomial since the 9d 4-forms are complex. Finally, the dimension of the Dirac rep. in 9d is the same as that of one Weyl in 10d, so unlike in standard anomaly cancellation we don't have to divide the index by two. So the parity anomaly will be obtained through descent procedure from a complex self-dual 4-form in 10d. The chiral anomaly of this is $-\frac{1}{2}L(R)$, and adding up all the states in the tower we get a contribution

$$X_8^{form} = \frac{1}{24}[L(R)]_8. \tag{5.2.7}$$

The total contribution to the $A \wedge X_8$ term after integrating out all these massive modes is

$$A \wedge (X_8^{form} + X_8^{v-s}) = A \wedge \left(\frac{1}{48}(trR^2)^2/(4\pi)^4 - \frac{1}{12}(trR^4)/(4\pi)^4\right), \tag{5.2.8}$$

which is the correct $A \wedge X_8$ term that we expect from T-duality. The type IIA polynomial $X_8$ is thus also given by the 9d type IIB parity anomaly, which in turn is directly related to the 10d type IIB anomaly polynomial by the descent relations.

In the coming sections, we use this logic (in reverse) to show that the supercritical IIA has a coupling that generalizes (5.2.1) and that includes the curvature of the normal bundle to the critical slice. We start with the computation of the supercritical IIB anomaly polynomial, compactify on $\mathbb{S}^1$ and compute the parity anomaly by the descent relations. Upon T-duality and decompactification, this allows to obtain the Chern-Simons coupling in supercritical IIA theory. We further show that the generalized polynomial is given by the anomaly polynomial of the NS-brane of supercritical IIA, as required by (a straightforward generalization of) the above worldvolume anomaly cancellation argument.

### 5.2.2 Anomaly cancellation in supercritical 0B

We start with the computation of anomaly cancellation of $(10 + 2k)$-dimensional supercritical IIB theory. The relevant sector is localized in the $(10+k)$-dimensional fixed locus $\Sigma$ of the $\mathbb{Z}_2$ orbifold; although the field content is non-chiral with respect to $(10+k)$ diffeomorphisms, it is chiral if we also consider diffeomorphisms in the directions of the $k$-dimensional normal bundle $N\Sigma$, and there is an associated anomaly polynomial. We denote by $R$ and $F$ the curvature 2-forms of the tangent bundle $T\Sigma$ and the normal bundle $N\Sigma$. Spinors of positive chirality will be denoted by $|s\rangle$, and those of negative chirality by $|c\rangle$.

The chiral spectrum (in the above sense) in the $(10 + k)$-dimensional fixed locus of the supercritical IIB theory is (see Chapter 2 for further details):

- (a) Two negative-chirality spinors of $SO(10+2k)$, with a Majorana condition, and charged in the vector representation $SO(10+k)$. The spinor decomposes as $|s\rangle|c\rangle + |c\rangle|s\rangle$ in terms of $SO(10+k) \times SO(k)$ reps. There is also a reality condition, giving a factor of 1/2 in the anomaly computation (which cancels with the fact that we have two such fields).

- (b) Two spinors of $SO(10+2k)$, charged as a vector of $SO(k)$. These become $|s\rangle|s\rangle + |c\rangle|c\rangle$ in terms of $SO(10+k) \times SO(k)$ reps.
- (c) A RR self-dual $(4+k)$-form, and an anti-self dual RR $(4+k)$ form.

To compute the anomaly polynomial, we introduce the Chern characters in the vector and spinor representations of $SO(n)$, $c(F)$,$c_s(F)$, $c_c(F)$. The contributions from the different fields are

- (a) There are $(10+k)$-dimensional gravitinos transforming as spinors of the normal bundle, which give:

$$-\Big[\hat{A}(R)\Big[\mathrm{ch}(R)-2\Big]\left(c_c(F)-c_s(F)\right)\Big]_{12+k} = \left[\frac{\hat{A}(R)}{\hat{A}(F)}\left[\mathrm{ch}(R)-2\right]\right]_{12} e(F). \tag{5.2.9}$$

Here, we have used the identity $(c_c(F) - c_s(F)) = e(F)/\hat{A}(F)$ for the Chern character for the spin cover of a bundle, with $e(F)$ being the Euler class of the normal bundle.

- (b) There are $(10+k)$-dimensional spinors, transforming as vector-spinors of the normal bundle, and give:

$$\Big[\hat{A}(R)(c_c(F)-c_s(F))c(F)\Big]_{12+k} = -\left[\frac{\hat{A}(R)}{\hat{A}(F)}c(F)\right]_{12} e(F). \tag{5.2.10}$$

- Self-dual forms: These are trickier to compute, as we need to know how the orbifold affects the forms to compute the anomalies. Recalling the earlier behaviour of spinor contributions, the natural result is

$$\left[\frac{1}{8}\frac{L(R)}{L(F)}\right]_{12} e(F). \tag{5.2.11}$$

Note that for $k=0$ this reproduces the standard result for 10d type IIB theory (using $e(F)=1$). The above generalization to arbitrary $k$ is also consistent with all our other computations. Indeed, this is the only choice ensuring cancellation of all anomalies, which are anyway located on the critical slice thanks to the $e(F)$ factor.

The total anomaly is

$$\left[\frac{\hat{A}(R)}{\hat{A}(F)}[ch(R)-2-c(F)] - \frac{1}{8}\frac{L(R)}{L(F)}\right]_{12} e(F) \tag{5.2.12}$$

In any situation leading to 10d IIB upon tachyon condensation, the bundles are such that $e(F)$ is merely localized onto the critical 10d slice. Hence, the above is the IIB anomaly polynomial, but with the characteristic classes evaluated for the $K$-theory element $(T\Sigma, N\Sigma)$ (rather than just the tangent bundle $T\Sigma$). Notice that this invariance depends crucially on the ansatz (5.2.11), which also ensures anomaly cancellation and reduction of the anomaly polynomial to the 10d one when the supercritical bundles are trivial.

### 5.2.3 Chern-Simons couplings

The strategy now is to compactify supercritical IIB theory on an $\mathbb{S}^1$, i.e. take supercritical 0B theory in $\mathbb{S}^1 \times \mathbb{R}^{9+k} \times \mathbb{R}^k/\mathbb{Z}_2$, and to compute the parity anomaly in the lower-dimensional theory. The only contributing fields are the KK modes of those fields entering into the anomaly of the uncompactified theory, by descent from the anomaly polynomial in $10 + 2k$ dimensions, as explained below.

All the fields contributing to the anomaly are sections of $T\Sigma \otimes N\Sigma$. Upon KK reduction we have the splitting $T\Sigma = T\Sigma' \oplus U(1)$, while $N\Sigma$ stays the same. We will refer to the curvature and gauge potential of the $U(1)$ factor as $G$, $A_G$, and denote by $R'$ the curvature of $T\Sigma'$.

Associated to each chiral field there is an anomaly polyform $P(R, F)$. Its degree-$(12 + k)$ part is the anomaly polynomial, whereas the degree-$(10 + k)$ part (modulo some factors which depend on the reality conditions for each field and which will be discussed later on) is the exterior derivative of the parity anomaly, as discussed above for the 10d case [116, 117]. The way to compute the CS term we are interested in is to expand $\mathrm{ch}(G)[P(R', F)]_{10+k}$, take the term linear in $G$, that is $GP(R', F)$, and then the anomaly term is $A_G P(R', F)$ where $dA_G = G$. We now compute these contributions from the different (KK towers of) fields:

- (a) For the vector-spinors of $SO(1, 9+k)$ with negative chirality, we get a term from

$$P_{10+n} = e(F)\frac{\hat{A}(R')}{\hat{A}(F)}[c(R')_{9+k} - 1] = \frac{\hat{A}(R')}{\hat{A}(F)}[c(R)_{10+k} - 2]. \qquad (5.2.13)$$

We only remove the contribution of one spinor instead of two because one of the constraints (invariance under $\psi \to \psi + p\xi$, with $\xi$ a spinor field, and $p$ the momentum) is absent for a massive spinor. The KK modes are not independent due to the reality condition in 10d, so we should count only half of the KK modes (say the positive ones), but this factor will cancel with the fact that there are two identical vector-spinors. Using a zeta function regularization $\sum_{n=0}^{\infty} n = -1/12$, we are left with the contribution

$$P^{(a)}_{8+n} = \frac{1}{12}\Big[\, e(F)\, \frac{\hat{A}(R')}{\hat{A}(F)}\, \big[c(R')_{10+k} - 2\big]\,\Big]_{8+k}. \qquad (5.2.14)$$

- (b) Similarly, for the vector-spinors of $SO(k)$, we get an anomaly

$$P^{(b)}_{8+k} = -\frac{1}{12}\Big[\, e(F)\,\frac{\hat{A}(R')}{\hat{A}(F)}\, c(F)\Big]_{8+k}. \tag{5.2.15}$$

- (c) The self-dual forms are trickier, but the natural guess is that they mimic an anti-self dual form for which the Chern character factor coming from the $N\Sigma$ contribution is precisely $\frac{e(F)}{L(F)}[\Sigma']c(G)$. With this proviso, we may analyze the parity anomaly. For the $(4+2k)$-form, we get a tower of complex massive KK forms in $9+2k$ dimensions. The reality condition restricts the sum to positive momenta, and we also drop the factor of $1/2$ in the anomaly polynomial since the 9d $(4+2k)$-forms are complex. Finally, the dimension of the Dirac spinor in $9+2k$ dimensions is the same as that of one Weyl spinor in $10+2k$, so unlike in standard anomaly cancellation we don't have to divide the index by two. So the parity anomaly is obtained through descent procedure from the chiral anomaly of a complex anti self-dual $(4+2k)$-form in $10+2k$ dimensions, namely $-\frac{1}{2}L(R)$. Adding up all the states in the tower plus the normal bundle contribution we get a contribution

$$P^{(c)}_{8+n} = -\frac{1}{24}\Big[\,\frac{L(R')}{L(F)}\, e(F)\Big]_{8+n}. \tag{5.2.16}$$

The total parity anomaly is

$$P_{8+k} = P^{(a)}_{8+k} + P^{(b)}_{8+k} + P^{(c)}_{8+k} = \frac{e(F)}{12}\left[\frac{\hat{A}(R')}{\hat{A}(F)}\big[\, c(R)_{10+k} - 2 - c(F)\,\big] - \frac{1}{2}\frac{L(R')}{L(F)}\right]_8. \tag{5.2.17}$$

Notice this nicely respects the $K$-theory invariance of the class $(T\Sigma', N\Sigma)$, as it only involves characteristic classes of the difference bundle.

The parity anomaly induces a Chern-Simons coupling $A_G P_{8+n}(R', F)$ in the supercritical type IIB theory. Upon T-duality on the $\mathbb{S}^1$, the gauge field $A_G$ becomes a component of the higher-dimensional NSNS 2-form $B_2$, so the above term arises from the KK reduction of a Chern-Simons coupling in the supercritical IIA theory of the form

$$B_2\, P_{8+n}(R, F) \tag{5.2.18}$$

Note that for $F$ trivial we recover the familiar $B_2 X_8$ term of 10d type IIA theory (5.2.1)

### 5.2.4 NS-brane worldvolume anomaly cancellation

The 10d type IIA CS coupling $B_2 X_8$ is a crucial ingredient in the cancellation of the NS5-brane worldvolume 6d anomalies by inflow, and explains that $X_8$ is actually the anomaly polynomial of this 6d theory [115]. In a similar spirit, we will show that the above CS coupling in the supercritical IIA theory precisely produces the inflow cancelling the worldvolume anomalies of the NS$(5+2k)$-brane (the object charged magnetically under $B_2$).

The chiral field content on the worldvolume of an NS$(5+2k)$-brane in supercritical type IIA theory can be guessed by the requirement that upon bulk closed tachyon condensation it becomes the standard NS5-brane worldvolume field content of 10d type IIA theory. Namely, this spectrum consists of two fermions of same chirality in 6 dimensions, plus a self-dual 2 form. The fermions should arise out of localized modes of supercritical fermions in the tachyon background, as briefly discussed in Section 5.1.1 and argued in [4]. As for the self-dual form, in the critical theory it arises as the reduction of the bulk anti self-dual 4-form along the anti self dual 2-form of the Taub-NUT background. One may expect the same logic to hold in the supercritical picture, so one expects to have a self-dual $(4+k)$-form in the supercritical theory, along with other lower $p$-forms. This sector will, upon tachyon condensation, yield the self-dual 2-form and one scalar.

In what follows we prefer to provide a more direct and rigorous derivation, from T-duality with a Taub-NUT configuration in supercritical IIB (which is easily understood to hold as in the critical case e.g. using Buscher's worldsheet derivation of T-duality). In fact, the set of localized modes at the center of Taub-NUT spaces can be obtained, for $k$ coincident centers, by replacing the near-center region of the Taub-NUT by an $A_{k-1}$ singularity, i.e. a $\mathbb{C}^2/\mathbb{Z}_k$ orbifold. Since we are interested in the chiral content, the anomaly is given by $n$ copies of that in a single Taub-NUT center (i.e. a single NS-brane in the type IIA theory).

For simplicity, we focus on the supercritical type IIB theory with a $\mathbb{C}^2/\mathbb{Z}_2$ along the critical slice. Since supercritical IIB is already a $\mathbb{Z}_2$ orbifold of supercritical 0B, we have to compute the spectrum in a $\mathbb{Z}_2\times\mathbb{Z}_2$ orbifold. As in Chapter 2, we denote by $g$ the generator of the $\mathbb{Z}_2$ to become the IIB GSO projection; we also use $h$ for the additional $\mathbb{Z}_2$ action $X\to -X$ on four of the 10d coordinates, and their worldsheet superpartner fermions. Let us now discuss the spectrum in sectors twisted by $h$ and $gh$, which lead to the localized matter:

- $h$-twisted sector: This contains NSNS and RR sectors, and chiral matter arises only from the latter. The zero point energies vanish, and there are fermion zero modes in all directions except those twisted by $h$. Massless states are the tensor product of two $SO(6+2k)$ spinors (with same relative chirality in the 0B theory). This produces a set of even-degree RR $p$ form potentials, including a chiral $(2+k)$-form, self-dual in the $(6+2k)$-dimensional fixed locus. These forms can be regarded as arising from the bulk RR forms supported on the harmonic 2-form associated to the collapsed 2-cycle, e.g. the self-dual $(2+k)$-form arises from integrating the bulk self-dual $(4+k)$-form over the 2-cycle.

Using the latter interpretation it is clear that the contribution to the anomaly polynomial is

$$X^{(d)}_{8+k} = -\frac{1}{8}\left[\frac{L(R)}{L(F)}\, e(F)\right]_{8+k} . \tag{5.2.19}$$

- $gh$-twisted sector: This contains NS-R and R-NS sectors, and behaves as a $g$-twisted sector with $k+4$ twisted dimensions instead of just $k$. The zero point energies vanish in both the NS and R sectors, so massless states arise from degenerate groundstates from the fermion zero modes.

  The spectrum consists of two $(6+2k)$-dimensional spinors of $SO(1,2k+5)$ of definite chirality, transforming as $(1/2,0)$ spinors of the $SO(4)$ normal to the fixed locus. The anomaly polynomial of the chiral matter of this sector is

$$X^{(e)}_{8+k} = 2\left[\frac{\hat{A}(R)}{\hat{A}(F)} e(F)\right]_{8+k} . \tag{5.2.20}$$

The total worldvolume anomaly polynomial of the $NS(5+2k)$-brane is

$$X^{NS}_{8+k} = X^{(d)}_{8+k} + X^{(e)}_{8+k} = e(F)\left[2\frac{\hat{A}(R)}{\hat{A}(F)} - \frac{1}{8}\frac{L(R)}{L(F)}\right]_{8} \tag{5.2.21}$$

We have the amazing fact that $X^{\rm NS}_{8+n} = P_{8+n}$, ensuring cancellation of anomalies by inflow from the Chern-Simons coupling (5.2.18). This is just the ordinary cancellation condition for the 8-form part of IIB anomaly polynomial, but the characteristic classes evaluated at the K-theory classes.

### 5.2.5 Other closed string tachyon solitons

In the previous subsections, we have described the endpoint of tachyon condensation for profiles carrying nontrivial charges in the integer-valued real K-theory groups $KO(\mathbb{S}^4)$ and $KO(\mathbb{S}^8)$. We have also provided evidence of the K-theory equivalence picture between the supercritical bundles upon tachyon condensation, for instance verifying that the anomaly polynomial of supercritical IIB or the $B_2X_{8+2k}$ coupling in supercritical IIA depend on the supercritical dimension bundles only through their K-theory class.

Keeping up with this picture, one should also be able to study solitons charged under the $\mathbb{Z}_2$-valued $K$-theory groups $KO(\mathbb{S}^k)$ for $k = 1,2,9,10$. Such objects are necessarily non-supersymmetric, and therefore much less understood than their integer-valued counterparts which, as we have seen, correspond to familiar objects in type II string theory. The great advantage of the K-theory description of the supercritical theory is that, once established, it allows us to describe these stable but exotic solitons on the same footing as the more well-known ones.

- Associated to $KO(\mathbb{S}^{10})$ we have a $\mathbb{Z}_2$ instanton. In the type II theory, this can be constructed as a gravitational instanton associated to the first factor of the homotopy group $\pi_9(SO) = \mathbb{Z}_2 \oplus \mathbb{Z}$ of the tangent bundle. The second factor is unstable, meaning that a representative of a nontrivial class of it can be made trivial by the addition of extra supercritical dimensions.

To our knowledge, this non-supersymmetric object has not been described before in the literature. It is a gravitational instanton which in principle should provide nonperturbative contributions to the 10-dimensional type II couplings, albeit to higher-derivative terms due to many fermionic zero modes due to the broken supersymmetries.

- For $KO(\mathbb{S}^9) = \mathbb{Z}_2$ we have a non-BPS particle in both type II theories. It is clearly different from the familiar ($\mathbb{Z}$-valued) type IIA D0-brane, or the (uncharged non-BPS) type IIB $\hat{D}$0-brane. Standard analysis as in [37, 119] suggest it transforms as a spinor (of both the supercritical bundles, and of the critical spacetime). As we argue later, the $\mathbb{Z}_2$ charge is fermion number mod 2, so it is possible that this particle simply decays into a perturbative spinor. If not, it would be interesting to follow this particle in dual pictures, in particular in the M-theory lift of type IIA.

- For $KO(\mathbb{S}^2)$, we get the dual string for the particle described above. Before the tachyon condensation it corresponds to a geometric configuration in which an $\mathbb{R}^2$ in the supercritical dimensions picks up a $2\pi$ rotation as one moves around the core of the soliton. This a non-trivial action on the spacetime fermions, which survives even after tachyon condensation. Therefore the string is associated to a $\mathbb{Z}_2$ discrete (gauge) symmetry under which only fermions are charged. It also acts non-trivially on the non-BPS particle described above, in nice agreement with our statement that it is a spinor (and thus a fermion).

- Finally, $KO(\mathbb{S}^1)$ corresponds to a domain wall. Before the condensation, it corresponds to a geometric configuration with one supercritical dimension fibered as in a Moebius strip. In analogy with the heterotic setting of Chapter 4 it seems reasonable to conjecture that amplitudes containing an instanton and a domain wall will get an extra phase when crossing each other.

## 5.3 A supercritical viewpoint on GLSMs

Non-trivial Calabi-Yau threefolds are often constructed as (complete intersections of) hypersurfaces in toric varieties. When these Calabi-Yau threefolds are taken to be backgrounds for type II string theory, the toric construction admits a beautiful realization in string theory [120]. Instead of trying to construct the CFT describing the non-linear type II sigma model in the target Calabi-Yau, we construct a particular non-conformal field theory in the worldsheet, the $(2, 2)$ Gauged Linear Sigma Model (GLSM). At low energies this field theory will flow to the non-linear sigma model, so we recover the previous physics, but the UV description makes the geometric construction as a hypersurface manifest. (Remarkably, the GLSM description does, in addition, allow for exploration of the aspects of string theory not captured by classical geometry, but this will not concern us here.)

A simple example is the quintic Calabi-Yau threefold, which can be constructed as a degree five hypersurface in the ambient toric space $\mathbb{P}^4$. In the GLSM description a slightly different description arises. We consider a $U(1)$ gauge group in our $(2,2)$ worldsheet theory, and take a set of six (complex) chiral superfields $z_i, P$ with charges

$$\begin{array}{c|cccccc} & z_1 & z_2 & z_3 & z_4 & z_5 & P \\ \hline U(1) & 1 & 1 & 1 & 1 & 1 & -5 \end{array} \tag{5.3.1}$$

In order to be in the geometric phase of the quintic, we also choose a value for the Fayet-Iliopoulos term for the $U(1)$ which forces $\sum |z_i|^2 > 0$, or in other words it disallows all the $z_i$ from vanishing simultaneously. We introduce, in addition, a superpotential of the form

$$W_{GLSM} = f_5(z_i)\, P \tag{5.3.2}$$

where $f_5(z_i)$ is a homogeneous polynomial of degree 5 in the $z_i$. We now look to the classical set of vacua for this theory. The set of fields in (5.3.1), plus the D-term constraint, parameterize a Calabi-Yau fivefold given by the $\mathcal{O}(-5)$ bundle over $\mathbb{P}^4$. The F-terms coming from (5.3.2) then impose $f_5(z_i) = 0$, and for smooth quintics (i.e. such that $df_5 = f_5 = 0$ has no solutions) also $P = 0$. Geometrically, it is most natural to take $P = 0$ first, which restricts us to the $\mathbb{P}^4$ base of the fibration, and then $f_5(z_i) = 0$ reproduces the classical description of the quintic.

This construction will be familiar to most readers, but given the focus of this Chapter we would like to reformulate it from the following supercritical type II viewpoint. We consider the GLSM in (5.3.1), initially with no superpotential, as describing a supercritical type 0 on the resulting $\mathcal{O}(-5) \to \mathbb{P}^4$ fivefold (therefore suitably dressed with a linear dilaton background). To connect with type II, we impose a $\mathbb{Z}_2$ quotient chosen to act as $P \to -P$, leaving the $z_i$ invariant. Notice that differently to previous cases, we make no distinction between "critical" and "supercritical" $z_i$ coordinates, all of them enter equally in the construction.[5]

Type 0 on the fivefold has a real closed string tachyon. We choose a profile for this tachyon given by

$$\mathcal{T} = \mathrm{Re}\,(W_{GLSM}) = \mathrm{Re}\,(f_5(z_i)\, P)\,, \tag{5.3.3}$$

ignoring as usual the lightlike profile. The solution of the equations of motion at late times is the subvariety cut out by

$$\frac{\partial \mathcal{T}}{\partial x_k} = \frac{\partial \mathcal{T}}{\partial y_k} = \frac{\partial \mathcal{T}}{\partial p} = \frac{\partial \mathcal{T}}{\partial q} = 0 \tag{5.3.4}$$

[5]Up to the same point in the argument, morally the same was true in our previous cases, since the distinction between $x$ and $X$ coordinates only mattered once we chose a tachyon profile and a branch of the resulting moduli space. In the current case, due to the non-trivial topology, it makes no sense to partition the $z_i$ coordinates, even in principle.

where we have introduced $z_k = x_k + iy_k$, $P = p + iq$. (Some subtler details of the dynamics of this system were worked out in [5].) Since $W_{GLSM}$ is holomorphic in $z_i, P$, one easily sees that (5.3.4) is equivalent to[6]

$$\frac{\partial W_{GLSM}}{\partial z_k} = \frac{\partial W_{GLSM}}{\partial P} = 0\,. \tag{5.3.5}$$

In words, we perfectly reproduce the GLSM construction, and end up at late times with ten-dimensional type II string theory on $P = f_5(z_i) = 0$ (for smooth $f$; see later for comments on singular cases).

Besides providing an interesting alternative viewpoint on the usual GLSM construction, the supercritical viewpoint has an interesting application to the study of discrete symmetries in string theory. The idea is a straightforward generalization of the discussion in Chapter 3: before the tachyon starts condensing, we have a rather large continuous symmetry group acting on $\mathcal{O}(-5) \to \mathbb{P}^4$ (in particular acting on the $\mathbb{P}^4$). The choice of the tachyon (5.3.3) breaks down this continuous group into a discrete subgroup for appropriate choices of $f_5(z_i)$. For instance, take the Fermat quintic

$$f_5(z_i) = z_1^5 + z_2^5 + z_3^5 + z_4^5 + z_5^5 - \psi z_1 z_2 z_3 z_4 z_5 \tag{5.3.6}$$

where $\psi$ is a complex structure parameter, which we take to be nonvanishing. In this case we have a non-abelian symmetry group given by $(S_5 \ltimes (\mathbb{Z}_5)^4)/\mathbb{Z}_5$ with $S_5$ the permutation group of 5 elements (see for example [121] for a detailed discussion of this point, and generalizations). This groups survives in the final type II background, and has now a natural embedding into the isometry group of the original non-compact Calabi-Yau fivefold.

The embedding of the discrete symmetries into a continuous group allows for a direct construction of the 4d charged strings (see [13] for a related but different approach). For a given $\mathbb{Z}_p$ string, one simply constructs a vortex of the Kibble mode associated to the corresponding $U(1)$ symmetry in the continuous group. It would be interesting to explore the string creation effects associated to mutually non-commuting strings.

It is also interesting to consider what happens to the background when the discrete symmetry gets enhanced to a continuous one, i.e. when we approach the core of the vortex. Choose for instance a continuous deformation of the Fermat quintic taking us to the singular space $X_s$ given by $f_5(z_i) = \psi z_1 z_2 z_3 z_4 z_5$. In this limit the discrete symmetry is enhanced to a continuous group containing $(S_5 \ltimes U(1)^5)/U(1)$. Topologically $X_s$ is a set of 5 intersecting $\mathbb{P}^3$s given by $z_i = 0$. This manifold is singular, and in fact it is an example of the ground state varieties with non-transversal polynomials described in [122]. Classically, in addition to the non-linear

[6]Here is an argument: write $W_{GLSM} = u + iv$. Since it is holomorphic on $z_i, P$, one has the Cauchy-Riemann conditions $\frac{\partial u}{\partial x_k} = \frac{\partial v}{\partial y_k}$, $\frac{\partial u}{\partial y_k} = -\frac{\partial v}{\partial x_k}$, and similarly for $P$. The supercritical equations of motion (5.3.4) are just $\frac{\partial u}{\partial x_k} = \frac{\partial u}{\partial y_k} = \frac{\partial u}{\partial p} = \frac{\partial u}{\partial q} = 0$. These equations together imply (5.3.5), if we expand it in real components. The reverse implication is also straightforward: using the real decomposition of $\frac{\partial W_{GLSM}}{\partial z_k} = 0$, together with holomorphy, one gets $\frac{\partial u}{\partial x_k} = \frac{\partial u}{\partial y_k} = 0$.

sigma-model branch, the $P \neq 0$ branch of the CFT opens up at the singularities, i.e. whenever two $\mathbb{P}^3$ components intersect. In related contexts [103, 123] there is evidence that the appearance of these branch structure signals the breakdown of worldsheet CFT, and that we need to include the effects of light brane states. This is also the case here. In the original quintic we had 204 independent homology three-cycles. These all become localized at the singularities of $X_s$, since $\mathbb{P}^3$ does not support any non-trivial three-cycle. In the type II theory the branes wrapping these cycles will become massless. Notice that differently from the usual conifold transition, here both electric cycles and their magnetic duals are becoming massless, so we expect string theory in this background to be described in four dimensions by a theory without a Lagrangian description.

The reinterpretation of the GLSM generalizes easily to spaces beyond the quintic. As a simple example, let us consider the non-compact Calabi-Yau sixfold given by the GLSM

$$\begin{array}{c|cccccccc} & z_1 & z_2 & z_3 & z_4 & x & y & z & P \\ \hline U(1)_1 & 1 & 1 & 1 & 1 & 8 & 12 & 0 & -24 \\ U(1)_2 & 0 & 0 & 0 & 0 & 2 & 3 & 1 & -6 \end{array} \tag{5.3.7}$$

Choose as above $\mathcal{T} = \mathrm{Re}\,(W_{GLSM})$ with

$$W_{GLSM} = P\Big(y^2 + x^3 + f(z_i)xz^4 + g(z_i)z^6\Big) \tag{5.3.8}$$

which after tachyon condensation, for a suitable choice of FI terms, leaves us with an elliptically fibered fourfold, with base $\mathbb{P}^3$. For homogeneity we need that $f(z_i)$ and $g(z_i)$ are homogeneous polynomials of $z_i$ of degrees 16 and 24 respectively. If we want to engineer a background with non-abelian discrete symmetries, we could choose for instance, in analogy with our discussion for the quintic

$$f(z_i) = \sum_{i=1}^{4} z_i^{16} + \psi_1(z_1z_2z_3z_4)^4 \tag{5.3.9}$$

$$g(z_i) = \sum_{i=1}^{4} z_i^{24} + \psi_2(z_1z_2z_3z_4)^6 \tag{5.3.10}$$

which are invariant under a manifest symmetry group $S_4 \ltimes \mathbb{Z}_2^4$ acting on the $\mathbb{P}^3$ coordinates, leaving $(x, y, z, P)$ invariant. Richer configurations can appear if we can induce discrete symmetries acting non-trivially on the fiber, which becomes particularly easy if we choose different realizations of the torus fiber [124, 125], such as a cubic in $\mathbb{P}^2$, or a quartic on $\mathbb{P}^{112}$. This construction has potential applications to the realization of non-abelian symmetries in F-theory.[7]

Another interesting observation in this context is that closed string tachyon condensation on flat space can give an alternative supercritical construction of ADE

[7]See [126, 127, 128, 129, 130, 131, 132, 133] for recent work on abelian discrete symmetries in F-theory, and [134] for work on non-abelian discrete symmetries in F-theory.

singularities for appropriate choices of the tachyon profile. For instance, start with $\mathbb{C}^4$ parameterized by $(z_1, z_2, z_3, P)$, with the $\mathbb{Z}_2$ orbifold acting as $(z_i, P) \to (z_i, -P)$. We choose a tachyon profile $\mathcal{T} = \text{Re}\left((z_1^2 + z_2^2 + z_3^2)P\right)$. After the tachyon condenses we end up with $z_1^2 + z_2^2 + z_3^2 = 0 \subset \mathbb{C}^3$, a well known description of $\mathbb{C}^2/\mathbb{Z}_2$.

## 5.4 Supercritical bundles for F-theory matrix factorizations

In [135, 136], type IIB tachyons describing particular backgrounds were related to particular matrix factorizations (MF) of polynomials living on an ambient auxiliary space sharing some resemblance to the total space of F-theory fibrations. The matrices of the matrix factorization are sections of bundles defined over the ambient space, much like our tachyon matrix $\mathscr{T}$ above is a section of the $x, y$ bundles over the critical slice. It is therefore tempting to try to give a physical meaning to this ambient space, and the associated matrix factorizations, by somehow reinterpreting it as a supercritical type 0 background, with tachyon condensation reducing again the dynamics to the type 0 critical slice (the ambient space).

We will first briefly review the recipe of [135] in the type F-theory context. An F-theory geometry is commonly specified by a possibly singular Calabi-Yau fourfold $\mathbb{X}_4$ with a torus fibration. This fourfold is part of an M-theory background $\mathbb{X}_4 \times \mathbb{R}^{2,1}$ which is related to the four-dimensional F-theory model by taking a limit in which the size of the torus fiber shrinks to 0. Enhanced gauge groups and chiral matter may arise at singularities of the fiber. As discussed above, a convenient class of backgrounds $\mathbb{X}_4$ can be constructed in terms of a hypersurface $f = 0$ on an ambient toric space $\mathcal{A}$.

The polynomial $f$ specifies the geometry of the fibration. However, this does not fully specify the M/F-theory background; the profile of the M-theory 3-form $C_3$ must also be specified. For backgrounds without $G$-flux the choice of $C_3$ is equivalent to picking a particular element in the Deligne cohomology of $H_3(X_4, \mathbb{Z})$ [137]. For singular fourfolds, there may be vanishing cycles on which one can wrap $C_3$; these backgrounds typically correspond to T-branes in the type IIB setup [137].

The proposal of [135] is that these "T-brany" degrees of freedom of the singular fourfold are captured by the different matrix factorizations of the polynomial $P$. A $n \times n$ MF of a polynomial $P$ is a pair $(A, B)$ of matrices satisfying

$$A \cdot B = B \cdot A = f \cdot \mathbf{I}_{n\times n}. \tag{5.4.1}$$

A particular (stable) matrix factorization gives a particular T-brane background. The reverse is not quite true: a T-brane background modulo complexified gauge transformations corresponds to a particular equivalence class of matrix factorizations (which furnish the so-called category of stable matrix factorizations).

The MF is naturally associated to a complex of vector bundles over the ambient space $\mathcal{A}$, which is formally very similar to those appearing in open tachyon condensation in brane-antibrane systems. In the present setup, however, there are no open string sectors. This seems to demand a setup in which there is there are bundles associated to geometry, and a mechanism to annihilate them. It is very tantalizing to suggest that the tachyon condensation of bundles in the complex defined by the MF is physically realized in terms of supercritical strings. In the following we give some suggestive hints in this direction.

A first step is to connect F-theory to the supercritical string setups considered here. This can be done by defining F-theory from M-theory on $\mathbb{X}$ as described above, and further compactifying on an extra $\mathbb{S}^1$ to connect with type IIA on $\mathbb{X}$. Conversely, we start with type IIA on an elliptically fibered variety $\mathbb{X}$, lift to M-theory on $\mathbb{X} \times \mathbb{S}^1$, decompactify the $\mathbb{S}^1$, and subsequently[8] shrink the elliptic fiber to connect to F-theory on $\mathbb{X}$. As usual, holomorphic information is preserved under these operations, so the physics of F-theory at singularities of $\mathbb{X}$ will be directly described by the physics of type IIA at singularities of $\mathbb{X}$.

We can describe these type IIA compactifications in terms of a $(2,2)$ GLSM with ambient space $\mathcal{A}$ and superpotential determined by the defining polynomial $f$. In this context, there is a natural proposal to realize the inclusion of the degrees of freedom of the MF. A typical trick to introduce extra bundles is to introduce extra $(2,2)$ multiplets in the GLSM, specifically $n$ pairs of multiplets $X_a$, $Y_a$, with $a = 1, \ldots, n$ (possibly charged under the 2d gauge group). Alternatively, these can be physically understood as extra supercritical dimensions[9] in a supercritical type 0 theory (or a supercritical type II theory, by simply orbifolding by $Y \to -Y$). In this language, the interpretation of the complex of bundles associated to the MF as a tachyon condensation motivates the introduction of closed tachyon background in the supercritical theory given by

$$\mathcal{T} = \mathrm{Re}\,\big(\sum_{ab} X_a M_{ab} Y_b\,\big) \tag{5.4.2}$$

Namely, the matrix $M$ in the MF plays as the tachyon matrix $\mathscr{T}$. This tachyon background leads to the annihilation of the extra supercritical dimensions, except at special loci at which the rank of the matrix drops, which occur precisely at singular points of $\mathbb{X}$. From the viewpoint of the supercritical worldsheet description, there are extra dimensions left over at these loci, which describe singular CFTs as we have described in earlier sections. A careful study of such singular CFTs should reproduce the formal prescription of looking to matrix factorizations (5.4.1). (Not entirely unlike how type II CFT reproduces the classification of D-branes by matrix factorizations [138], although we expect the details to differ significantly).

Some interesting hints from translating results of the previous sections to this setup are that complex codimension 2 defects are associated to the singular CFT of $A_n$ singularities, and therefore are naturally associated to the appearance of enhanced gauge symmetries. Also, complex codimension 4 defects are associated to

[8]We do not consider other limiting procedures, which may lead to other possibly interesting outcomes.

[9]This is in addition to possibly including the normal dimension to $\mathbb{X}$ in $\mathcal{A}$ as supercritical, as in Section 5.3.

the presence of F1's in the type IIA picture, which maps under the dualities to the presence of D3-branes in the F-theory setup; this nicely connects with the presence of (anti) D3-brane sources detected in [135, 136] by the presence of matter localized at points.

Clearly, there are many points that deserve further clarification. For instance, the fact that F-theory MFs are defined by the stable category requires that widely different tachyon backgrounds lead to equivalent physics, at least in what concerns the singular structure of the model. Also, we have been deliberately ambiguous about whether the supercritical string theory is defined on $\mathbb{X}$ or on $\mathcal{A}$ (times the extra 'bundle' dimensions). Finally, it is not clear what computations in the (eventually singular) CFT could match the remarkably tractable computation of non-perturbative spectra in [135, 136] in terms of exact sequences.

# Part II

# The Weak Gravity Conjecture

# 6 The Weak Gravity Conjecture

In the second part of this thesis, we will focus on a very different tool for obtaining information about interesting charged states in string theory: The Weak Gravity Conjecture. This conjecture, first proposed in [139], has turned out to be a very powerful tool for constraining phenomenological models. Its power relies on its generality: whereas the constructions of the first part of the thesis construct charged states explicitly, the WGC demands the existence of certain charged states in any theory. Arguably, then if a theory does not include such states its in trouble with quantum gravity, likely belonging to the Swampland mentioned in Section 3.1.

This Chapter reviews the essentials of the WGC which will be used in Chapters 7 and 8. We will start with a motivation and introduction to the WGC, and then proceed to survey the different versions of the conjecture in different scenarios. Finally we will summarize in a convenient way the relationships between the different versions of the conjecture.

## 6.1 Rationale for the WGC & the electric form

The WGC is closely related to the considerations about global symmetries and quantum gravity discussed in Chapter 3. There, it was argued that under mild assumptions one cannot have a global symmetry in a quantum theory of gravity, since that would lead to infinitely many black hole remnants and a pathological theory.

Then, in any theory with a $U(1)$ symmetry this symmetry must be gauged (as described in Chapter 3, this means that the symmetry is the global part of a $U(1)$ gauge group). For concreteness, let us focus on the four-dimensional case. A $U(1)$ gauge symmetry is described by a dimensionless gauge coupling $g$. The coupling enters into the covariant derivative, so that the lagrangian for a field of charge $q$ is built using

$$D_\mu \phi_q \equiv (\partial_\mu - iqgA_\mu)\phi_q. \tag{6.1.1}$$

The global part of the gauge group acts on the field $\phi_q$ as $\phi \to \exp(2\pi i q\alpha)\phi_q$.

Now, consider tuning the gauge coupling all the way to 0. Then the gauge field decouples from the rest of the system, but we retain the symmetry acting on charged fields as multiplication by a phase $\exp(2\pi i q\alpha)$. This is a forbidden $U(1)$ gauge symmetry, so it follows that it is not consistent to tune the gauge coupling to zero in a quantum theory of gravity.

This argument in fact works for non-abelian gauge fields in any number of dimensions, and it is easy to verify in many examples in string theory. For instance, consider heterotic string theory, which has a non-abelian gauge field in 10 dimensions with coupling [10, 12]

$$g^2 = \frac{4}{\alpha'}\kappa^2, \tag{6.1.2}$$

where $\kappa$ is the square of the 10-dimensional Newton's constant. We see that taking $g = 0$ amounts to either $\kappa = 0$, so that gravity is decoupled and there are no black holes to give us trouble, or $\alpha' = 0$ so that the strings are tensionless and we do not have a local effective field theory description. In other words, one cannot take $g = 0$ without decoupling gravity. This is just a particular example of a generic behavior in string theory.

The WGC stems from exploring what happens in the $g \to 0$ limit without actually reaching this value. For a global symmetry, we have the problem that we can build black holes with arbitrarily high values of the charge under the global symmetry, and they are mass-degenerate. For a gauge symmetry we can also build charged black holes, described by the famous Reissner-Nordstrom solution [140, 52]

$$ds^2 = -f_{RN}(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2, \quad f_{RN}(r) \equiv 1 - \frac{2M}{M_P^2 r} - \left(\frac{gQ}{M_P r}\right)^2. \tag{6.1.3}$$

This metric presents two coordinate horizons at the roots of $f_{RN}$. The causal structure of the solution is controlled by the discriminant of $f_{RN}(r) = 0$, which turns out to be (in terms of the non-reduced Planck mass $M_P$)

$$\Delta = \frac{4}{M_P^2}\left[M^2 - (gQM_P)^2\right]. \tag{6.1.4}$$

For $\Delta > 0$, spacetime is globally hyperbolic. There are two coordinate horizons; the exterior one is an usual event horizon, while the interior is a Cauchy horizon which will be "covered" by matter in any charged black hole produced via spherical collapse [140]. We will not speculate about the physics behind the horizon. These black holes are called subextremal.

At $\Delta = 0$, the two horizons coincide; the black hole is said to be extremal. An extremal black hole has a number of interesting properties. For instance, its Hawking temperature is zero, so that it does not radiate; it is a stable object. Its entropy computed by the usual Bekenstein-Hawking formula vanishes. In string theory, extremal solutions are often the supergravity limit of interesting nonperturbative objects such as D-branes.

For $\Delta < 0$, the black hole is superextremal. There is no event horizon; instead, a naked singularity is visible to asymptotic observers. This spacetime is highly pathological, having timelike curves and the associated loss of global hyperbolicity. This is not to say that objects resembling superextremal black holes do not exist, but rather that they cannot be described by the weakly coupled semiclassical description we have been using. The Cosmic Censorship hypothesis (see [140, 141] among many others for a review) states that no naked singularities can form dinamically within the confines of general relativity coupled to classical fields.

Given a charged black hole of mass $M$ and charge $Q$, it will lose both charge and mass via Hawking evaporation [142, 44, 143, 144], in such a way that it will only stop radiating when approaching the extremal limit. In a theory with a charged particle it is easy to build a very massive subextremal charged black hole with whatever charge we want; then, it will Hawking-evaporate all the way to the extremality bound.

We thus arrive at the conclusion that, in a theory where Hawking's calculation can be trusted for large enough black holes, we will have extremal black holes of mass $M = gM_PQ$ for every value of $Q$ or, at least, for every $Q$ large enough so that the semiclassical approximation and the description in terms of the metric (6.1.4) is valid.

Now, imagine tuning the coupling $g$ close to zero, while keeping $M_P$ constant. The number of extremal black holes with masses in the range $[M, M + \Delta M]$ is

$$n(g) = \frac{\Delta M}{gM_P}, \tag{6.1.5}$$

which goes to infinity as $g \to 0$. If these extremal black holes are exactly stable (we have seen that they seem to be at the semiclassical level), as $g \to 0$ they become exactly the kind of remnants that posed trouble for global symmetries in Chapter 3. Namely, the thermal bath at the Unruh temperature will see these objects suppressed by a factor $n(g)\exp(-M/T)$. For $g$ sufficiently small their contribution will overrun the partition function, leading to a theory out of control.

A simple way out of this trouble is to demand the existence of a decay channel for extremal black holes, so that they are not stable. We expect unstable states to only contribute significantly to a thermal bath for temperatures larger than their decay width $\Gamma$(it is only then that the particle can exist in the thermal bath for a sufficiently long time). Semiclassically, an extremal black hole is exactly stable, but quantum-mechanically it may have a finite lifetime. These lifetimes can conspire to make the overall contribution of the extremal black holes to the thermal bath negligible.

This is far from actually showing that extremal black holes do not become remnants. For this, one should actually compute $\Gamma$, the decay width of each black hole, and show that in the IR they decouple from the physics. Nonetheless, demanding the existence of a decay channel turns out to be a very powerful tool for analyzing theories with gauge symmetry coupled to gravity.

Let us work out the constraints on the particle which allows the black holes to decay. For this, it is useful to define the charge-to-mass ratio $\xi$ of an object as

$\xi = (QgM_P/)M$, so that extremal black holes have $\xi = 1$. An extremal black hole of mass $M$ and charge $Q$ decays via emission of a particle of mass $m$ and charge $q$. We require the final black hole to also be subextremal, so that we remain in control of the effective field theory. We then have

$$gM_P(Q-q) \geq M - m \quad \Rightarrow \quad m \leq qgM_P \tag{6.1.6}$$

where we have used the fact that the original black hole was exactly extremal, $M = QgM_P$. This is the Weak Gravity Conjecture: in a $U(1)$ theory in four dimensions, there should be a charged superextremal particle (extending the nomenclature from charged black holes) in the theory. Although the argument has been made in four dimensions, nothing is particular to this case: In $D$ dimensions, the gauge coupling has dimension $g^{\frac{D}{2}-2}$, and the WGC demands the existence of a particle with mass lower than $g/\sqrt{G}$, where $G$ is Newton's constant in $D$ dimensions.

We emphasize that the WGC is a kinematical requirement, meant to allow the charged black holes to decay in principle via the emission of a charged particle. Whether or not they actually do this is a much more complicated question.

There is ample evidence for this conjecture in various string theory setups. For instance, in type IIA string theory, there is a $U(1)$ gauge field in 10d. The WGC conjecture is saturated by the $D0$ brane, which indeed is charged under this $U(1)$. Satisfaction of the inequality is not so surprising in theories with extended supersymmetry, since in this context the extremality condition $M = QgM_P$ is precisely the BPS bound. Thus, superextremal objects are forbidden in these theories, and it is enough that they contain BPS objects to satisfy the WGC. This is not a trivial requirement; from a field theory point of view it is perfectly fine to have a supersymmetric theory without BPS multiplets.

There is one interesting nonsupersymmetric example, introduced in the original WGC paper [139], in which the WGC is satisfied by superextremal states. Consider again heterotic string theory in 10 dimensions, and let us focus on the $U(1)^{16}$ Cartan torus. This theory has only $\mathcal{N} = 1$ in ten dimensions, so it does not have an exact BPS bound.

The excited levels of the fundamental string are charged under the Cartan $U(1)$. The left-moving part of a heterotic worldsheet contains a current algebra, which gives rise to the spacetime gauge group. The allowed string states have their charges living on an even, self-dual lattice; we will take the $SO(32)$ example in what follows. The contribution of the current algebra to the Virasoro generators is given by the Sugawara construction [10, 12], so that

$$\frac{\alpha'}{2}M_L^2 = \frac{1}{2}|\vec{Q}_L|^2 - 1 + N_L. \tag{6.1.7}$$

Here, $N_L$ is the oscillator number of the higher harmonics of the string. Let us focus along a single $U(1)$ out of the 16 present in the Cartan algebra, such as the first one, to keep in line with the discussion of previous paragraphs. In this case, the massive states of the string give a family of charged states, with charge vector of the form

$(Q, 0, \ldots)$ and $Q$ even. They have a charge-to-mass ratio ($M^2 = 2M_L^2$ due to level matching)

$$\xi^2 = \frac{Q^2 g_{U(1)}^2}{\kappa^2 M^2} = \frac{Q^2}{Q^2 - 2}, \tag{6.1.8}$$

where the $U(1)$ gauge coupling $g_{U(1)}^2 = g^2/2$, where $g$ was given in (6.1.2). The extra factor of two comes from the canonical normalization $\mathrm{Tr}(T^a T^b) = 2\delta^{ab}$ of the generators of $SO(32)$. The charge-to-mass ratio of these perturbative heterotic states is bigger than one, and so the state is superextremal. These arguments depend only on the fact that the heterotic sigma model has a nontrivial current algebra, which is true for any heterotic model. This can be used to prove the Weak Gravity Conjecture for any perturbative heterotic compactification.

The above arguments are sketchy and at best constitute a mere motivation for the WGC. They have serious loopholes. On one hand it has been noted that they rely on the validity of the semiclassical description of black hole evaporation all the way to the extremal limit. On the other hand, to make the WGC argument we also need to take the $g \to 0$ limit. But maybe this is not a possibility in a quantum theory of gravity. In that case there is no justification for WGC particles coming from generic black hole decay and remnant considerations. However this does not seem to be the case in string theory, at least in generic situations. It is easy to build supersymmetric configurations in which the gauge coupling is a modulus and we can get arbitrarily close to $g = 0$ in moduli space, while keeping Newton's constant fixed. For instance, consider again heterotic string theory in 10 dimensions. This is a $\mathcal{N} = 1$ SYM plus gravity, with coupling $g$ satisfying (6.1.2). It is possible to take $g \to 0$ while keeping $M_P$ constant at the expense of lowering the string scale $(\alpha')^{-1/2}$, which acts as a cutoff of the EFT theory (incidentally, this is the cutoff predicted by the magnetic form of the WGC, as we will discuss in the next section). However, since the Schwarzschild radius grows with the mass, the effective field theory (valid over distances greater than $\sqrt{\alpha'}$) contains large near-extremal black holes of charge $Q \gtrsim (\kappa^{-1}/\sqrt{\alpha'})$ and mass $M \gtrsim (gQ\kappa^{-1}) = 2Q/\sqrt{\alpha'}$, to which the above arguments apply.

The version of the WGC discussed in this Section is the mild, electric version of the conjecture. It is mild because there are stronger versions of the conjecture, not motivated by remnant considerations, as we are about to discuss. It is electric because it predicts the presence of light objects coupled electrically to the gauge field; we will discuss now the magnetic version.

## 6.2 The magnetic WGC

A $U(1)$ gauge theory can couple to both electric and magnetic sources. Reissner-Nordstrom solutions with magnetic charge do exist [52], so we can apply the same arguments as above to magnetic charges. In this case, the WGC demands the existence of a monopole with mass $M \lesssim M_P/g$. However, this is now true in the strong coupling $g \to \infty$ limit. It is not an interesting statement for the weakly coupled theory (although it is for the magnetic dual).

What is usually called the magnetic WGC comes instead from a related but different line of reasoning. In a weakly coupled theory, monopoles are described as solitonic solutions to the equations of motion. Since the monopole charge enclosed by a surface $\Sigma$ is measured by the integral

$$\int_\Sigma F, \tag{6.2.1}$$

we can specify that this integral is nonvanishing at infinity, and integrate the field equations all the way to $r = 0$ to obtain a solution describing a monopole. For a weakly coupled $U(1)$ gauge theory, with lagrangian $\frac{1}{2}|F|^2$, the magnetic field goes as $1/r^2$, so that the energy density diverges at the origin. If the field theory has a cutoff at some scale $\Lambda$, then at a distance of roughly $\Lambda^{-1}$ from the core the UV completion takes over, smoothing out the description. A prototypical example of this is the $SU(2)$ 't Hooft-Polyakov monopole [145, 146]. It describes a particular configuration of nonabelian $SU(2)$ gauge theory, Higgsed down to a $U(1)$. At scales below the Higgs vev, the solution describes an abelian magnetic monopole. As the core of the monopole is reached the symmetry is restored, and the solution is everywhere smooth.

In these circumstances, it is possible to estimate the mass of the monopole as the sum of two contributions,

$$M_{\text{mon}} = M_{\text{UV}} + M_{\text{outer}}, \quad M_{\text{outer}} \sim \left(\frac{n}{g}\right)^2 \int_{\Lambda^{-1}}^{\infty} \frac{dr}{r^2} \sim \frac{n^2\Lambda}{g^2}. \tag{6.2.2}$$

Here, we have split the mass of the monopole into an UV contribution coming from the core region, and another coming from the outer layers in which our effective theory is valid. $n$ is the monopole charge; by Dirac quantization it is an integer, if we normalize the dual gauge coupling to $2\pi/g$.

Now, as mentioned above, we do have Reissner-Nordstrom black holes with magnetic charge, with an extremality condition $M \gtrsim 2\pi M_P/g$. The semiclassical description is impervious to Dirac quantization, which is a quantum effect. As a result, we expect these magnetic RN black hole solutions to be valid only in the regime of large magnetic charge. Put another way, the $n = 1$ monopole should not be black, meaning that its Schwarzschild radius is lower than the cutoff length scale $\Lambda^{-1}$. Using (6.2.2) and assuming $M_{\text{UV}} \geq 0$, we arrive at the so-called magnetic Weak Gravity Conjecture,

$$\Lambda \lesssim g M_p. \tag{6.2.3}$$

Thus, the effective field theory must have a cutoff lower than the mass of the electric particle. Like the electric form, this also extends to higher dimensions.

Equation (6.1.2) is precisely the magnetic WGC for the 10d heterotic theory; the mass of the electrically charged particle is of order the string scale $1/\sqrt{\alpha'}$, precisely when the supergravity effective description with action quadratic in the fields starts to fail. Similar considerations can be made for the other examples.

Like its electric counterpart, one can make objections to the reasoning which led to (6.2.3). Notably, the existence of a low cutoff (6.2.3) does not always mean that we lose control of the theory. The argument above merely tells us that by the time we reach the scale $gM_P$, the theory is no longer the quadratic action $\frac{1}{2}|F|^2$ that we took. But in some circumstances this is just fine; for instance, in D-brane setups the Dirac-Born-Infeld action takes into account all the $\alpha'$ corrections in the open-string sector in the limit of low acceleration. In fact, the argument leading to (6.2.3) is very dependent on canonical kinetic terms. For isntance, take instead nonlinear DBI electrodynamics, with lagrangian

$$\mathcal{L} \equiv -T \int (\sqrt{-\det(\eta + \sigma F)} - 1), \quad \sigma \equiv 2\pi\alpha'. \tag{6.2.4}$$

We will now compute the mass and field of a DBI monopole solution, in order to try to reproduce the argument which led to (6.2.3). The determinant in (6.2.4) can be evaluated explicitly in 4d (e.g. treating it as a characteristic polynomial in $\sigma$) to yield

$$T \int (1 - \sqrt{1 - 2|\sigma F|^2 - |(\sigma F)^2|^2}). \tag{6.2.5}$$

For a monopole, the equations of motion give

$$Td \left( \frac{2\sigma^2 F}{\sqrt{1 - 2|\sigma F|^2 - |(\sigma F)^2|^2}} \right) = \delta^{(3)}(r) \tag{6.2.6}$$

which can be integrated to

$$T \left( \frac{2\sigma^2 F}{\sqrt{1 - 2|\sigma F|^2 - |(\sigma F)^2|^2}} \right) = d\Omega. \tag{6.2.7}$$

From this, we get $F \wedge F = 0$ and therefore

$$2(\sigma F)^2 = \frac{1}{1 + 2T^2\sigma^2 r^4}. \tag{6.2.8}$$

The field is everywhere finite, and the mass of the monopole is

$$M = -4\pi T \int_0^\infty \left( \frac{1}{\sqrt{1 + (2T^2\sigma^2 r^4)^{-1}}} - 1 \right) r^2 dr \approx 4.6T(\sigma T)^{-3/2}. \tag{6.2.9}$$

Unlike in the previous case, the mass does not have an UV divergence. The fact that DBI electrodynamics makes the self-energy of the electron (and also that of the monopole) finite is in fact the reason it was first considered by Born and Infeld. As a result, the mass of the elementary monopole and hence its Schwarzschild radius are not sensitive to any further cutoff that the theory might have. So there is no magnetic WGC for DBI electrodynamics. (6.2.3) does hold with $\Lambda \sim 1/\sqrt{\alpha'}$, since from the point of view of the quadratic theory we have a cutoff at a scale $\Lambda^{-2} = 2\pi\alpha'$.

But now the theory is under control and there is no reason for any new physics at a scale $gM_P$, apart from the presence of the charged particle predicted by the electric version of the WGC, which in principle poses no trouble.

Other loopholes are the assumption $M_{\text{UV}} \geq 0$, which does not always hold in string theory (for instance, orientifold planes source RR fields but their overall tension is negative), or the fact that in the end there is no evident logical inconsistency with having the $n = 1$ monopole being an extremal magnetic RN black hole. In spite of these considerations, the magnetic WGC holds in many examples in string theory.

## 6.3 Extension to $p$-forms

Both the electric and the magnetic forms of the WGC studied above are valid for $U(1)$ gauge fields in an arbitrary number of dimensions. However, string theory contains a wider class of generalized gauge fields, of which the $B$-field and the Ramond-Ramond forms are examples. A $p$-form gauge field $C_p$ can couple to an extended object of worldvolume $V$ as

$$g \int_V C_p, \tag{6.3.1}$$

generalizing the coupling of an electric particle to a 1-form gauge field familiar from electrodynamics. $g$ is the gauge coupling which, for a canonically normalized $p$-form potential in $D$ dimensions, has dimension $\frac{D-2}{2} - p$ so that (6.3.1) is dimensionless.

In many cases, one can write down charged black $p$-brane solutions, analogous to the RN metric (6.1.4). These black branes evaporate via Hawking radiation much like their lower-dimensional counterparts, in such a way that the argument in Section 6.1 still works, albeit replacing mass with tension of the brane. Thus, the WGC can be extended to some $p$-form fields to demand the existence of a charged object with superextremal tension.

We now summarize, following [147], for which $p$-forms is this argument possible and what does subextremal exactly mean in this context. The action in the Einstein frame will be taken to be

$$S = \int \frac{R}{2\kappa^2} + \frac{1}{2}|F|^2. \tag{6.3.2}$$

As explained in [147] one can obtain the solution for general $p$ and $D$ from dimensional reduction of the Horowitz-Strominger solution in [148] (for more literature on dilatonic black branes, see [149, 150, 151, 152, 153, 154, 155]). We will construct a membrane with magnetic charge $n$; the electric one will follow by considering the dual gauge potential. A magnetic source for a $p$-form potential in $D$ dimensions has a worldvolume of dimension $D-p-2$. The black brane solution is obtained by splitting the spatial coordinates into $y^i$, $i = 1, \dots D-p-3$, parallel to the worldvolume of the brane, and using generalized spherical coordinates for the $p+2$ transverse spatial directions. The metric is

$$ds^2 = -f_{BB+}(r)f_{BB-}(r)^{\frac{2p}{D-2}\gamma} + f_{BB+}^{-1}(r)f_{BB-}^{-1}(r)dr^2 + r^2 d\Omega_{p+1}^2 + \delta_{ij}dy^i dy^j, \quad (6.3.3)$$

$$\gamma \equiv \frac{D-2}{p(D-p-2)}. \quad (6.3.4)$$

Here, $r_\pm$ are directly the location of the horizons, which can be determined in terms of the asymtotic charge and mass. These can be determined in a standard way [147], and after combining all the results we obtain an extremality condition

$$\gamma \frac{4\pi^2}{g^2} n^2 = \kappa^2 T_p^2. \quad (6.3.5)$$

By electric-magnetic duality we also obtain the extremality condition for electric branes with electric charge $Q$,

$$\gamma g^2 Q^2 = \kappa^2 T_p^2. \quad (6.3.6)$$

From these, one can write down the WGC for $p$-form fields,

$$T_p \leq \sqrt{\frac{D-2}{p(D-p-2)}} \frac{gQ}{\kappa}. \quad (6.3.7)$$

The formula has the same parametric dependence than the WGC for $U(1)$ gauge fields, as expected. This is not suprising since one can obtain ordinary gauge fields from dimensional reduction of higher gauge fields; in that case the formula (6.3.7) must reduce to (6.1.6). Also, in the context of string theory, different $p$-form fields are related by $T$-duality; the WGC formula has to be covariant under this operation if it is to hold in string theory.

We can also mention the higher form version of the magnetic WGC discussed in Section 6.2: We simply claim there is a cutoff roughly at a mass scale $\Lambda \sim T_{\text{el.}}^{1/p}$ given by the tension of the electric brane. The same derivation and caveats as in Section 6.2 apply here as well.

An important point is that (6.3.7) (and thus the magnetic WGC) clearly cannot hold for any value of $p$ and $D$. For $p = 0, D-2, D-1, D$ it gives nonsensical results. The solution (6.3.4) is not valid in these cases. This can be understood in physical terms: for $p = 0$, the electrically charged object is an instanton. Instantons do not decay via Hawking radiation, nor have an event horizon, so WGC considerations do not apply directly. The cases $p = D-2, D-1$ correspond to a string and a domain wall, respectively. The codimension of these objects is so low that they cannot have flat asymptotics, nor a horizon in the usual sense. Thus, the WGC is again ill-defined. For $p = D$ the object fills the entire spacetime (and is also subject to charge cancellation conditions), so this is not a very interesting case.

We will return to the cases $p = 0, D-1$ in four dimensions in the next two Chapters. Although they pose trouble for the black hole rationale for the WGC advocated in this Chapter, the T-duality argument above suggest that some form of the conjecture holds at least in string theory. We will see that these two case in fact turn out to be the ones of greatest phenomenological relevance.

## 6.4 Strong and mild forms

So far, we have discussed the only form of the WGC supported by black hole evaporation arguments, which says there should be a superextremal object in the theory. Henceforth we will call this the mild version of the WGC. Although very interesting, this form of the conjecture is not very useful for the phenomenologist: It can never constrain any field theory model, since it only constrains the charge-to-mass ratio of the WGC particle. One can satisfy it in principle with a charged object with arbitrarily high mass, far above any cutoff our effective field theory might have.

In the original WGC paper [139], other two possibilities were considered:

- The state of *least* charge under the $p$-form already satisfies (6.1.6); we will call this the *first strong form* of the WGC. One can always define the gauge coupling so that the state of least charge in the theory actually has $Q = 1$; hence, this form of the WGC tells us that there is a state with unit charge which is superextremal. The tension of this state gives the magnetic WGC cutoff.

- The *lightest* state charged under the $p$-form field already satisfies (6.1.6); this is the *second strong form* of the WGC.

Clearly, the first strong form implies the second, for any charged state lighter than a superextremal state with $Q = 1$ must also be superextremal. The converse does not hold; picture a theory in which the state with $Q = 1$ is very massive and does not satisfy the WGC, but the state with $Q = 2$ does. Both strong forms naturally imply the mild form, and as noted above the converse is not true: A theory in which every state of low charge is below the extremality bound, but there is some high charge state which is not would satisfy the mild WGC but neither of the strong forms. All this is illustrated in Figure 6.1.

The original WGC paper [139] claimed to have a counterexample for the first strong form. It is based on the heterotic example of Section 6.1. There, we considered a $U(1)$ along the Cartan generator $(1, 0, 0\ldots)$. The charge of any state under this Cartan is the first component of their charge vector $\vec{Q}_L$. However, the lattice also allows a vector $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$, which by this normalization would have charge $1/2$. This vector has norm $|\vec{Q}_L|^2 = 4$, so that its mass squared is $8/\alpha'$. According to the formulae of Section 6.1, its squared charge-to-mass ratio is $1/16$. However, it has been recently noticed [147] that this counterexample involve a particle charged under several $U(1)$'s. Once this fact is properly taken into account, as we will do in the next section, the first strong form of the WGC still survives.

It is notoriously difficult to find counterexamples to the second strong form in string theory; to my knowledge, none are known. Again, we get some help whenever there is extended supersymmetry, since presence of BPS states with every value of the charge then guarantees both strong forms to hold. Nevertheless, with the only counterexample to the first strong form gone, so far both strong forms are fine in the string theory context.

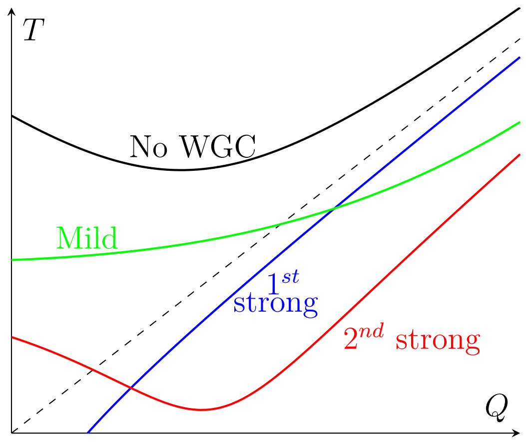


**Figure 6.1:** An illustration of the charge-tension relationship of theories satisfying different versions of the WGC. The extremality bound is the dashed diagonal line; states above it are subextremal, whereas those below are superextremal. Although we have depicted continuous lines, actual theories only have states at integer values of the charges due to Dirac quantization.

## 6.5 The WGC for several $U(1)$'s

So far we have only been concerned with theories with a single $U(1)$. However, as hinted at the end of the previous section, this is not the end of the story. In theories with several $U(1)$'s, we can have black holes charged under several of them at the same time. This has consequences both for the strong and mild forms of the WGC, which we now review.

### 6.5.1 The Convex Hull Condition

The extension of the mild WGC to the case of several $U(1)$'s was worked out in [156], and it goes under the name "Convex Hull Condition". Consider, for simplicity, an example with two $U(1)$ gauge fields in four dimensions, and couplings $g_1$, $g_2$. A RN black hole can be charged under both $U(1)$'s, with integer charges $(Q_1, Q_2)$ living on an integer lattice. The extremality condition is now (assuming canonical kinetic terms for both $U(1)$'s)

$$M_{\text{extremal}} = \sqrt{(g_1 Q_1)^2 + (g_2 Q_2)^2} M_P. \tag{6.5.1}$$

Instead of a single charge-to-mass ratio, we now have a charge-to-mass vector, $\vec{\xi}$, with components $\xi_i = g_i Q_i M_P / M$. The extremality condition becomes $|\vec{\xi}| = 1$. Subextremal black holes lie in the region $|\vec{\xi}| < 1$, while superextremal states satisfy the opposite inequality $|\vec{\xi}| > 1$.

In more general setups, however, extremal black holes will not lie on a sphere in $\vec{\xi}$-space. In general, (6.5.1) will be replaced by a complicated relationship between mass and the charges,

$$M_{\text{extremal}} = f_e(g_i Q_i) M_P. \tag{6.5.2}$$

The classical low-energy theory is invariant under a rescaling of every gauge coupling and mass scale by the same constant. This means that $f_e$ is homogeneous of degree 1 in every variable. As a result, we can rewrite the extremality condition as $f_e(g_i Q_i M_P/M) = 1$, or $f_e(\vec{\xi}) = 1$, a function of the charge-to-mass vector only. We will call this curve in the $\vec{\xi}$ plane the extremal curve. As an important particular example, take the black holes to be mutually BPS. Then,

$$M_{\text{extremal}} = (g_1|Q_1| + g_2|Q_2|)M_P, \tag{6.5.3}$$

so that the extremal black holes describe a diamond in the $\vec{\xi}$ plane. However we will stick to (6.5.1) for the examples in the remainder of this section.

With an extremality bound, we can redo the WGC argument. To avoid trouble with remnants, we clearly should allow black holes with charges $(Q_1, 0)$ and $(0, Q_2)$ to decay via the corresponding WGC particle. This is just applying the mild WGC to the two $U(1)$'s separately, but it is not enough. Consider a black hole with identical charges under both $U(1)$'s, and let us imagine it tries to release its charge by emitting a WGC particle of charges $(q_1, 0)$ and mass $m_1$, with charge-to-mass ratio vector $\vec{\xi}_p$. After the emission, the black hole sees a change in its charge-to-mass vector (here $\vec{u}_i$ is the unit vector in direction $i$)

$$\Delta\vec{\xi} \approx \frac{1}{M}\left(g_i \Delta Q_i \vec{u}_i M_P - \Delta M \vec{\xi}\right) = \frac{m_1}{M}\left(\vec{\xi} - \vec{\xi}_p\right) \tag{6.5.4}$$

so that

$$\Delta|\vec{\xi}|^2 = 2\vec{\xi}\cdot\Delta\vec{\xi} = \frac{2m_1}{M}\left(1 - \frac{q_1 g_1 M_P}{\sqrt{2}m_1}\right). \tag{6.5.5}$$

This should be $\leq 0$, to ensure that the black hole remains subextremal. However, the WGC for the first $U(1)$ only demands $m_1 \leq q_1 g_1 M_P$, which is not enough. Clearly, the WGC for several $U(1)$'s poses more stringent requirements than just applying the WGC to each particle individually.

For a general extremality relation $f_e(\vec{\xi}) = 1$, we should demand that, for each point on the extremality bound, there should be a particle allowing the black hole to decay while remaining subextremal. This just means that for every $\vec{\xi}$ on the extremality curve, there is a particle with a $\vec{\xi}_p$ so that

$$\nabla f_e|_{\vec{\xi}} \cdot (\vec{\xi} - \vec{\xi}_p) \leq 0. \tag{6.5.6}$$

This is the general statement of the WGC for several $U(1)$'s. It has a nice geometrical interpretation, illustrated in Figure 6.2.

Phrased in this way, it would seem that we need an infinite amount of new particles to satisfy the WGC. While one can build examples in which this is the case [147], it is not a necessity. Imagine we have two particles with charge-to-mass vectors $\vec{\xi}_{p_1}, \vec{\xi}_{p_2}$. The black hole can emit these two particles at once, and the bound state behaves as a new particle. If we neglect any interaction energy between the emitted

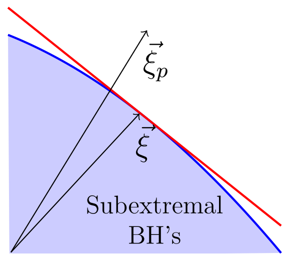


**Figure 6.2:** For any point on the extremal curve, draw the tangent hyperplane to the curve. The hyperplane partitions space in two regions, and the WGC demands the existence of a particle with charge-to-mass vector in the region not containing the subextremal black holes in a neighborhood of the chosen point.

particles (this will be the case if both are mutually BPS, or in the far future if the two particles end up far away from each other), then the charge-to-mass vector of the composite particle is simply the mass-weighted average of the two charge vectors. In the general case when we have $n_1$ particles of type 1 and $n_2$ particles of type 2, the charge-to-mass vector of all the particles together is

$$\vec{\xi}_{\text{comp.}} = \frac{\vec{\xi}_{p_1} n_1 m_1 + \vec{\xi}_{p_2} n_2 m_2}{n_1 m_1 + n_2 m_2}. \tag{6.5.7}$$

By taking large enough $n_1$,$n_2$ we can cover densely the line $x\vec{\xi}_{p_1} + (1-x)\vec{\xi}_{p_2}$, for $x \in [0,1]$. Thus, considering the bound states of a finite set of particles, with charge vectors $\Xi = \{\vec{\xi}_{p_i}\}$, we obtain every charge-to-mass vector lying on the boundary of the convex hull of $\Xi$. We will suppose that $\Xi$ contains at least as many linearly independent vectors as there are $U(1)$'s in the theory. In this case, any line $\alpha\vec{\xi}$ will intersect the boundary of the convex hull of $\Xi$.

In (6.5.6), let $\vec{\xi}_p$ the intersection of the line generated by $\vec{\xi}$ with the boundary of the convex hull. Then (6.5.6) is satisfied if the convex hull of $\Xi$ contains the region $f_e(\vec{\xi}) \leq 1$. This is the Convex Hull Condition, which allows us to satisfy the WGC with a finite number of particles. It is clearly a sufficient condition to satisfy (6.5.6); it is necessary for any differentiable extremal curve, since any branch of the extremal curve outside the convex hull contains at least one point for which (6.5.6) is not satisfied, as illustrated in Figure 6.3[1]. Figure 6.4 shows one example that satisfies the convex hull condition and another which does not.

The CHC works for any $p$-form field, just like the electric mild form does. The cutoff of magnetic version of the CHC could be taken as the smallest cutoff obtained for any bound state in the convex hull of $\Xi$.

[1]The differentiability condition is essential for this argument, but relaxing it does not lead to counterexamples since all the analysis surrounding (6.5.6) then becomes invalid. In any case, one should be able to replace any extremal curve by a smooth approximation to arbitrarily good accuracy, so this really covers the general case.

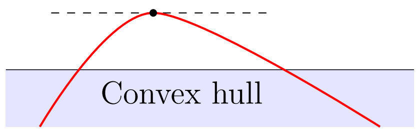


**Figure 6.3:** The CHC is necessary to satisfy (6.5.6) for a smooth curve. If the extremal curve, depicted in red, does not lie completely within the convex hull (the blue shaded region), the black dot in the Figure does not satisfy (6.5.6) (there are no superextremal particles above the dashed line).

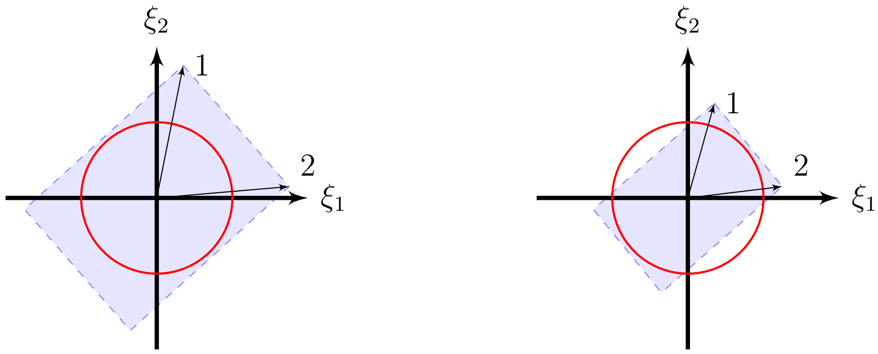


**Figure 6.4:** Illustration of the CHC condition. In the first picture, the extremal curve, in red, is contained within the convex hull spanned by particles 1 and 2, and so the WGC is satisfied. In the second case, it is not.

### 6.5.2 Strong forms for several $U(1)$'s: the Lattice WGC

Unlike in the previous case, there have recently been essentially two different proposals for a strong version of the WGC for many $U(1)$'s [157, 158, 147]. Just as in the case of a single $U(1)$, these stronger forms are not supported by black hole decay arguments.

The most natural extension of the second strong form is demanding that, for any direction on the charge lattice, the lightest state (be it a single particle, or a bound state of several particles) is superextremal. In this way, the usual strong form of a single $U(1)$ is satisfied for any particular $U(1)$ we might choose in the charge lattice. This was proposed by the Madison group in [157], and hence we will call it the Madison strong form. Notice that, unlike the mild version for several $U(1)$'s, which can be satisfied by a finite set of particles and their bound states via the CHC, the Madison strong form is a statement about an infinite set of particles (one per direction in the charge lattice).

Another candidate to strong WGC for several $U(1)$'s was briefly mentioned in [158]: the lightest set of particles whose charge-to-mass vector span the whole space must satisfy the CHC. This has the advantadge that it is again a statement about a finite set of particles. However, it does not imply the single $U(1)$ strong form for every $U(1)$, as the Madison form does. To show this, consider a theory with two identical $U(1)$'s such that two lightest particles along the directions $(1, 0)$ and

$(0,1)$ have charge vectors $(q,0)$ and $(0,q)$, and indentical mass $m = \frac{q}{\sqrt{2}} g M_P$. These particles satisfy the CHC, and let us assume that they are the lightest particles that do so. The WGC particle along the charge direction $(1,1)$ has mass $2m$ and charge $(q,q)$. But then, suppose the theory has a particle with mass $2m-\epsilon$ and charges $(1,1)$. This particle is very subextremal, and can also be the lightest particle along the direction $(1,1)$, so that the usual second strong form does not hold for this $U(1)$. In any case, this candidate strong form has been ruled out by a explicit M-theory example in [147].

The previous examples try to generalize the second strong form to systems of several $U(1)$'s. The reason these proposals did not tackle the first strong form was because for a long time it was thought not to be correct, based on the heterotic example given at the end of Section 6.4. However, in [147] it was pointed out that this example actually does not disprove the strong form, since it essentially involves particles charged under several $U(1)$'s at a time.

A "Madison-like" first strong form of the conjecture would be to require that, along every direction in charge space, there is a superextremal object with lowest possible charge. As in the single $U(1)$ case, this automatically implies the Madison strong form described above. It is now clear how the would-be counterexample to the first strong form given in Section 6.4 does not actually apply: The troublesome vector $(\frac{1}{2}, \frac{1}{2}, \ldots)$ does not lie in the same direction in the charge lattice as the $U(1)$ under consideration, which was $(1,0,\ldots)$. In fact, for any vector in the heterotic charge lattice with $N_L = 0$, we have, generalizing (6.1.8) slightly,

$$|\xi|^2 = \frac{|\vec{Q}_L|^2}{|\vec{Q}_L|^2 - 2} > 1, \qquad (6.5.8)$$

so that the strong form for several $U(1)$'s is indeed satisfied. While the charge-to-mass ratio diverges for $|\vec{Q}_L|^2 = 2$ (the $SO(32)$ W bosons), this can be remedied by the introduction of Wilson lines which break the non-abelian gauge group to the Cartan torus, introducing a shift in the denominator of (6.5.8) which removes the divergences. One should do this anyway, since the whole analysis of the WGC has been restricted to abelian gauge groups.

In fact, in light of (6.5.8), a stronger version is satisfied in this case: We see that for every point in the charge lattice there is a superextremal particle. This last statement goes under the name of the Lattice WGC [147], and it is the most general conjecture of the WGC family formulated so far. Like all the other candidates, it holds in every example in string theory known so far. This version of the conjecture blurs the distinction between black holes and elementary particles: along a given direction in the charge lattice, the states satisfying the LWGC for large charges will be (approximately) extremal black holes, whereas for low charges they become elementary particles.

The Lattice WGC is true in any perturbative heterotic CFT: The same argument that holds for the mild WGC in this case, as described in Section 6.1, can be applied to arbitrary points in the charge lattice, since the argument around eq. (6.5.8) works for any current algebra.

The extra assumption compared to the Madison first strong form, i.e. the existence of charged states at every point as opposed to at every direction in the charge lattice, is motivated if one requires the WGC to be covariant under compactification, as we now review.

Consider a theory with a $U(1)$ in $D+1$ dimensions, which satisfies the (mild WGC). Upon dimensional reduction on a circle to $D$ dimensions, there will be two different $U(1)$'s in the system: One coming from dimensional reduction of the higher-dimensional $U(1)$, and the KK photon coming from the metric. The electric charge coupled to this $U(1)$ is precisely the KK momentum. The kinetic terms are diagonal, so that the extremal curve is a circle of radius 1.

If the WGC is well behaved under dimensional reduction, the theory in $D$ dimensions should satisfy the CHC. For a particle of mass $m_{D+1}$ and charge $q$ in $D+1$ dimensions, we will obtain particles with charge

$$(q,k) \quad \text{and mass} \quad m^2 = m_{D+1}^2 + \left(\frac{k}{R}\right)^2, \tag{6.5.9}$$

the KK tower associated to the $D+1$-dimensional particle. If the gauge coupling of the $U(1)$ in $D+1$ dimensions is $D+1$, then in $D$ dimensions it is $g_{D+1}/\sqrt{2\pi R}$, and Newton's constant is corrected by the same factor. The gauge coupling of the KK photon is simply $2\kappa_D^2/R$, but the extremality condition for KK black holes has an extra factor of $1/2$ because of the coupling to the radion. Putting everything together we find particles with charge-to-mass ratios given by [147]

$$\vec{\xi}_{q,k} = \frac{(m_{D+1}R\xi_{D+1}, n)}{\sqrt{m_{D+1}^2 R^2 + n^2}}, \tag{6.5.10}$$

where $\xi_{D+1} = g_{D+1}/\kappa_{D+1}$ is the charge-to-mass ratio of the particle in $D+1$ dimensions.

No matter what the value of $\xi_{D+1}$ is, the set of charge vectors (6.5.10) will never satisfy the CHC (see [147] for a detailed proof). This proves that the WGC is not covariant under dimensional reduction. However, if the $D+1$-dimensional theory satisfies the Lattice WGC, there is actually an infinite number of KK towers in the dimensionally reduced theory; these are enough to satisfy the WGC.

In this sense, demanding a good behavior under dimensional reduction provides evidence for the Lattice WGC. An interesting observation is that particles of high charge in the $D+1$ theory must have at least some KK modes outside of the convex hull of particles of lower charge, and which are hence stable. In general, the parent $D+1$ dimensional particles which give rise to these KK modes will be unstable, so that dimensional reduction has stabilized them. To decay, the KK modes should emit some particle with KK momentum, but this also increases its mass, in such a way that the decay channel is no longer available. This is similar to the inhibition of spontaneous emission for atoms placed between two mirrors, which greatly enhances the lifetime of some short-lived excited atomic states [159].

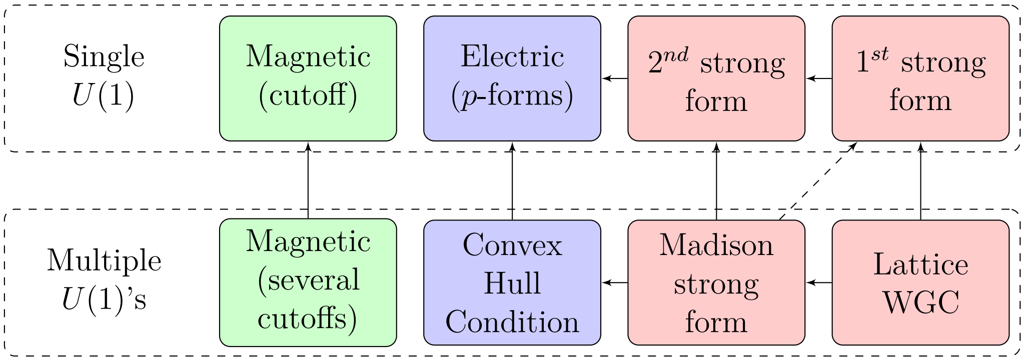


**Figure 6.5:** The different forms of the WGC and the relationships among them. In blue, forms of the WGC motivated by black hole evaporation arguments; in red, forms that are not. In green, the magnetic forms, which are supported by semiclassical considerations plus some mild assumptions about the effective field theory. Only a slight generalization of the Madison form (replacing "lightest charge" with "least charge"), not discussed in [157], implies the first strong form.

## 6.6 Summary and final comments

There are many forms of the WGC which apply in different contexts. While only the mild form and the Convex Hull Condition are motivated by generic quantum gravity considerations, all the other forms are conjectured to hold at least in string theory examples. This is supported by a variety of evidence. Figure 6.5 illustrates the different versions of the WGC reviewed in this Chapter and the relationships among them.

There are two possible generalizations of the WGC which come to mind which have not been discussed. The first is a WGC-like statement for angular momentum. Black holes carry angular momentum as they carry electric charge. Like electric charge, it is quantized in the quantum theory. However, there is no analog of the gauge coupling which we can tune to zero to get some remnant trouble: the number of extremal black hole states below a mass scale $M$ is fixed by the theory, and can only be tuned by changing $M_P$. On the other hand, in Minkowski space, it is possible to emit any given particle with orbital angular momentum, and thus as long as there is a fermion in the theory (to allow for the emission of half-integer angular momentum) a rotating black hole can kinematically lose its angular momentum without the addition of extra particles.

The second is the generalization to non-abelian gauge groups. While we have discussed in depth an heterotic example in which the underlying gauge group is non-abelian, the standard technique in this case is to turn on Wilson lines to break the gauge group to the Cartan torus. Representations of a non-abelian gauge group are indexed by their Dynkin labels, which are the charges under the Cartan generators of the group. Since the WGC is only concerned with conserved charges, it is reasonable to think there will not be much difference between $U(1)^{\operatorname{rank}\mathfrak{g}}$ with a particular metric (Killing form) on the charge lattice and the full $G$ as far as the mild WGC is concerned. Nevertheless, this is an interesting question to explore.

Another interesting project would be finding evidence for the WGC outstide of the umbrella of string theory, to figure out whether it is a property of the general QG Swampland, or just of the stringy portion of it. This will be difficult to ascertain in the near future, both because all of the QG Swampland might actually be stringy, and because as mentioned in Chapter 1 string theory is so far the only framework in which the WGC conjecture predictions can be checked. Holography may also shed new light on the WGC [160, 161].

The WGC and its many forms constitute an interesting crossroads at which generic quantum gravity considerations, string phenomenology, holography and field theory meet. No full formal proof of the conjecture has been given so far, and as discussed throughout this Chapter the arguments in its favor are not devoid of loopholes. Are some of its assumptions, such as weak coupling, validity of effective field theory, etc. essential, or can they be relaxed? Is the conjecture really valid for $p$-forms of low codimension or axions? These questions make the WGC an interesting and active area of research. In any case all, the stringy evidence gathered so far points out to the correctness of the conjecture, and as we will see in the next Chapters it has important implications for phenomenology.

# 7 Transplanckian axions

In this Chapter we will apply the ideas of the WGC, developed in the previous Chapter, to axion models of inflation. The first obstacle we meet is that the WGC does not directly apply to axion theories, since the charged objects are instantons. Nevertheless, one can still construct electrically charged solutions of the Einstein-axion system, and these instantons are enough to constrain the model in some cases.

We will start with a brief introduction, motivating the popularity of large field inflation models recently. We will then describe the shortcomings of the WGC in this system, and discuss the gravitational instanton solutions which show up for large charges in their stead. After that, these gravitational instantons (or their stringy counterparts, when needed), will be used to constrain some models of large-field inflation, notably the Kim-Nilles-Peloso (KNP) alignment proposal [162]. We will also discuss other popular models of large-field inflation which can evade our constraints.

The results of this Chapter were presented in [163], which appeared at around the same time as other works discussing quantum gravity constraints to large field inflation [164, 165, 166, 163, 157, 167, 168, 169, 170, 158, 171, 147, 172, 173, 174].

## 7.1 Why transplanckian axions?

The present precision era in the measurement of CMB observables is triggering a vigorous activity in model building of early universe inflationary models (see [175] for a recent string-motivated review). The most recent example is provided by the possibility of a sizable tensor to scalar ratio $r$, raised by the initial BICEP2 detection claim of B-mode polarization from primordial gravitational waves [176], and still allowed by the joint Planck/BICEP2 upper bound $r < 0.12$ [177]). In single field inflation models, this ratio is correlated with the inflaton field range in a very direct way by the Lyth bound [178],

$$\frac{\Delta\phi}{M_P} \approx N\sqrt{\frac{r}{8}}, \qquad (7.1.1)$$

where $\Delta\phi$ is the inflaton field range traversed during inflation, and $N$ is the number of e-folds. Because of this, there has been an intense activity in building inflation models in which the inflaton travels transplanckian distances in field space. The notes [179] provide a self-contained and clear introduction to the Lyth bound and large field inflation.

Such large-field inflation models are sensitive to an infinite number of corrections to the inflaton potential even if they are suppressed by the Planck mass scale, so their construction requires accounting for corrections due even to quantum gravity effects. Given the difficulties in addressing the latter, a sensible approach is to invoke symmetries protecting the model against such corrections; therefore, many large-field inflation models are based on axions, i.e. scalars $\phi$ with an approximate continuous shift symmetry, broken by non-perturbative effects $e^{i\phi/f}$ to a discrete periodicity

$$\phi \sim \phi + 2\pi f. \tag{7.1.2}$$

This idea was originally proposed at the phenomenological level in so-called natural inflation [180], assuming a cosine gauge instanton potential, and requiring axions with transplanckian decay constant $f$ (see [181] for recently suggested exceptions). The 'empirical' case-by-case realization that this requirement is not obviously realized in string theory compactifications [182] (which provide a template for a quantum gravity framework, as discussed in Chapter 6), motivated the construction of models with multiple axions and potentials from gauge instantons, in which some axion linear combination effectively hosts a transplanckian field range in its basic period, even if the original periodicities are taken sub-planckian [162, 183] (see [184, 185, 186, 187, 188, 189, 190, 191, 192, 193] for recent works).

These models are formulated in purely phenomenological field theory terms, and therefore there remains the question of whether their features survive in actual embeddings in theories including quantum gravitational corrections i.e. whether they are in the Swampland or not. If one were to apply the WGC naively to the axion case, there would be additional instanton effects leading to higher harmonics in the axion potential, i.e. $e^{in\phi}$, in any consistent theory of quantum gravity (see [165, 166]). Because the presence of gravity is essential in the argument, these contributions may be beyond those captured by gauge instantons in the phenomenological model. If sufficiently strong, such contributions could spoil the transplanckian field range by introducing additional maxima along the axion period.

However, as mentioned in the previous Chapter, one cannot apply directly the WGC to axions. To go around this we undertake a direct approach to this problem, by considering euclidean Einstein gravity and explicitly constructing gravitational instantons coupling to the relevant axions, and evaluating their contributions to the axion potential. The configurations are wormhole solutions, and for sufficiently large axion charge have low curvature and should provide good effective descriptions of the corresponding configuration in any UV quantum theory containing Einstein gravity. [157] takes a different approach, arguing for the validity of the WGC for axions in string theory by turning them into more ordinary $p$-forms via a chain of dualities. While this approach is more precise, it is automatically restricted to the string theory context. However, we will see in the next Chapter that the approach taken here does not work for 3-forms. In that case, the T-duality argument still holds.

The gravitational instantons are effective descriptions of configurations in concrete models of quantum gravity like string theory. For instance, in realizations of natural or aligned inflation, similar effects can arise from euclidean D-brane instantons beyond those corresponding to gauge instantons (dubbed exotic or stringy [194, 195, 196], see [197, 198] for a review). An important aspect, which we emphasize and has seemingly been overlooked in the literature (see [199, 200, 201, 202, 203, 204] for recent attempts to embed such models in string theory), is that the corresponding instantons may not correspond to BPS instantons in the vacuum. Non-BPS instantons are not often discussed in the literature, because in supersymmetric setups they do not lead to corrections to the superpotential, but rather to higher F-terms (higher derivative or multi-fermion terms) [205]. However, in non-supersymmetric situations like inflation, their extra fermion zero modes are lifted or equivalently the external legs are saturated (by insertions of susy breaking operators), and their contributions descend to the scalar potential [206]. These considerations should be implemented to analyze actual string theory realizations of natural and aligned inflation (see [207] for a possible exception).

## 7.2 Weak gravity conjecture and axions

As discussed in the previous Chapter, the electric WGC implies that for any abelian $p$-form field there must be a charged $p$-dimensional object with tension

$$T \lesssim \frac{g}{\sqrt{G_N}}. \tag{7.2.1}$$

Here, $g$ is the coupling of the $p$-form field to its sources, and it has dimensions of $[\text{mass}]^{p+1-D/2}$. For $p=0$, that is, an axion, the coupling $g$ is nothing but the inverse of the axion decay constant $f$, and so the above conjecture formally implies the existence of an instanton with action $S \lesssim \frac{M_P}{f}$. This is usually taken to imply [139] that the axion cannot have a parametrically flat potential. However, as mentioned in the previous Chapter, there is a very strong difference between $p=0$ and $p>0$; whereas in the latter we do have stable black hole (in general, black brane) remnants unless an object satisfying (7.2.1) exists, there is no similar argument for the axion, since there is no would be stable black hole solution: the electrically charged object is a gravitational instanton. When $f \gg M_P$, we have many such instantons with small action, and computing their effect will be a challenge; however, there does not seem to be any inconsistency of the kind described in [139].

Although this shows that the WGC for axion fields is not on such a firm ground, it does highlight the fact that in a consistent quantum theory of gravity one expects to have gravitational instantons coupled to the axion. These can be found just by solving the Euclidean equations of motion for the axion-gravity system, while specifying the right asymptotic charge. This is in accordance with the considerations in [42] that in a quantum theory of gravity one expects to have all possible charged objects.

A similar argument works also for Euclidean theories. As discussed above, one can build a solution to the euclidean equations of motion with nontrivial axion charge. This is an instanton. The question now is wether it is possible to build a consistent theory in which these instantons are excluded of the path integral, or on the contrary their presence cannot be avoided. This seems unlikely, since in a quantum theory of gravity we are expected to add up all geometries with the right asymptotics [208]. Although a single instanton does not fulfill this condition, since its axionic electric charge may be measured at infinity, an instanton- anti instanton pair does. Once we allow the instanton- anti instanton amplitude in the path integral, it is difficult to exclude the single instanton as a configuration mediating a transition between two different states; one could always cut a time slice between the instanton and the anti-instanton. This constant time configuration defines a state within the theory, so it makes sense to consider finite action configurations connecting it to the vacuum (that is, a single instanton).

The smallest such instanton will involve Planckian geometries, and so it should not be taken very seriously. However, some instantons may turn out to be within the reach of semiclassical Einstein gravity. Their effects on axion dynamics can be safely studied even in the absence of an UV completion.

We will now start constructing gravitational instantons coupling to axions, and their effects, in theories of gravity coupled to axions. These do not include other fields like scalar partners of the axions to turn them into complex fields, fermions or other superpartners, etc. The motivation for this is twofold: First, for applications to single-field inflation, the inflationary dynamics requires the axion to be the only (non-gravitational) dynamical field; in fact, the main challenge in realizing inflation in UV complete theories, like string theory, lies in giving hierarchically large masses to all fields except for the inflaton, so that they do not interfere with inflation. Hence, gravity coupled to axions provides the appropriate setup to address questions pertaining the possible gravitational corrections to the flatness of the inflationary potential. Second, we would like to keep our framework as minimal and general as possible, to explore the interplay of gravitational instantons, axions, and the Weak Gravity Conjecture in general gravitational theories, in a model-independent way.

Throughout this Chapter we assume that the axion which is to act as the inflaton somehow has an adequate potential to provide successful inflation. In typical models this comes e.g. from a gauge sector coupled to the axion. We will not address how this potential is obtained: we will take it as given and then ask ourselves if gravitational effects can spoil it.

We first study the consequences of gravitational instantons for a single axion with periodicity $\phi \sim \phi + 2\pi$. The Minkowskian action for the theory is

$$\int \left( \frac{-1}{16\pi G} R dV + \frac{f^2}{2} d\phi \wedge * d\phi \right) + \dots \tag{7.2.2}$$

where the dots stand for any extra light fields. The euclidean counterpart of (7.2.2) is

$$\int \left( \frac{-1}{16\pi G} RdV - \frac{f^2}{2} d\phi \wedge *d\phi \right) + \dots \tag{7.2.3}$$

Notice that naive Wick rotation would have produced the opposite kinetic term for the axion. This extra minus sign for an axion as compared to a standard scalar field is related to it being dual to a three-form; duality involves the Hodge star operator and since $*^2 = (-1)^{k(n-k)}s$ for $k$-forms in $n$-dimensional space with metric signature $s$, dualizing to a three-form and Euclidean rotation do not commute. The best semiclassical approximation is obtained by first dualizing to a three-form, then performing euclidean rotation, and then dualizing back to a scalar; as a result the extra minus sign arises. For a detailed explanation of this point, see [209].

### 7.2.1 The instanton

We are looking for a solution of (7.2.3) which is asymptotically flat and with an axion profile which at large distances from its core has the asymptotic form

$$\phi \sim \frac{n}{4\pi^2 f^2} \frac{1}{r^2}. \tag{7.2.4}$$

This means it has electric charge $n$, since asymptotically $*d\phi = \frac{n}{2\pi^2 f^2} d\Omega_3$.

With everything discussed above, we only need to find the right gravitational instanton. This will be a spherically symmetric solution with the right axionic charge. To do this we take a spherically symmetric ansatz for the metric,

$$ds^2 = f_I(r) dr^2 + r^2 ds^2_{S^3}, \tag{7.2.5}$$

where we have redefined the radial coordinate so that constant-$r$ surfaces behave like proper spheres. The only dependence of the metric is through the radial factor $f(r)$. Similarly, the axion profile will be taken as a function only of $r$. We thus get a system of two coupled second-order ODE's which is easy to analyze in detail.

The stress-energy tensor of an axion field is

$$f^{-2} T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} (\partial_\mu \phi \partial^\mu \phi) g_{\mu\nu}. \tag{7.2.6}$$

The only nonvanishing components of these are

$$T_{rr} = \frac{1}{2} f^2 (\phi')^2, \quad T_{ii} = -\frac{1}{2 f_I} (f\phi')^2 g_{ii}, \quad i = \psi, \theta, \phi. \tag{7.2.7}$$

Einstein's equations yield

$$\begin{aligned} \frac{3}{r^2}(1 - f_I) &= -4\pi G (f\phi')^2, \\ -f_I + 1 - r\frac{f_I'}{f_I} &= 4\pi G r^2 (f\phi')^2. \end{aligned} \tag{7.2.8}$$

The solution is

$$f_I(r) = \frac{1}{1-\frac{a}{r^4}}, \quad d\phi(r) = \frac{n}{2\pi^2 f^2}\sqrt{f(r)}\frac{dr}{r^3}, \quad a \equiv \frac{n^2}{3\pi^3}\frac{G}{f^2} = \frac{n^2}{3\pi^3}\left(\frac{M_P}{f}\right)^2 M_P^{-4}. \tag{7.2.9}$$

The parameter $a^{1/4}$ gives the typical size of the instanton; at this radius the metric seems to become singular. This is a coordinate singularity, as the curvature scalar

$$R = 8\pi GT = 8\pi G f^2 \partial_\mu\phi\partial^\mu\phi = 8\pi\frac{G}{f^2}\left(\frac{n}{4\pi^2}\right)^2\frac{1}{r^6} = \frac{3}{2}\frac{a}{r^6}. \tag{7.2.10}$$

is perfectly regular there. Notice that for $r = a^{1/4}$, which corresponds to the maximum curvature at the throat of the wormhole, $R = \frac{3}{2}a^{-1/2}$. Thus the single parameter controlling wether the solution can be trusted is $a$; it has to be larger than the Planck scale for the solution to be reliable within the context of effective field theory. Notice also that the wormhole radius $\sim a^{1/4}$ is of order Planck for $n = 1$; we need wormholes with higher axionic charge.

Since we have defined our radial coordinate in terms of the areas of 3-spheres, what is happening is that as we go to $r = a$ from infinity these areas reach a minimum, then start growing again. In fact, by making the change of coordinates $r^4 = a + t^2$, the metric becomes

$$ds^2 = r^2\left[\frac{1}{4}dt^2 + ds^2_{S_3}\right], \tag{7.2.11}$$

which is conformally equivalent to a flat metric on $S^3\times\mathbb{R}$; in other words, our instanton is actually a wormhole. In fact, we have just rederived the Strominger-Giddings wormhole [210]. As discussed there, it is up to us to choose which geometries is the wormhole connecting: Two asymptotically flat regions is clearly unphysical. We might think of an euclidean wormhole connecting two points of spacetime, with an arbitrarily long tube between them; thus, we might better think of half our solution describing the "entrance" of a wormhole (instanton), and a solution with negative axionic charge describing the exit (anti-instanton).

In [210], these half-wormholes are regarded as describing the nucleation of baby universes, which then evolve on their own towards a Big Crunch. The interpretation is very similar to a gauge instanton, which changes the asymptotic configuration of the gauge fields, in particular the Chern-Simons number of the vaccum

$$n_{CS} = \frac{1}{3}\int_{\mathbb{R}^3}\mathbf{A}\wedge\mathbf{A}\wedge\mathbf{A}. \tag{7.2.12}$$

Since two states with different $n_{CS}$ are degenerate in energy, the true vacuum is a $|\theta\rangle$ vacuum of the form

$$|\theta\rangle = \sum_{n_{CS}} e^{in_{CS}\theta}|n_{CS}\rangle. \tag{7.2.13}$$

The analog of the Chern-Simons number in the gravitational case would be the number of connected components of space. A half-wormhole nucleates a baby universe which undergoes a Big Crunch; after the wormhole has closed, we end up with a Universe with two connected components. The instanton thus increases this number by one. Again by analogy with the gauge case, one could think that the actual ground state is one of the so-called "$\alpha$-states" of Coleman [211, 212, 213]; in this case, the half-wormholes would generate a potential for the axion in exactly the same way as the instantons in gauge theory. We will discuss the effect of the instanton in the axion potential further below.

The only other relevant quantity of the solution is its action. This can be computed by noting that the equations of motion imply $R = 8\pi GT$, and $T = -f^2(\partial\phi)^2$, so that[1]

$$S_E = \frac{f^2}{2}\int dV\ \partial_\mu\phi\partial^\mu\phi = \pi^2 f^2 \int_{a^{1/4}}^{\infty} f(r)^{-1/2} r^3 (\phi'(r))^2\ dr = \frac{\sqrt{3\pi}}{16}\frac{M_P n}{f}. \quad (7.2.14)$$

where in the first equality we have divided by two to take into account the fact that our instanton is only half a wormhole.

We now have to argue that this gravitational instanton can indeed generate the right potential for the axion. Let us expand the axion field in the background of an instanton of charge $n$ as $\phi \approx \langle\phi_n\rangle + \varphi$, where $\varphi$ is the fluctuation. The axion kinetic term thus yields

$$\begin{aligned} &d(\langle\phi_n\rangle + \varphi) \wedge *d(\langle\phi_n\rangle + \varphi) \\ &= d\langle\phi_n\rangle \wedge *d\langle\phi_n\rangle + d\varphi \wedge *d\varphi + d\langle\phi_n\rangle \wedge *d\varphi + d\varphi \wedge *d\langle\phi_n\rangle. \end{aligned} \quad (7.2.15)$$

The last two terms may be rearranged as

$$\begin{aligned} d\langle\phi_n\rangle \wedge *d\varphi + d\varphi \wedge *d\langle\phi_n\rangle = 2d\varphi \wedge *d\langle\phi_n\rangle &= 2d(\varphi \wedge *d\ \langle\phi_n\rangle) - 2\varphi \wedge (d * d\langle\phi_n\rangle) \\ &\sim -2n\varphi \wedge (d * d\langle\phi_n\rangle), \end{aligned} \quad (7.2.16)$$

where we dropped the total derivative term. We thus get the usual axion-instanton coupling term, with the instanton supported on the form $(d * d\langle\phi_n\rangle)$. In the dilute-gas approximation, summing over all these instantons and fluctuations around them gives a contribution to the path integral of the form (switching back from $\varphi$ to $\phi$)

$$S_{inst.} = \int d^4x\sqrt{-g}\mathcal{P}e^{-S_E}\cos(n\phi), \quad (7.2.17)$$

where $\mathcal{P}$ is a prefactor whose computation whose precise computation has only been achieved in a few cases [214, 215]. We will take $|\mathcal{P}| \sim M_P^4$, since it is at this scale when the effective description breaks down. This is in accordance with previous literature, see [166].

However, it is also essential to know whether or not $\mathcal{P}$ has an imaginary part. This will be the case if the spectrum of fluctuations around the instanton background

[1] We thank [167] for pointing out a missing factor of 1/2 in an earlier version of this manuscript.

has a negative eigenmode. In this case, (7.2.17) does not have an interpretation as a potential. The instanton is not mediating transitions between different vacua, rather it is telling us that Minkowski space is an unstable configuration which can decay to something else, in the spirit of Witten's bubble of nothing [216] and of Coleman and DeLuccia's [217, 218] bounce solutions.

In [219, 220], it is argued that the Giddings-Strominger wormhole has exactly one negative eigenmode, and thus $\mathcal{P}$ is purely imaginary. However, as discussed above, the instanton under consideration would be something like half a Giddings-Strominger wormhole; we would rotate back to an Euclidean solution after the wormhole throat. It is not clear that the negative mode survives this cutting; however, since we have not computed the instanton prefactor, we do not know wether this solution is truly an instanton or a bounce.

Although the physical effects of an instanton and a bounce are very different, the distinction is not very important if we are concerned only with the possibility of transplanckian axions as candidates for inflation. The instanton would generate a huge potential, cutting down the effective field range to $\phi$'s satisfying

$$|\mathcal{P}|e^{-S_E}\cos(n\phi) \sim V_0, \tag{7.2.18}$$

where $V_0$ is the required height of the inflaton potential in our model. Similarly, the bounce will initiate a decay to the true vacuum, spoiling inflation and the Universe with it, unless we restrict ourselves to the same field range (7.2.18). For convenience from now on we will take $\mathcal{P}$ to be real and discuss the physics in terms of axion potentials. The reader should keep in mind the possibility that we are actually talking about bounces.

### 7.2.2 Consequences for a transplanckian axion

If we want the full range of the axion $f$ to be available for inflation, we need (7.2.17) to be very suppressed. In other words, $S_E \gg 1$. By looking at (7.2.14), this means $M_P n \gg f$. Thus, effects of gravitational instantons constrain the effective axion decay constant. If $f \sim M_P$, the gravitational instantons with low $n$ will not be suppressed. Although strictly speaking we cannot trust our computation for low $n$ since the wormhole throat is of Planckian size, it is reasonable to say $f \sim 10M_P$ is ruled out. In any case, this kind of effects do rule out parametrically flat axion potentials, within the regime of low energy effective field theory.

The effective field range for the action is determined by the first instanton contribution to become relevant. This is around $S_E \sim 1$, or in other words $n \sim \frac{16f}{\sqrt{3\pi}M_P} \approx 5\frac{f}{M_P}$. This generates a potential of the form $\cos(n\phi)$ which means that in terms of the canonically normalized axion $\zeta = f\phi$, the effective field range is just $\sim M_P/5$.

This bound is quite similar to those obtained directly from the WGC (7.2.1), where it was argued that there should be instantons in the theory which creates a potential with higher harmonics which makes the field range subplanckian, for all intends and purposes. The gravitational instantons discussed here are not those demanded by the WGC; the latter are postulated to allow the "decay" of the former.

Axions arising from string theory compactifications typically couple to extended instantonic objects in the theory, for instance euclidean D-branes. The supergravity perspective regards D-branes as extended, extremal black hole solutions [52], and so they are in a sense the stringy version of the instantons discussed above. On the other hand euclidean D-branes are often claimed to be the instantons required by the WGC. We will elaborate further on this point in Section 7.5.

The above calculations assumed the validity of Einsteinian gravity up to the Planck scale. This is not generically true in string compactifications, for which this description breaks down at the compactification scale. To be more specific, the range of validity of the instanton calculation is bounded by both sides: In order to trust the gravitational instanton solution the wormhole throat should be larger than some scale $\Lambda$ at which extra moduli become light or we start to see stringy physics in some way; at the same time the above computations neglect the contribution of the axion potential coming from the gauge instantons to the stress-energy of the solution. As most of the contribution to the action comes from the vicinity of the throat, we must demand that it is larger than $V_0^{-1}$, where $V_0$ is the scale of the inflationary potential. Thus, we have the bounds

$$\Lambda^{-1} < a^{1/4} < V_0^{-1} \quad \Rightarrow \quad \Lambda^{-1} < \frac{\sqrt{n}}{(3\pi^3)^{\frac{1}{4}}\sqrt{M_P f}} < V_0^{-1}. \tag{7.2.19}$$

## 7.3 Gravitational instantons for multiple axions

These bounds on transplanckian decay constants coming from purely quantum gravitational arguments agree with the difficulties found in string theory to provide a transplanckian axion. In [182] was shown for several examples that in order to have a transplanckian decay constant for an axion, one is always forced to go beyond the safe and controlled regime of the effective theory. New contributions, usually in the form of higher harmonics to the potential, become relevant at scales $\mathcal{O}(M_P)$. In [221] the same question was addressed in the context of F-theory, finding similar harmonics that would forbid parametrically large decay constants. People have tried to evade these difficulties by considering models of multiple axions in which you might hope to engineer a direction in the moduli space with an enhanced effective field range. However there is not a completely successful embedding of any of these models yet in string theory (see [199, 200, 201, 202, 203, 204] for recent attempts). While the failure to get a single transplankian axion in string theory is quite widely accepted, the same question when there is more than one axion is not clear at all. Here we generalize the computation of the previous Section to the case of multiple axions and study the gravitational constraints that our instanton imposes over the effective field range available for inflation. We will see that while parametrically flat directions are inconsistent with the presence of gravitational instantons, we can still have a certain numerical enhancement proportional to $\sqrt{N}$ with $N$ being the number of axions, which survives to the gravitational effects.

### 7.3.1 Axion-driven multiple field inflation

Let us consider an effective field theory containing N axions $\phi_i$. The shift symmetries will be broken by non-perturbative effects inducing a scalar potential which is still invariant under discrete shifts $\phi_i \to \phi_i + 2\pi$. The periodicities of the axions define a lattice in the field space of side length $2\pi$. The effective lagrangian of the system takes the form

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \vec{\phi})\mathcal{G}(\partial^\nu \vec{\phi}) - \sum_{i=1}^{N} \Lambda_i^4 (1 - \cos(\phi_i)) \tag{7.3.1}$$

where in general the kinetic term is not necessarily canonical. We will refer to this basis as the lattice basis. We can go now to the physical or kinetic basis, in which the kinetic terms are canonically normalized by making a rotation to diagonalize

$$\mathcal{G} = \mathbf{R}^T diag(f_i^2)\mathbf{R} \ , \quad \vec{\zeta} = \mathbf{R}\vec{\phi} \tag{7.3.2}$$

followed by a field redefinition given by $\hat{\zeta}_i = f_i \zeta_i$. The effective lagrangian in the kinetic basis reads

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \hat{\zeta}_i)^2 - \sum_{i=1}^{N} \Lambda_i^4 (1 - \cos(R_{ji}\hat{\zeta}_i/f_i)) \ . \tag{7.3.3}$$

If we diagonalize now the potential we get a new lattice which depends non-trivially on the kinetic eigenvalues $f_i$ and the rotation angles. The maximum physical displacement in field space is given by half of the diagonal of this new lattice. If $\Lambda_i = \Lambda$ for all $i$, this theory gives rise to $N$ fields with masses $m_i^2 = \Lambda^4/f_i^2$ and it turns out that the lightest field is also the one with the largest available field range. In general, in order to obtain successful inflation along the largest direction in field space we need to tune the parameters $\Lambda_i$. This can lead to some difficulties when trying to embed the model in string theory. However, since this issue is very model dependent, we will ignore it and assume that the parameters $\Lambda_i$ can always be tuned such that inflation would occur along the direction of interest. Our focus will be in finding out whether the flat direction survives or it is shortened when the gravitational instantons are taken into account.

The different ways to achieve large field inflation with multiple axions can be divided in three main classes:

- N-flation [183]
- Kinetic alignment [188]
- Lattice alignment [162]

In N-flation models the kinetic metric is assumed to be diagonal already in the lattice basis. Then both lattice and kinetic bases are proportional to each other and related by the field redefinition $\hat{\zeta}_i = f_i \phi_i$, so the original lattice is mapped to a new rectangular lattice with side lengths $2\pi f_i$ (see fig.7.1 (left)). The maximum displacement is then given by a collective mode traveling along the diagonal,

$$\Delta\Phi = \pi\sqrt{f_1^2 + \cdots + f_N^2}\ . \tag{7.3.4}$$

In the favorable case in which all the decay constant are of similar order $f_1 \sim \cdots \sim f_N$ we get an enhancement of $\sqrt{N}$.

In models based on kinetic alignment the kinetic metric is not assumed to be diagonal when written in the lattice basis. In fact, one can used this freedom to choose the eigenvector $\zeta_i$ with the largest eigenvalue $f_i^2$ pointing along a diagonal of the lattice, such that the maximum displacement is then given by $\pi\sqrt{N} f_i$. This relaxes the requirements to get large field inflation with $N$ axions, since it is enough to have one entry of the kinetic matrix large, while the others can remain small. For concreteness let us consider two axions. The kinetic basis is given by

$$\vec{\hat{\zeta}} = \begin{pmatrix} f_1 & 0 \\ 0 & f_2 \end{pmatrix} \mathbf{R}\vec{\phi} = \begin{pmatrix} f_1(\phi_1 \cos\alpha + \phi_2 \sin\alpha) \\ f_2(\phi_2 \cos\alpha - \phi_1 \sin\alpha) \end{pmatrix} \tag{7.3.5}$$

where $f_1, f_2$ are the eigenvalues of the kinetic matrix and we have parametrized the rotation matrix of (7.3.2) as

$$\mathbf{R} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{7.3.6}$$

It can be easily checked that the maximum physical displacement can be achieved if the eigenvector $\zeta_1$ (assuming $f_1 > f_2$) points in the direction of the diagonal of the new lattice, ie. $\alpha = \pi/4$, obtaining

$$\Delta\Phi = \pi\sqrt{2} f_1\ . \tag{7.3.7}$$

The situation of perfect kinetic alignment is depicted in fig.7.1 (right), in which the inflationary trajectory (red arrow) coincides with the direction determined by $\hat{\zeta}_1$. For later use the metric in the lattice basis can be written in general as

$$\mathcal{G} = \begin{pmatrix} f_1^2\cos^2\alpha + f_2^2\sin^2\alpha & (f_1^2 - f_2^2)\cos\alpha\sin\alpha \\ (f_1^2 - f_2^2)\cos\alpha\sin\alpha & f_1^2\sin^2\alpha + f_2^2\cos^2\alpha \end{pmatrix} \tag{7.3.8}$$

where $\alpha = \pi/4$ for perfect kinetic alignment and $\alpha = 0$ for the original models of N-flation.

The Kim-Nilles-Peloso proposal of lattice claims to achieve a parametrically flat direction in the potential starting from subplanckian decay constants. However this large direction, although hidden in some choice of basis for the lattice, is there

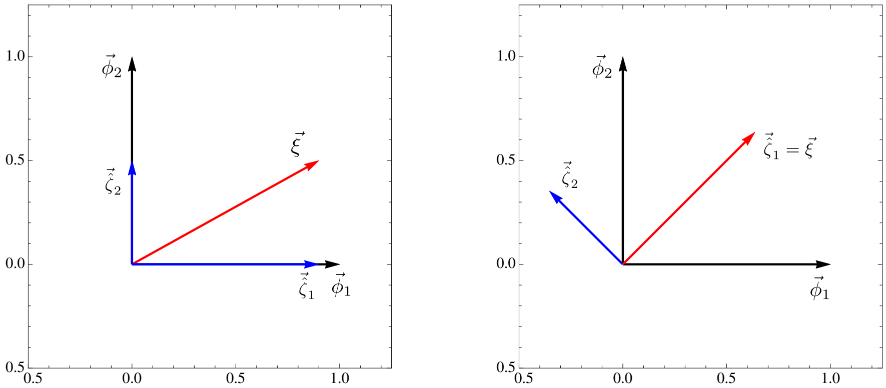


**Figure 7.1:** Different choices of basis for N-flation (left) and kinetic alignment (right) models.

from the beginning and can be made manifest by going to the kinetic basis. Let us consider the lagrangian proposed in the original paper [162]

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \theta_1)^2 + \frac{1}{2}(\partial_\mu \theta_2)^2 - \Lambda_1^4\left(1 - \cos\left(\frac{\theta_1}{f_1} + \frac{\theta_2}{g_1}\right)\right) - \Lambda_2^4\left(1 - \cos\left(\frac{\theta_1}{f_2} + \frac{\theta_2}{g_2}\right)\right) \tag{7.3.9}$$

Consider for simplicity $f = f_1 = f_2$. Perfect lattice alignment occurs if in addition $g_1 = g_2$, leading to a flat direction in the potential corresponding to the linear combination $\xi \propto \frac{\theta_1}{g_1} - \frac{\theta_2}{f}$ which is orthogonal to the one appearing in (7.3.9). A slight misalignment $g_1 - g_2 \simeq \epsilon$ gives rise to a nearly flat direction in field space with effective field range parametrized by $f_{eff} = g_2\sqrt{f^2 + g_1^2}/\epsilon$. This effective decay constant can be done a priori parametrically large by sending $\epsilon \to 0$. The problem is that by doing this we are also making the kinetic metric diverge. Let us reformulate the model in terms of the lattice and kinetic basis introduced above. The KNP basis $\theta_i$ is related to the lattice basis as follows

$$\vec{\phi} = \mathbf{M}\vec{\theta}\,, \quad \mathbf{M} = \begin{pmatrix} 1/f & 1/g_1 \\ 1/f & 1/g_2 \end{pmatrix} \tag{7.3.10}$$

In the lattice basis the kinetic term is not canonical and is given by

$$\mathcal{G} = (\mathbf{M}^T)^{-1}\mathbf{M}^{-1} = \frac{1}{(g_1 - g_2)^2}\begin{pmatrix} g_1^2(f^2 + g_2^2)^2 & -g_1 g_2(f^2 + g_1 g_2) \\ -g_1 g_2(f^2 + g_1 g_2) & g_2^2(f^2 + g_2^2)^2 \end{pmatrix} \tag{7.3.11}$$

By diagonalizing $\mathcal{G}$ we can get the kinetic eigenvalues $f_1, f_2$ and the rotation angle $\alpha$ in terms of the initial decay constants $f, g_1$ and $g_2$. Interestingly it turns out that one of the eigenvalues scale as $f_1 \sim 1/\epsilon$ while the other as $f_2 \sim \epsilon$. Hence an effective transplanckian field range ($\epsilon < 1$) is possible only if one of the kinetic eigenvalues is transplanckian. Besides in the limit $\epsilon \to 0$ we also have that $\tan(\alpha) \to 1$ so the rotation angle $\alpha \to \pi/4$ and the system approaches perfect kinetic alignment.

In fact, by diagonalizing the metric and doing the field redefinition explained in (7.3.2) (to go to the kinetic basis) it can be checked that the kinetic eigenvector with the largest eigenvalue corresponds indeed to the flat direction that KNP uses

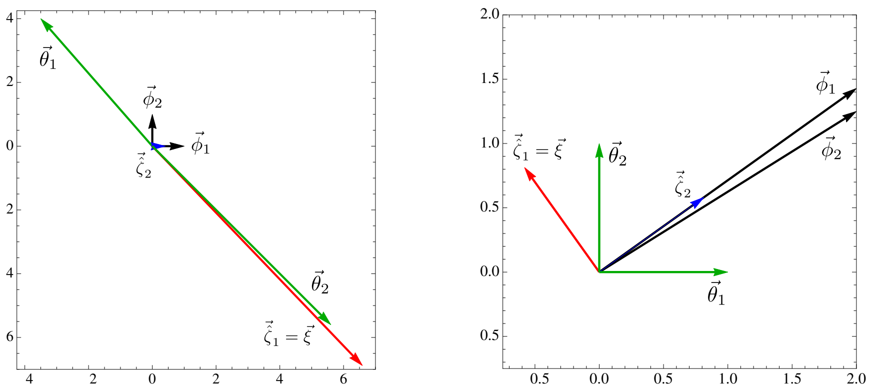


**Figure 7.2:** Different choices of basis in KNP models. The red line always correspond to the inflationary trajectory.

for inflation, ie. $\hat{\zeta}_1 \equiv \xi$. Therefore the maximal displacement is still given by (7.3.7) with the crucial difference that now the eigenvalue $f_1$ is transplanckian.

Figure 7.2 illustrates the relation between the different choices of basis which play a role in the discussion. In the left Figure we show the lattice generated by the potential (basis $\{\phi_1, \phi_2\}$), and the transformation required to go to the KNP basis. In the right Figure we take the KNP basis as canonical and show its relation with the other basis. The degeneracy of the matrix $\mathbf{M}$ in the limit of lattice alignment is pictorially translated into the almost degenerate vectors $\{\phi_1, \phi_2\}$.

To sum up, a priori we could get a parametrically large flat direction by doing $\epsilon \to 0$ which is equivalent to increase parametrically the value of the kinetic eigenvalue $f_1$, so the large scale is hidden in the model from the beginning and is not a consequence of a clever change of basis. The remaining question is then if a transplanckian $f_1$ is consistent in a quantum theory of gravity.

In the following we turn to study the effects of the gravitational instantons over these multiple field inflationary models.

### 7.3.2 Gravitational effects

Let us generalize the computation performed in Section 7.2 for the case of $N$ axions with generic non-canonical kinetic terms. The gauge potential sets the periodicities of the axionic fields, but has no relevant effect in the resolution of the Einstein's gravitational equations as was discussed at the end of Section 7.2.2. Hence we neglect it from now on as was done in the single field case. The stress-energy tensor of the axion system is now

$$T_{\mu\nu} = \partial_\mu \vec{\phi} \mathcal{G} \partial_\nu \vec{\phi} - \frac{1}{2}(\partial_\mu \vec{\phi} \mathcal{G} \partial^\mu \vec{\phi}) g_{\mu\nu} \tag{7.3.12}$$

with nonvanishing components

$$T_{rr} = \partial_r \vec{\phi} \mathcal{G} \partial_r \vec{\phi}, \quad T_{ii} = -\frac{1}{2f(r)} \partial_r \vec{\phi} \mathcal{G} \partial_r \vec{\phi} g_{ii}, \quad i = \psi, \theta, \phi. \tag{7.3.13}$$

Recall that the fields $\phi_i$ are adimensional and define a lattice of periodicities $2\pi$. Dirac quantization then implies

$$\int_{S^3} d\vec{\phi} = \mathcal{G}^{-1}\vec{n} \tag{7.3.14}$$

with $\vec{n}$ some vector of integer entries. This implies that at large distances the axions must have have an asymptotic profile of the form

$$\vec{\phi} \sim \frac{\mathcal{G}^{-1}\vec{n}}{4\pi^2}\frac{1}{r^2} \tag{7.3.15}$$

By solving the Einstein's equations we get

$$f_I(r) = \frac{1}{1-\frac{a}{r^4}}, \quad \partial_r\vec{\phi}\,\mathcal{G}\,\partial_r\vec{\phi} = \frac{\vec{n}^T\mathcal{G}^{-1}\vec{n}}{4\pi^4}\frac{f(r)}{r^6}, \quad a \equiv \frac{\vec{n}^T\mathcal{G}^{-1}\vec{n}}{3\pi^3}G \tag{7.3.16}$$

which reduces to the solution of a single field obtained in Section 7.2 upon setting $\mathcal{G} = f^2$. The action of the gravitational instanton becomes

$$S_E = \frac{\sqrt{3\pi}}{16}M_P\sqrt{\vec{n}^T\mathcal{G}^{-1}\vec{n}} \tag{7.3.17}$$

and induces a potential for the system of axions given by

$$V = \mathcal{P}'e^{-S_E}\cos(n_i\phi_i) \tag{7.3.18}$$

Notice that the fundamental periodicities of the potential generated by the gravitational instantons are the same as those of the original lattice. Thus if there is some vector $\vec{n}$ for which the action $S_E \sim 1$, the corresponding gravitational induced potential will spoil inflation. Since we have encoded all the information about the periodictiies of the physical lattice in $\mathcal{G}$, the effect of gravitational instantons on the realization of large field inflation reduces to see if the "tricks" used to get large displacements in multiple field inflation also imply that $S_E$ is small.

The situation of two axions is depicted in Figure 7.3 for $\alpha = 0$ (left) and $\alpha = \pi/4$ (right), the latter corresponding to the perfect kinetic alignment situation described above. The Figure illustrates the dual lattice, so for instance the vectors $\vec{n}_1, \vec{n}_2$ correspond to the instantons with integer charges (1,0) and (0,1) respectively. The blue ellipse represents the region for which $\tilde{S}_E = \frac{16}{\sqrt{3\pi}}S_E \leq 1$ and the gravitational instantons induce a non-negligible potential for the axions[2]. The axis lengths of the ellipse are determined by twice the eigenvalues of the kinetic matrix $\mathcal{G}$ and it is rotated with respect to the fundamental lattice by an angle given by $\alpha$. The red line represents the would be inflationary trajectory in each case. If the physical lattice is such that some point corresponding to a vector $\vec{n}$ lies inside the ellipse, then the gravitational instantons will induce a potential which spoil inflation if the vector $\vec{n}$ is close enough to the inflationary trajectory. This is not the case if the

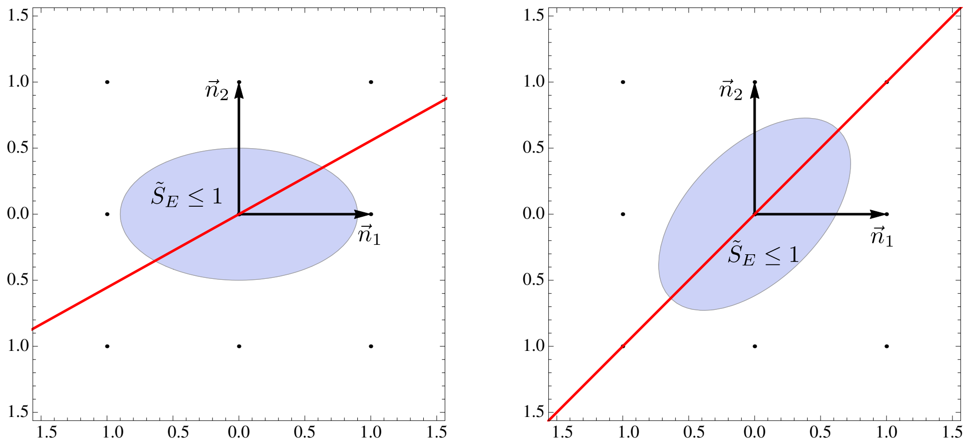


**Figure 7.3:** Instanton action for a model of N-flation with two axions (left) and a model of kinetic alignment (right). The red line corresponds to the direction of the inflationary trajectory.

metric eigenvalues remain subplanckian and therefore no point lies within the ellipse, as can be deduced from fig.7.3 for the cases under consideration. But let us take a closer look to understand how this works in general.

In N-flation or kinetic alignment, all the eigenvalues remain subplanckian and the enhancement is generated by traveling along a diagonal of the field space. For simplicity let us consider two axions in the most favorable situation of perfect kinetic alignment, ie. $\alpha = \pi/4$. The kinetic metric is given by (7.3.8), leading to

$$\vec{n}^T \mathcal{G}^{-1} \vec{n} = \frac{1}{2 f_1^2 f_2^2} \left( f_1^2 (n_1 - n_2)^2 + f_2^2 (n_1 + n_2)^2 \right) \tag{7.3.19}$$

The maximum displacement takes place along the direction of the eigenvector with the largest eigenvalue. For concreteness, let us consider $f_1 > f_2$ so this eigenvector is $\hat{\zeta}_1 = f_1(\phi_1 + \phi_2)/\sqrt{2}$, corresponding to a direction with $n \equiv n_1 = n_2$ in the lattice basis. Along the inflationary trajectory, the action for the gravitational instanton becomes

$$S_E = \frac{\sqrt{3}\pi}{16} M_P \sqrt{\vec{n}^T \mathcal{G}^{-1} \vec{n}} = \frac{\sqrt{3}\pi}{16} M_P \frac{\sqrt{2} n}{f_1} \tag{7.3.20}$$

Notice that this action is a factor $\sqrt{2}$ bigger comparing to the one obtained in the single field case (7.2.14), while the maximum displacement (7.3.7) is a factor $\sqrt{2}$ larger. This result can be generalized to $N$ axions obtaining that the same factor $\sqrt{N}$ which enhances the effective field range, also appears in the instanton action suppressing the amplitude of the induced gravitational potential.

The same results apply for N-flation, where the only difference is that both kinetic and lattice basis are proportional so $\alpha = 0$ and the decay constant entering in $S_E$ is not the largest one but the Pythagorean sum of all of them.

[2] The prefactor $\frac{\sqrt{3}\pi}{16}$ is characteristic of the gravitational instanton under consideration, so it is the same both for one or multiple axions. Hence we plot $\tilde{S}_E$ (the action ommiting $\sqrt{3}\pi/16$) in order to compare directly the constraints for multiple axions with those for a single axion.

The case of lattice alignment is however different. We have seen that a large effective decay constant in the proposal of KNP implies a parametrically large eigenvalue of the kinetic matrix $\mathcal{G}$. Hence the inverse of the metric has a nearly zero eigenvalue and the effects of the gravitational instantons are huge. In more detail, along the kernel direction the action will scale as $S_E \sim \epsilon^2$ being $\epsilon$ the misalignment parameter, so if $\epsilon < 1$ (in order to have a transplanckian effective field range) then the action is automatically $S_E \lesssim 1$. It can be checked that the kernel direction is indeed the nearly flat direction used for inflation, so there is no way out. A vector $\vec{n}$, no matter how big its components, will have nearly vanishing action as long as it is close enough to the kernel direction and will spoil inflation.

The result becomes even more clear by looking at Figure.7.4. In the limit of lattice alignment one of the kinetic eigenvalues tends to zero while the other diverges. This implies that the ellipse is very thin in one axis and very long in the other one, so it captures most of the points with $n_1 = -n_2$ (all of them if there is no misalignment). This implies that the instantons with charges $n_1 = -n_2$ will have a nearly zero action and will induce a huge potential for the axions. In addition, the inflationary trajectory is identified with the direction of the eigenvector with the largest eigenvalue, implying that these instantons indeed generate a potential along the inflationary trajectory and spoil inflation. Each of the points lying within the ellipse can be seen as a higher harmonic contribution which reduces dramatically the field range.

One could hope to engineer a concrete model in which all these contributions are suppressed enough to get successful inflation, but then the advantage of invoking an axion in order to have control over higher order corrections is partially lost. One should really keep trace of all these non-negligible non-perturbative effects, which is clearly not an easy task. In this sense a model based in lattice alignment does not provide any improvement with respect to a model of a single axion.

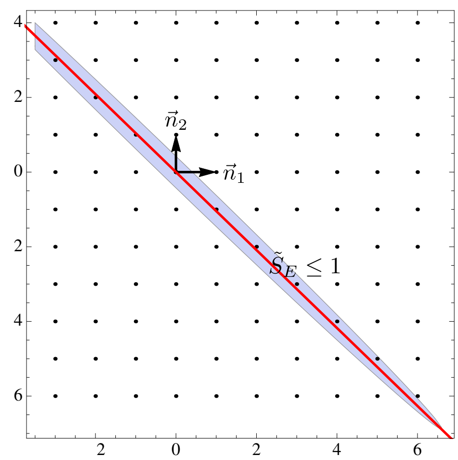


**Figure 7.4:** Gravitational effects for a model of lattice alignment. The blue ellipse represents the region for which the instanton action is small and its effect over the axion can not be neglected. The red line corresponds to the direction of the inflationary trajectory.

These results indicate that having multiple axions might help to relax the gravitational constraints found for the effective field range of a single axion, as long as there are not intrinsic transplanckian scales hidden by any change of basis. To check that, one can always go to the lattice basis such that all the information about the length scales is encoded in the kinetic matrix, and check that the eigenvalues remain subplanckian. If this is the case one can still have an enhancement of the decay constant by traveling along a diagonal of the field space. However this enhancement is given at most by a factor of $\sqrt{N}$, so for $N$ fixed one can never get a parametrically large field range. Even so, one could think that this result is in contradiction with the results of Section 7.2, since one could always integrate out all massive degrees of freedom, except the one corresponding to the inflationary trajectory, ending up again with a single inflationary model with a transplanckian field range but this time without a gravitational instanton to constrain it. Therefore we must be missing some information when integrating out naively the massive degrees of freedom. The solution to this 'paradox' is clarified in Section 7.4, but before that, let us comment about two assumptions we have made in our computation.

First, we have assumed that the fundamental periodicities of the axions are the ones determined by the scalar potential. However, if $V \propto \cos(\phi/f)$, the potential is also periodic under shifts $\Delta\phi = 2\pi m f$ where $m$ can be any positive integer. This is indeed what happens in most of the attempts to embed alignment inflation in string theory, where the induced potential is $V \propto \cos(m\phi/f')$ with $f' = mf$ and $m > 1$. Thus a natural question would be what changes in the study of the gravitational instantons if $m > 1$. The answer is that all of our results still apply in the same way, but we are missing in addition some gravitational instantons which could modify the profile of the potential but not reduce the field range. The generalization of the potential induced by the instantons if $m > 1$ reads

$$V \sim e^{-S_E} \cos\left(\frac{n\phi\, m}{f'}\right) \;, \quad S_E \sim \frac{nM_P\, m}{f'} \tag{7.3.21}$$

The instanton to which we were assigning the unit charge $n = 1$ actually would correspond to that of charge $n' = m$. It is clear from (7.3.21) that by redefining $f' = mf$ we are still capturing the effects of all instantons with $n' \geq m$, but we are missing the information about instantons with $n' < m$. However these latter instantons induce a potential periodic only under bigger shifts $\Delta\phi > 2\pi f$. Hence although they might modulate the inflationary potential in a non negligible way if the instanton action is small enough, they are not relevant for our discussion since they do not reduce the effective field range. This can be easily generalized to more than one axion. The periodicities set by the potential generate a lattice $L$, but they are also compatible with any sublattice $S \subset L$. Since the dual lattice $L^* \subset S^*$ is contained in the dual of $S$, $S^*$, by considering $L$ we are missing some instantons present in $S^*$ but not in $L^*$, which do not respect the periodicity of the potential but instead have a bigger period.

Finally, here we have assumed $M_P$ to be a constant. However, the Planck mass will also receive quantum corrections from diagrams involving the $N$ axions running in the internal loops, so $M_P$ will grow as $\sqrt{N}$, canceling the Pythagorean enhancement from traveling along the diagonal [183, 199]. It has been argued that suitable cancellations can occur such that even if a parametrically large field range is not possible, one might still get a sufficient large field range to provide inflation. Studies regarding the diameters of axion moduli spaces in Calabi-Yau manifolds have been carried out in [165], showing the difficulties arising to have an enhancement of a moduli space diameter while keeping the overall volume small (see also [191] for a more optimistic view). We just want to remark that the gravitational effects which are usually argued to be behind the difficulties from getting transplanckian field ranges in string theory might actually be suppressed for the case of multiple axions. Some of the gravitational instantons which are assumed to be present in the IR, might actually not exist due to the selection rules of some $\mathbb{Z}_N$ discrete gauge symmetries in the UV, as we proceed to explain in the following.

## 7.4 Discrete symmetries in multiple axion systems

We now want to point out what seems to be a paradox between the results of sections 7.2 and 7.3. We will consider a system with two axions $\vec{\phi}$. In the basis in which the lattice is diagonal, we will have a nontrivial kinetic term $\mathcal{G}$. The lagrangian will be

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \vec{\phi})\mathcal{G}(\partial^\nu \vec{\phi}) - \sum_{i=1}^{N} \Lambda_i^4 (1 - \cos(\phi_i)). \tag{7.4.1}$$

We will couple this system minimally to gravity, and also assume $\Lambda_i \ll M_P$. As discussed above, kinetic alignment provides successful inflation. In other words, we take $\Lambda_i = \Lambda$ and

$$\mathcal{G} = \mathbf{R}^T \begin{pmatrix} f^2 & 0 \\ 0 & g^2 \end{pmatrix} \mathbf{R}, \tag{7.4.2}$$

where $\mathbf{R}$ is the rotation matrix which takes the vector $(1, 0)$ to the diagonal $\frac{1}{\sqrt{2}}(1, 1)$. Indeed, if one defines $\vec{\zeta} = \mathbf{R}\vec{\phi}$, then the potential becomes

$$\Lambda^4 \left(2 - \sum_{j=1}^{2} \cos(2\pi R_{ji}\zeta_i)\right) = 2\Lambda^4 \left[1 - \cos(\frac{\zeta_1}{\sqrt{2}})\cos(\frac{\zeta_2}{\sqrt{2}})\right]. \tag{7.4.3}$$

By changing to $\hat{\zeta}_1 = \sqrt{2} f \zeta_1$, $\hat{\zeta}_2 = \sqrt{2} g \zeta_2$, we see that this field theory describes 2 fields with masses $\Lambda^4/g^2$ and $\Lambda^4/f^2$. If $g \ll f$, we may take the latter as an inflaton. The potential along its direction has a periodicity $\sqrt{2} f$.

So far, we have only described again the kinetic alignment mechanism of [188] discussed in the previous section. However, we may now integrate out all degrees of freedom with mass $\Lambda^4/g^2$ and higher, as long as we are away of $\zeta_1 \sim \pi/\sqrt{2}$, which is where inflation ends. We will be left with a theory of a single transplanckian axion which would seem to have an enhanced periodicity $\sqrt{2} f$.

From the results of Section 7.2, we know that such a theory has gravitational instantons contributing potentials of the form $\cos\left(n\frac{\hat{\zeta}_1}{\sqrt{2}f}\right)$, with $n$ any integer. As discussed there, these gravitational instantons reduce the effective field range, constraining $f < M_P$ and thereby spoiling inflation.

However, the gravitational instantons contributing to the action can also be studied before integrating out the heavy axion, and we know from the results of Section 7.3 that in this case inflation with a transplanckian field range is indeed possible. What is going on?

A look at the instanton action (7.3.17) tells us that any instanton not in the direction of $(1,1)$ will pick up a contribution proportional to $1/g$ and hence will be tremendously suppresed. When computing instanton corrections to the path integral, a consistent criterion is to pick all the configurations with action smaller than some pre-stablished value; in this Section we will drop any instantons whose action goes as $1/g$, in accordance with our initial statement that $f \gg g$. Notice however that configurations with actions of order $M_P/f$ will be large (since $f$ is subplanckian) but will be retained nevertheless; otherwise, we would not be able to analyze the potential generated by the gravitational instantons. Also, $M_P/f \gg 1$ is the regime when the instanton throat, etc. are larger than the Planck length and so our effective field theory computations can be trusted.

Let us compute the action of an instanton with $\vec{n} = n(1,1)$:

$$S_E^{\text{instanton}}(n) = \frac{\sqrt{3\pi}}{16}\frac{n\sqrt{2}M_P}{f} = \frac{\sqrt{3\pi}}{16}\frac{M_P}{\sqrt{2}f}(2n). \tag{7.4.4}$$

The instanton generates a potential of the form $\cos\left(2n\frac{\hat{\zeta}_1}{\sqrt{2}f}\right)$. That is, gravitational instantons only generate even harmonics in terms of the periodicity of the effective transplanckian axion established by the gauge potential.

Clearly, the effective field theory of one axion is missing some constraint coming from the UV. A selection rule forbidding odd-order harmonics is highly reminiscent of a discrete symmetry [42, 75, 76, 77]. We will now study the precise way in which the IR theory forbids the presence of these odd charge instantons: it can do so because in the $g \to 0$ limit there is another set of instantons of low action which contribute the path integral, namely, charged axionic strings.

### 7.4.1 The string

Gravity not only enables us to build an instanton for any axion we may have, but it also allows us to build the dual object: a string around which the axion has the asymptotic profile $\vec{\phi} = \theta\vec{m}$, where $\theta$ is the angle around the string. The instanton couples electrically to the axion field, whereas the string has a magnetic coupling of the kind

$$\int_{\text{string}} d\vec{\phi}. \tag{7.4.5}$$

So, we are led to study solutions of the coupled Einstein-axion system with cylindrical symmetry and monodromy $2\pi\vec{m}$ as we circle around the string. Solutions to this system have been considered before in the literature [222, 223, 224, 225]. Asymptotically, the stress energy tensor in cylindrical coordinates is of the form

$$T_\mu^\nu = \frac{C}{r^2}\text{diag}(1,-1,1,1). \tag{7.4.6}$$

It turns out [224] that of the solutions of Minkowskian Einstein's equations corresponding to the above mondromy only one has locally flat asymptotics, and it corresponds to a very unphysical configuration in which there is an event horizon outside which the metric is not static. Physically, what happens is that the constant stress-energy implied by an asymptotic profile with the above monodomy acts as a fluid with larger and larger stresses as we go away from the string, which in turn curves the space at asymptotic regions. To compensate and get a flat solution one must put large negative sources of stress-energy at the core of the string, and these cause the losss of staticity of the solution.

Although it is not possible to have an asymptotically flat Minkowskian metric describing one such string, it may be possible to have one describing two strings with opposite monodromies, so that asymptotically the stress-energy vanishes. This suggests (and this is indeed the case) that although there are no asymptotically flat Minkowskian solutions, there might be asymptotically flat Euclidean ones. Physically, these would describe the nucleation of a string- anti string pair, or one could have closed strings as loop diagrams contributing to vacuum to vacuum amplitudes. The crucial element ensuring the existence of well-behaved asymptotically flat Euclidean solutions when there are no Minkowskian ones is the swap of the sign of the kinetic term of an axion when continued to the euclidean [209]. This was also crucial to ensure the existence of our instantons above.

We thus propose to consider a metric tensor for the string of the form (analytic continuation of [224, 222], whith “wrong” kinetic term sign)

$$ds^2 = \left(1+\frac{1}{\sigma}\log\left(\frac{r}{r_0}\right)\right)dt^2 + dz^2 + \left(1+\frac{1}{\sigma}\log\left(\frac{r}{r_0}\right)\right)^{-1} dr^2 + \mu\sigma r^2 d\theta^2 \tag{7.4.7}$$

The axion string must satisfy $\int_{S^1} d\vec{\phi} = 2\pi\vec{m}$, where the $S^1$ is any $S^1$ which encloses the string. This immediately implies by symmetry that $\partial_\theta\vec{\phi} = \vec{m}$, or $\vec{\phi} = \vec{m}\theta$. Also, $\mu = \vec{m}^T\mathcal{G}\vec{m}$. Notice the branch cut for the axion; this creates a “Dirac domain wall”, similar to the Dirac string in the singular gauge description of monopoles. Instantons crossing this wall should contribute a trivial phase. The parameter $r_0$ can be related to the total length of the Dirac string in some contrived way; we will leave it and study the results as a function of $r_0$.

This metric will describe the near-field limit of a closed string. The Dirac domain wall will therefore be a disc, whose boundary is the actual string (see Figure 7.5).

The metric (7.4.7) has a coordinate singularity at $r_h = r_0\exp(-\sigma)$, but as usual when continuing to the Euclidean the space is much nicer. We must take the

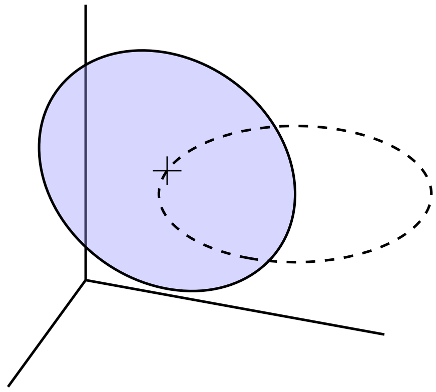

**Figure 7.5:** Pictorial representation of the euclidean string (thick black line) and its associated Dirac domain wall (blue shaded circle). When we move an instanton around the dashed circle, it picks an extra phase as it crosses the domain wall. This phase must be trivial for the theory to be consistent.

time coordinate as periodic in order to avoid a conical singularity [140]. The period of the imaginary time coordinate is $4\pi r_h$, i.e. $2\pi$ times twice the horizon radius, as usual.

On the other hand, the metric we have built for the euclidean string is not asymptotically flat. This is not a problem, as it describes the fields near the core of the string; one expects that an asymptotically flat solution exists for a closed string, and that the above metric is a well approximation near the core of the string. However, this means that for realistic solutions one should impose an upper cutoff $R$ after which the approximation (7.4.7) is no longer good; typically, $R$ will be of order of the radius of the closed string.

We have

$$\frac{1}{2}\sqrt{g}(\partial_\mu\vec{\phi})\mathcal{G}(\partial^\mu\vec{\phi}) = \frac{\mu}{\sqrt{g_{\theta\theta}}} = \frac{\sqrt{\mu}}{2\sqrt{\sigma}r}. \tag{7.4.8}$$

And thus the action per unit length is

$$\frac{S_E^{\text{string}}}{2\pi R} \approx 8\pi^2 r_h \int_{2r_h}^{R} \frac{\sqrt{\mu}dr}{\sqrt{\sigma}r} = \frac{8\pi^2 r_h\sqrt{\mu}}{\sqrt{\sigma}}\log\left(\frac{R}{2r_h}\right) = \frac{8\pi^2 r_h\sqrt{\mu}}{\sqrt{\sigma}}\left[\frac{1}{\sigma} + \log\left(\frac{R}{2r_0}\right)\right]. \tag{7.4.9}$$

The point here is that, as a function of $\sigma$, the action of the string can be made arbitrarily small for $\sigma \to \infty$, which corresponds to vanishing small mass per unit length. So in principle the action of the above instanton can be made arbitrarilly small. A cutoff however is imposed by the fact that $r_h$ should be greater than the Planck length, otherwise the above solution should not be trusted. Since $r_0$ will be of order $R$, this imposes $\sigma < \log(r_0/l_p)$ so we arrive at

$$\frac{S_E^{\text{string}}}{2\pi R} > 8\pi^2\sqrt{\mu}\frac{1}{\sqrt{\log\left(\frac{r_0}{l_p}\right)}}\left[\frac{1}{\log\left(\frac{r_0}{l_p}\right)} + \log\left(\frac{R}{2r_0}\right)\right] \tag{7.4.10}$$

The good thing about the logarithms is that they contribute things of order 1, 2 at most. Also, the radius of the string must be several orders of magnitude larger than the Planck length (i.e, $R \gtrsim l_p$), so we end up with a bound (restoring the Planck mass)

$$S_E^{\text{string}} \gtrsim 16\pi^3\sqrt{\frac{\vec{m}^T\mathcal{G}\vec{m}}{M_p^2}}. \tag{7.4.11}$$

### 7.4.2 Putting everything together

Now that we have the action of the Euclidean strings, we have to ask ourselves under which circumstances do they contribute significantly to the path integral. Recall, as discussed above, that we will neglect contributions to the path integral of action $\sim M_P/g$ but not those which depend on $f$.

With this criterion, instantons with $\vec{n} = (n_1 - n_2)$ with $n_1 + n_2$ have an action which scales as $M_P/g$ and are therefore suppressed. However, their dual strings (say, the string with $\vec{m} = (1,0)$) are not suppressed; from (7.4.11) their action goes as

$$S_E^{\text{string}} \gtrsim 16\pi^3 M_P\sqrt{\frac{f^2(m_1+m_2)^2 + g^2(m_1-m_2)^2}{2}}. \tag{7.4.12}$$

Thus, when we compute the IR effective theory for $\hat{\zeta}_1$, it is not consistent to integrate out $\hat{\zeta}_2$ and retain the gravitational instantons which generate the even-harmonics potential for $\hat{\zeta}_1$. If we want the latter, we must allow for any other configuration which goes as $f/M_P$ in the path integral, such as the strings. The effective theory in the IR is not a single axion of periodicity $\sqrt{2}f$ coupled to gravity; it also contains euclidean strings which change the quantum theory.

These strings constrain the allowed gravitational instantons, forbidding those of odd charge. As discussed above, the strings generate a Dirac domain wall. Consider a configuration with a string with vector $\vec{m}$ and an instanton with $\vec{n}$. If we move the instanton so as to cross the Dirac domain wall, the phase should be (a multiple of) $2\pi$. Actually, the phase the instanton gets upon moving along an $S^1$ with linking number 1 with the string is

$$\int_{S^1} \vec{n}\cdot d\vec{\phi} = 2\pi\,\vec{n}\cdot\vec{m}. \tag{7.4.13}$$

To write this amplitude in terms of IR quantities, we must substitute $\vec{n} = \frac{n}{2}(1,1)$. $n$ could be any integer and still respect the periodicity of the lattice. However, for

$\vec{m} = (1, 0)$ that the above phase becomes just $\pi n$. Consistency of the path integral thus requires that only even-order harmonics contribute.

From the UV point of view, what is happening is clear. The instanton discussed in the previous paragraph would lift to an instanton with $\vec{n} = \frac{1}{2}(1, 1)$. This is not a vector of the instanton lattice, and hence the instanton is not allowed. However, the $\mathbb{Z}_2$ discrete symmetry operating in the IR is harder to see. Notice that the absence of odd harmonics in the potential is only strictly true in the limit $g \to 0$; otherwise we have gravitational instantons breaking the $\mathbb{Z}_2$ symmetry, even if they are very suppressed.

In the limit $g \to 0$, the two antidiagonal corners of the axion lattice are at zero distance from each other. The lattice in which the canonically normalized fields $\vec{\phi}$ live, which is spanned by the vectors $\frac{1}{\sqrt{2}}(f, g)$ and $\frac{1}{\sqrt{2}}(f, -g)$ degenerates to a single straight line generated by $\frac{1}{\sqrt{2}}(f, 0)$ (see Figure 7.6). As a result, an extra lattice point appears on the diagonal of the unit cell. So although it indeed seems that the axion lattice has a diagonal of length $2\pi\sqrt{2}f$, in the degenerate limit the extra lattice point cuts that periodicity by half. As a result, only even-order harmonics are allowed in the potential. In other words, a theory with two axions in which one becomes very massive cannot be reduced to the effective theory of the remaining axion, but rather requires the inclusion of something else, an additional $\mathbb{Z}_2$ symmetry.

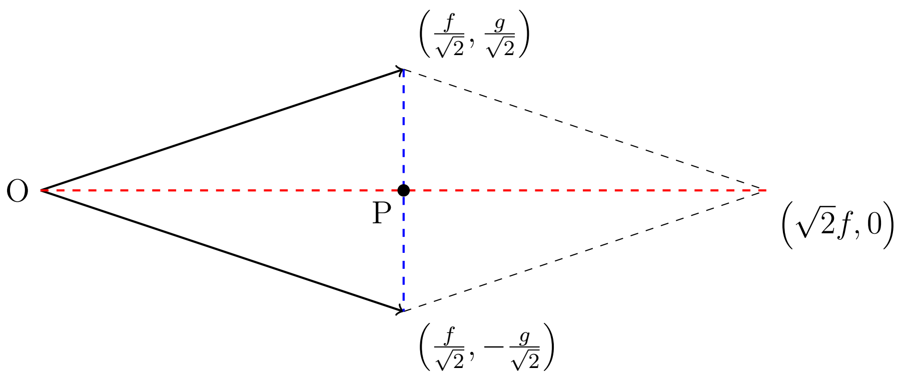


**Figure 7.6:** Basis of the axion lattice with canonically normalized axions. The diameter of the axion lattice is given by the red dashed line. As $g \to 0$, the two elements of the lattice basis move along the blue dashed line to degenerate to the segment $\overline{OP}$, resulting in a one-dimensional lattice with half the periodicity.

If we had actually taken the $g \to 0$ limit the $\mathbb{Z}_2$ gauge symmetry would have been exact and even the gauge potential used for inflation would have had to obey it. The theory would contain a single effective axion of subplanckian periodicity $2\pi f/\sqrt{2}$, and that would be the end of the story. By allowing a small, explicit breaking of the symmetry, we can have a transplanckian axion with highly suppressed gravitational contributions, allowing for tranplanckian inflationary regimes. Speaking loosely, one may degenerate the axion lattice in a precise way so that the gravitational effects "think" that the effective axion is subplanckian, but the gauge effects know it is not.

The above discussion generalizes in a straightforward manner to $N$ axions. In this case, there will be a $\mathbb{Z}_N$ gauge symmetry, which in the UV will come from the degeneration of all the lattice basis vectors to the form $\frac{1}{\sqrt{N}}(f, 0, \ldots)$.

This "loophole", which allows for axions with large decay constants, works because we have assumed the existence of two separate sets of instantons: the gravitational ones (which for low charges will be superseeded by their stringy counterparts related to the WGC, as described in the next section) and the gauge ones, whose action we are taking as a free parameter. This is only possible if the strong forms of the WGC for several $U(1)$'s discussed in the previous Chapter do not hold. This is discussed in detail in [169], but it is easy to see: The gauge potential (7.4.3) must have the right value of $\Lambda$ to result in successful inflation. But this means that the instanton generating it has an action bigger than 1 in units of $M_P/f$. Thus, the "lightest" charged object in the theory is subextremal (in the sense that its action is bigger than that of the gravitational instantons), so that the second strong form does not apply.

However, this does not mean that transplanckian axions are excluded the moment we require any strong forms the WGC. There are several caveats, all of them related to the fact that there is no clear definition of extremal curve for 0-forms (indeed, this is why we avoided discussion of charge-to-mass ratios in this context). For instance, [164] provides a controlled example in which instanton prefactors are smaller than assumed here, and as a result one can have a slightly transplanckian axion. [226] provides a recent example in a controlled setup in string theory where the effects of many instantons are taken into account, resulting in a slightly superplanckian field range. Taking all of this into account, the question of exactly which models of multiple axion inflation allow for transplanckian displacements is still open.

## 7.5 Aspects of string theory realizations

In earlier sections of this Chapter we have considered gravitational instantons in theories of gravity coupled to axions, as part of the effects present in a complete quantum gravitational theory. In this section, we focus on the interplay between our results and expectations in string theory compactifications realizing the axion models described.

### 7.5.1 Generalities

A first observation is that, in order to produce viable single-field inflation models, the string theory compactifications must deal with the challenge of rendering all fields except the axion(s) massive at a hierarchically large scale, so that they do not interfere with the axion inflationary scenario. This typically requires a relatively large breaking of supersymmetry, in order to get rid of the saxion superpartners. Therefore, in our consideration we pay special attention to often-forgotten features arising in the absence of supersymmetry.

In string theory, axions are ubiquitous. For instance they can arise very generically from the KK compactification of higher-dimensional $p$-form fields over $p$-cycles in the internal space (see e.g [227, 228, 229, 230, 231, 232, 233, 234, 235, 236] for such axions as inflaton candidates); they also arise from D-brane positions in toroidal compactifications or in large complex structure limits of Calabi-Yau compactifications (see [237, 238, 239, 240] for such axions as inflaton candidates). For concreteness and for their genericity, we focus on axions coming from the RR $p$-form potentials integrated over $p$-cycles.

In general, many of these axions can get masses from fluxes in general compactifications. In applications to inflation, one assumes that this happens at a high scale[3], so that one is left with a low energy theory with a reduced set of axions.

The non-perturbative contributions to the potential for these axions (the gaugino condensate in phenomenological models like natural inflation or its diverse aligned multi-axion variants) are realized in terms of non-perturbative effects on gauge sectors arising from wrapped D-branes in the model. For instance, many models are based on type IIB theory with axions coming from the RR 4-form on 4-cycles, and the 10d type IIB axion. The relevant gaugino condensate non-perturbative effects arise from gauge sectors on D7-branes on 4-cycles (and possibly with worldvolume gauge backgrounds inducing D3-brane charge coupling to the 10d type IIB axion).

In addition to these field theory non-perturbative effects, there are nonperturbative effects from euclidean D-brane instantons. In general, for any axion in string theory there are such D-brane instanton effects which induce corrections in the 4d effective action [4]. In the above example, of type IIB with axion from the RR 4-form (and the universal axion), the corresponding instantons are euclidean D3-branes wrapped on 4-cycles (possibly magnetized, inducing D(-1)-brane instanton charge, coupling to the universal axion). Actually, the non-perturbative gaugino condensate superpotential can be described in terms of "fractional"[5] D3-brane instantons wrapped on the D7-brane 4-cycle (and with the same worldvolume gauge background). D3-brane instantons which do not have a gauge theory counterpart are dubbed exotic or stringy in the literature [194, 195, 196, 197].

In supersymmetric backgrounds, the kind of superspace interaction induced by euclidean D-brane instanton effects depends on the structure of exact fermion

[3]Actually, is is possible to consider models in which some axion receives a hierarchically small contribution to its potential from the fluxes. In those cases however the structure of the flux potential is such that it induces axion monodromy (e.g. the F-term axion monodromy inflation models [232], see also [233, 234]), for which our analysis does not apply.

[4]Actually, for axions made massive by fluxes, the corresponding instantons are absent due to Freed-Witten consistency conditions [241]; but these are precisely the axions which we have assumed have been integrated out in our effective theories.

[5]The fractionalization of instantons in an $SU(N)$ gauge theory is manifest in the F-theory lift, in which the elliptic fibration over $N$ D7-branes pinches off into $N$ component 2-cycles, each of which can be wrapped by an instanton (described as an M5-brane on the 2-cycle of the elliptic fiber fibered over the base 4-cycle).

zero modes (see [205] for discussion). For instance, superpotential terms (and therefore contributions to the scalar potential) are generated only by instantons with exactly two fermion zero modes (to saturate the $d^2\theta$ measure) [242]. Instantons with additional fermion zero modes induce non-perturbative corrections to higher F-terms (which locally in moduli space can be written like D-terms), namely higher-derivative terms or multi-fermion operators [243]. However, in the presence of supersymmetry breaking effects, these higher F-terms can descend to contributions to the superpotential (or in non-susy language, to the scalar potential), by contracting the additional external legs with insertions of the supersymmetry breaking operators [206]; more microscopically, the susy breaking sources can lift the additional fermion zero modes of the instanton, to allow them to contribute to the scalar potential (see e.g. [244, 245] and references therein). Therefore, a general lesson in non-supersymmetric situations is that fairly generically all D-brane instantons induce corrections to the scalar potential; the only price to pay is the appearance of suppression factors in $M_{\rm susy}/M_s$, where $M_{\rm susy}$ is the susy breaking scale. Since our main interest lies in applicactions to large-field inflation, in which the inflation scale is high and supersymmetry is broken substantially (e.g. to remove the saxion degrees of freedom), our setup is that all D-brane instantons contribute to the potential of the corresponding axions[6]. We will often omit the very model-dependent prefactors depending on $M_{\rm susy}/M_s$.

### 7.5.2 D-brane instantons and gravitational instantons

String theory thus provides a set of non-perturbative objects coupling to the low-energy axions, and inducing non-perturbative contributions to the scalar potential of the form

$$V_{\rm D-inst} \sim \mathcal{P} e^{-S_{\rm D-inst}} \left[\, 1 - \cos(\sum_i n_i \phi_i) \,\right] \tag{7.5.1}$$

where $n_i$ are integer charges under the axions $\phi_i$, normalized to $2\pi$ periodicity, $\mathcal{P}$ is an instanton dependent prefactor, and $S_{\rm D-inst}$ is the D-brane instanton action.

For instance, in the example of axions arising from the RR 4-form (and the universal axion), we can introduce a basis of 4-cycles $\Sigma_i$, $i = 1, \ldots$, and define

$$\phi_i \;=\; \int_{\Sigma_i} C_4 \tag{7.5.2}$$

and let $\phi_0$ denote the universal axion. In this setup, the instanton generating (7.5.1) is an euclidean D3-brane wrapped on the 4-cycle $\sum_i n_i \Sigma_i$, and with a worldvolume gauge background with second Chern class $n_0$.

The close resemblance of (7.5.1) and (7.3.18), suggests the gravitational instantons may correspond to an effective description of non-perturbative effects which are

[6]This is similar to the implicit assumption we are taking for gravitational instantons, which in the presence of supersymmetry may have extra fermion zero modes and contribute to higher-derivative terms [246].

microscopically described in terms of D-brane instantons. This would be much in the spirit of the correspondence between gravitational black hole solutions and systems of D-branes. In the following we provide some hints supporting this interpretation, at the semi-quantitative level (which is presumably the best one can hope for, in these non-supersymmetric setups).

An interesting observation is that BPS euclidean D-brane instantons (in general spacetime dimensions) can be described as solution to the euclidean supergravity equations of motion. These solutions source only the gravitational field and the axion(s) (complexified by a dilaton-like partner), have precisely the axionic charge structure of our solutions, and moreover in the string frame they are described by wormhole geometries, see e.g. [247, 248], see also [249, 250]. It is tantalizing to propose that our wormhole solutions correspond to a low-energy version of the D-brane instanton wormhole solutions, once the dilaton-like scalar partners are removed (by a hypothetical integrating out mechanism once these fields are made massive in the string setup). Notice that in the absence of dilation-like scalars, the string and Einstein frames are equivalent, so the wormhole nature of the D-brane instanton gravity solution becomes more physical.

An interesting hint in this direction comes from the dependence of the instanton actions on $g_s$. Let us consider the single axion case for simplicity. For gravitational instantons, the dependence of the action (7.2.14) on $g_s$ is hidden in the 4d quantities $M_P$ and $f$. To make it explicit, let us consider the case of an axion from the RR 4-form, and perform the KK reduction of the 10d action for gravity and the RR 4-form, which (modulo the familiar self-duality issues) has the structure

$$S_{10d} \sim \frac{1}{\alpha'^4} \int d^{10}x\sqrt{-G}\,(e^{-2\Phi}R + |F_5|^2) \qquad (7.5.3)$$

with $\Phi$ the 10d dilaton. Defining the axion by $C_4 = \phi\,\omega_4$ for some harmonic 4-form, in 4d we get

$$S_{4d} \sim \int d^4x \left( M_P^2 \sqrt{-g}\, R + f^2\, |d\phi|^2 \right) \qquad (7.5.4)$$

with

$$M_P^2 \simeq \frac{V_6}{\alpha'^4 g_s^2} \quad , \quad f^2 \simeq \frac{1}{\alpha'^4} \int_{\mathbb{X}_6} |\omega_4|^2 \qquad (7.5.5)$$

Hence the charge-$n$ gravitaional instanton action has a parametric dependence

$$S_E \sim \frac{M_P\, n}{f} \sim \frac{n}{g_s} \qquad (7.5.6)$$

where in the second equation we have ignored the $\omega_4$-dependent prefactor.

The above expression has the precise parametric dependence expected for a charge-$n$ euclidean D-brane instanton (namely, a D3-brane multiply wrapped $n$ times on the corresponding 4-cycle). We take this agreement as a suggestive indication of the above indicated relation between gravitational instantons and D-brane instantons.

This analysis for the dependence on $g_s$ generalizes straightforwardly to the multi-axion case, using the action (7.3.17). The agreement on the parametric dependence is in general not exact, but holds for large charges as follows. Consider an euclidean D-brane with axion charge vector of the form $\vec{n} = p\vec{n_0}$, i.e. a large multiple of a basic vector $\vec{n_0}$. On general grounds, we expect the D-brane instanton action to scale as $S_{\text{D-inst}} \sim p\, S_0$, with $S_0$ the action of the D-brane instanton associated to $\vec{n}_0$. The parametric dependence on $p$ is straightforwardly reproduced from (7.3.17).

In some particularly simple instances, the action of D-brane instantons agrees with the precise dependence of the charges in (7.3.17), as we describe in a particular setup in the next section.

### 7.5.3 Charge dependence in flat multi-axion setups

In some very specific models, the specific dependence of the gravitational instanton action (7.3.17) on the charges of the axion lattice can be shown to survive the embedding into string theory.

Consider an euclidean $(p-1)$-brane wrapping a $p$-cycle $\mathcal{C}$ of the compactification manifold. Even if the brane is not BPS, one may estimate the action of the instanton via the usual Dirac-Born-Infeld action. In the particular case there are no gauge or $B$-fields turned on the brane worldvolume, this is just

$$S_{DBI} \approx \frac{\mu_p}{g_s} \int_{\mathcal{C}} dV_p = \frac{\mu_p}{g_s}\, V_p, \tag{7.5.7}$$

proportional to the volume of the $p$-cycle. We will now see that, under some (admittedly very specific) circumstances, the action (7.5.7) has a dependence on the charges of the axion lattice similar to that of (7.3.17).

In order to do this, it will be convenient to have an alternative expression for the volume element $dV_p$ of the brane. Let us assume that a globally defined $(n-p)$ form exists such that $*\omega_{n-p} = \lambda dV_p$, where $\lambda$ is a constant and $n$ is the dimension of the ambient space (typically $n = 6$, although we may also consider submanifolds of the compactification manifold if we like). That is, $*\omega_{n-p}$ is proportional to the volume form of the $p$-cycle. Using the condition $dV_p \wedge *dV_p = dV_n$, one gets $\lambda = \sqrt{*(\omega_{n-p} \wedge *\omega_{n-p})}$, and thus we may write

$$dV_p = \frac{*\omega_{n-p}}{\sqrt{*(\omega_{n-p} \wedge *\omega_{n-p})}}. \tag{7.5.8}$$

The usefulness of the above expression depends largely on the possibility of finding a suitable $\omega_{n-p}$. This is a hard thing to do in general, since $\omega_{n-p}$ cannot be guaranteed to exist, but it is a particularly simple thing to do for tori. In what follows, we will take the ambient space to be $T^n$, the $n$-dimensional torus, and the $p$-cycle will be a flat $p$-dimensional hiperplane.

As discussed above, an easy way to obtain axions in the effective theory coming from this string compactification is via the zero modes of RR $p$-forms. Namely, one expands

$$C_p = \sum \phi_i \beta_i \tag{7.5.9}$$

with $\{\beta_i\}$ being the harmonic representatives of a basis of $H_p(T^n, \mathbb{Z})$. Grouping the axions into a vector $\vec{\phi}$, their kinetic term will be of the form $\vec{\phi}^T \mathcal{G} \vec{\phi}$, with

$$\mathcal{G}_{ij} = \langle \beta_i, \beta_j \rangle = \int_{T^n} \beta_i \wedge *\beta_j. \tag{7.5.10}$$

The instantons coupled to these axions will be euclidean $(p-1)$-branes wrapped on cycles of $T^n$. $H^p(T^n, \mathbb{Z})$ and $H_{n-p}(T^n, \mathbb{Z})$ are naturally isomorphic via Poincaré duality. There is a single flat hyperplane in each class of $H^p(T^n, \mathbb{Z})$, and these may be expanded as $\mathcal{C} = \sum n_i \tilde{\alpha}_i$, where $\{\tilde{\alpha}_i\}$ is the basis of $H^p(T^n, \mathbb{Z})$ dual to $\{\beta_i\}$. Then

$$\omega_{n-p} = \sum n_i \alpha_i \tag{7.5.11}$$

satisfies that $*\omega_{n-p} \propto dV_p$. Here, $\alpha_i$ is the (unique) covariantly constant representative of the image of $\tilde{\alpha}_i$ via Poincaré duality. Furthermore, the $\alpha_i$ are harmonic. This means that the action of the instantons is, in this case

$$\begin{aligned} S_{DBI} &\approx \int_{\mathcal{C}} \frac{*\omega_{n-p}}{\sqrt{*(\omega_{n-p} \wedge *\omega_{n-p})}} = \int_{T^n} \frac{*\omega_{n-p} \wedge \omega_p}{\sqrt{*(\omega_{n-p} \wedge *\omega_{n-p})}} \\ &= \sqrt{V_{T^n}} \sqrt{\int_{T^n} \omega_{n-p} \wedge *\omega_{n-p}}, \end{aligned} \tag{7.5.12}$$

where the volume factor arises because $*(\omega_{n-p} \wedge *\omega_{n-p})$ is constant, and hence

$$*(\omega_{n-p} \wedge *\omega_{n-p}) = \frac{1}{V} \int_{T^n} \omega_{n-p} \wedge *\omega_{n-p}. \tag{7.5.13}$$

However, we now have

$$\int_{T^n} \omega_{n-p} \wedge *\omega_{n-p} = \mathcal{G}^{-1}_{ij} n^i n^j, \tag{7.5.14}$$

since $\langle \alpha_i, \alpha_j \rangle$ is the inverse matrix to $\langle \beta_i, \beta_j \rangle$ by virtue of the duality between the two bases.

In this way, the dependence of the gravitational instanton action on its axionic charges is reproduced in the stringy setup. Strictly speaking however this result only holds for a torus. Exact computation of the action of euclidean non-BPS D-brane instantons in a generic Calabi-Yau manifold is a challenging matter into which we will not delve.

One expects the gravitational instanton solution to be approximately valid when the charges are large and the curvature at the throat of the wormhole is low. If we fix a particular $p$-cycle and wrap $n$ branes around it, the action (7.5.12) scales with $n$, as it did in the single-axion case, and at the end of the previous section.

Notice that to recover the precise matching of the charge dependence it is crucial that the internal geometric objects are actually constant, so that the integrands are constant and integration amount to picking up volume factors. In this sense, the gravitational instanton can be regarded in the general case as an average assuming the internal geometric information is truncated to constant pieces.

### 7.5.4 D-brane instantons and alignment

Let us now apply the D-brane instanton viewpoint to the analysis of string theory realizations of transplanckian alignment models. In particular, we now show that string realization of the lattice alignment models described in Section 7.3 necessarily contain D-brane instantons which spoil the naive transplanckian inflaton field range allowed by the gauge non-perturbative axion potentials. In other words, the D-brane instantons take up the job carried out by the gravitational instantons in the description in earlier sections.

Given the close relation between the axion charges of D-brane instanton and gravitational instanton, the analysis is simple. For a lattice alignment model, we recast it in the form shown in Figure 7.4, and consider D-brane instantons corresponding to the charges $(n_1, n_2) = p(1, -1)$, which in fact is of the form considered above for D-brane instantons in multi-axion setups. This effectively corresponds to a single-axion model with a transplanckian decay constant $f$, so that the D-brane instanton action

$$S_{\rm D-inst} \sim \frac{p}{g_s} \sim \frac{p\, M_P}{f} \tag{7.5.15}$$

becomes small for any point inside the ellipse.

It is illustrative to consider a concrete setup. Several attempts to embed these alignment models in string theory are based on considering two stacks of D7-branes wrapped (possibly a multiple number of times $n$) on homologous 4-cycles (i.e. in the same class $[\Sigma_4]$), and carrying worldvolume gauge backgrounds with slightly different instanton numbers, $k$, $k'$. In terms of the homology class of the 4-cycle and the class of the point, the non-perturbative effects on these gauge sectors correspond to D-brane instantons with charge vectors $\vec{v}_1 = (n, k)$, $\vec{v}_2 = (n, k')$. Actually, we are interested in considering the lattice generated by these vectors, namely $n_1\vec{v}_1 + n_2\vec{v}_2$, i.e. we use the potential to define the periodicities of the axions (using the more refined lattice defined by the homology amounts to including more instanton effects, but these do not reduce the field range, as explaine at the end of Section 7.3.2). The relevant D-brane instantons are given by $n_1 = -n_2$, namely are of the form $p(\vec{v}_1 - \vec{v}_2)$ which in terms of the underlying homology lattice corresponds to $(0, p(k - k'))$. In other words, the D-brane instanton is, at the level of charges, a superposition of $p$ D3-branes on $\Sigma_4$, each with gauge instanton number $k$, and $p$ anti-D3-brane on $\Sigma_4$ with worldvolume instanton number $k'$. The lowest-energy state in this charge sector is a set of $p(k - k')$ D$(-1)$-brane instantons, whose action is essentially given by $p(k - k')M_P/f$.

The bottomline is that in string theory D-brane instantons provide the microscopic description of the effects which have been interpreted in terms of gravitational instantons in the effective theory in the earlier sections. As in the case with higher $p$-form gauge fields, one may relate the existence of these small instantons with the WGC (the connection becomes clearest when one uses dualities to transform the axions into $p$-form gauge fields, $p > 0$).

# 8
# Monodromic relaxation

In this Chapter we continue to apply the WGC to derive interesting phenomenological consequences. Here we will focus on a different problem from inflation, but where too cosmology and large field ranges for axions are relevant: The relaxion proposal [251] to resolve the hierarchy problem.

We will start with a brief discussion of the hierarchy problem and of the proposed relaxion solution. We will argue that single-field relaxion models immediately lie in the Swampland, unless they are examples of what is called axion monodromy models. We will describe these models and their nice properties. Monodromy models include a 3-form in the spectrum, to which one can apply the WGC. This turns out to have interesting consequences in the relaxion scenario, providing stringent constraints in the parameter space of the model. Lastly, some ideas towards an embedding of the relaxion in string theory are discussed.

The results of this Chapter are published in [252].

## 8.1 Relaxions and the hierarchy problem

The electroweak hierarchy problem (see [253] among many others) is perhaps one of the most famous fine-tuning questions in physics, leaving aside perhaps the cosmological constant problem. To pose it, let us regard the SM as an effective field theory valid until some cutoff scale, $\Lambda$. The fact that the SM lagrangian observed experimentally is renormalizable suggests strongly that this scale $\Lambda$ must be somewhat above the electroweak (EW) scale of around 100 GeV. This is so because one generically expects the physics above the cutoff scale to induce non-renormalizable operators in the effective field theory, suppressed by some power of $\Lambda$. Since the effects of such operators have not been detected conclusively so far, their suppression cannot be too small.

Physics at the cutoff however not only enters the low energy effective field theory by inducing non-renormalizable operators: it also generates contributions to renormalizable ones, which are however not observable because they should be renormalized to match the theory to experimental data. In spite of this, the structure of these corrections can give us hints about the physics at the cutoff scale.

For instance, the physical Higgs quartic coupling or the Yukawa couplings in the SM are logarithmic, depending on the cutoff only via $\log(\Lambda)$. In this way, we get values for these parameters of roughly the same order of magnitude, no matter what the value of the cutoff is.

The case for the $m^2$ of the Higgs boson is different, since it receives quadratic corrections, proportional to $\Lambda^2$. Thus, if we take a large cutoff, say of the order of the Planck scale, we will get a quantum correction of $10^{19}$ GeV which will have to cancel up to 18 significant digits with the bare contribution to give the observed mass of 126 GeV. Equivalently, we would expect the theory at the cutoff scale to produce parameter values which are "natural", by which we mean of the same order of magnitude than the appropriate power of the cutoff. This is clearly not the case for the Higgs mass, if the cutoff is very large, and the resulting fine-tuning problem is called the hierarchy problem.

The hierarchy problem is not a problem from the effective field theory point of view: There is no inconsistency in taking the bare mass to cancel the quantum corrections to any desired precision. It is a problem for the UV completion, which has to explain this extreme cancellation. For instance, typical string compactifications with the gauge structure and Higgs mechanism of the SM generically do not solve the hierarchy problem, unless some extra ingredients such as a low scale of supersymmetry breaking or a Randall-Sundrum mechanism are assumed. Therefore, the hierarchy problem is giving the string phenomenologist a clue, namely to look for atypical compactifications which can accommodate these large hierarchies.

Many solutions to the hierarchy problem, such as low-scale supersymmetry, can be posed purely in field-theoretic terms. Recently [251] (see also [254, 255, 256, 257, 258, 259, 260, 261, 262, 263] for later versions) proposed a new mechanism to solve the EW hierarchy problem under the name of "cosmological relaxation". Its main appeal is that it does not require the presence of new physics near the EW scale, while providing at the same time a natural dynamical mechanism to keep the Higgs hierarchically lighter than the cutoff of the theory. Basically the proposal is to extend the SM Higgs scalar potential by including a coupling to an axion field, leading to a potential

$$V = V(g\phi) + (-M^2 + g\phi)|h|^2 + \Lambda^4 \cos\left(\frac{\phi}{f}\right). \tag{8.1.1}$$

Here $V(g\phi) = gM^2\phi + g^2\phi^2 + \ldots$, and $\Lambda = \Lambda_0 + \Lambda(h)$ depends on the vev of the Higgs field $h$. Also, $f$ is the usual axion decay constant, and $M^2$ is a cutoff coming from SM loop effects.

In the minimal version the field $\phi$ is the QCD axion, the cosine potential arises from the usual $SU(3)$ instanton effects breaking the Peccei-Quinn symmetry, and $\Lambda_{QCD} \sim \Lambda(h = v)$ depends on the Higgs vev through quark masses. We will however also consider more general axion-like particles to play the relaxion role.

During inflation, the non-perturbative effects are negligible, hence the dynamics of $\phi$ is controlled by $V(g\phi)$; the axion $\phi$ starts out at a large positive value and

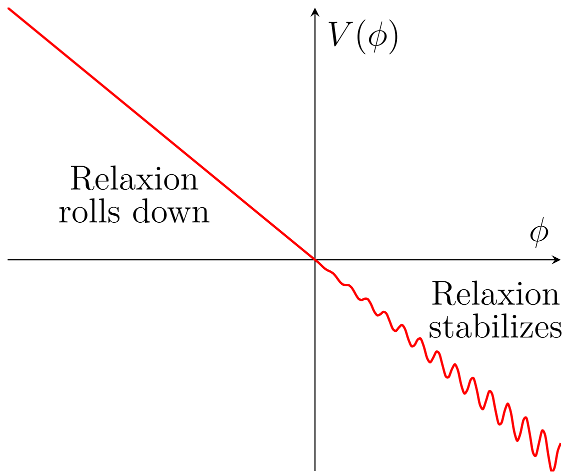


**Figure 8.1:** Pictorial representation of the relaxion mechanism. At early times, the relaxion $\phi$ rolls down its potential, until $\equiv (-M^2 + g\phi)$ vanishes and triggers EW symmetry breaking (in the plot this happens where the two axes meet). This in turn switches on nonperturbative effects which stabilize the relaxion.

slow rolls down its potential, thus scanning values for the Higgs mass. When crossing $m_h \sim 0$, namely at $\phi \sim M^2/g$, the Higgs develops a vev, triggering electroweak symmetry breaking, and the barrier of the cosine potential increases stabilizing the axion shortly after $m_h \sim 0$, as illustrated in Figure 8.1. Hence the Higgs mass is dynamically set to a value much lower than the cutoff $M$. In order for the instanton term to stop the rolling of $\phi$, the barrier $\Lambda(h)$ evaluated at $h = v$ should be comparable to the slope of the axion potential. Parametrizing $\Lambda(h)^4 = ch^2$, this requires

$$gM^2 \sim \frac{\Lambda^4(h = v)}{f} \quad \longrightarrow \quad g \sim \frac{cv^2}{fM^2} \tag{8.1.2}$$

For the relaxion being the QCD axion, $c = f_\pi^2 y_u^2$, and $f > 10^9$ GeV according to astrophysical bounds, leading to a very small coupling

$$g \sim 10^{-16} \frac{m_{\rm EW}^2}{M^2} \text{ GeV}. \tag{8.1.3}$$

Therefore a big hierarchy between the cutoff $M$ and the EW scale is translated into a very small coupling $g$. In fact, the smallness of this parameter is common in all the versions of the cosmological relaxation mechanism, with $g \sim 10^{-34}$ GeV being a typical value.

This mechanism to generate a hierarchically small Higgs mass is argued to be technically natural, since the smallness of $m_h$ comes from the smallness of the parameters $g$ and $\Lambda$, which are associated to symmetry breakings. Indeed, the parameter $g$ is the only source of breaking of the global shift symmetry of the axion, and therefore its smallness is expected to be technically natural.

There are two important questions unanswered by the above description:

- The smallness of $g$ implies a field excursion $\Delta\phi$ during inflation much larger than the UV cutoff of the theory. This possibly endangers the stability of the potential against higher dimensional operators, a familiar issue in large field inflation models (see [175] for a review). One may argue that this problem is solved by appealing to the continuous perturbative axion shift symmetry. However, given the general belief that quantum gravity violates all global symmetries, this mechanism seems unrealizable in actual embedding of this effective theory in UV completions including quantum gravity.

- On the other hand, for $\phi$ to describe an axion, it should have a discrete periodic identification under $\phi \to \phi + 2\pi f$. As emphasized in [258], this should correspond to a gauge symmetry in a consistent theory of quantum gravity [1]. This is however not respected by the coupling to the Higgs field in (8.1.1), implying that $\phi$ can not be a pseudo-Nambu-Goldstone boson (pNGB).

It is interesting that these two questions have already been addressed in the context of axion monodromy inflation models [227, 228, 264, 265, 232, 234] (see also [229, 266, 231, 233, 238, 237, 267, 235, 236, 268, 239, 240, 269, 270, 271, 272, 273, 274]), and this motivates the suggestion to UV-complete relaxion models in this framework. In this Chapter we take several important steps towards fleshing out this proposal, and analyzing the problems of embedding it into a consistent quantum theory of gravity like string theory. Among other things, we will study the constraints that the Weak Gravity Conjecture implies in the viability of the models.

## 8.2 Axion monodromy

In [258] the authors argue that the discrete shift symmetry of $\phi$ has to be necessarily gauged in a consistent quantum theory of gravity and therefore can not be broken by any term in the action. This implies that if $\phi$ is an axion or a pseudo-Nambu-Goldostone boson (pNGB), the coupling $g$ is not naturally small but indeed theoretically inconsistent. However, the authors in [258] miss the possibility that $\phi$ is not a pNGB but an xion with multi-branched potential, so that the theory is consistent with a mass term and interactions for the axions while preserving an underlying discrete shift symmetry (see e.g. [227, 228, 264, 234, 232, 265], also [229, 266, 231, 233, 238, 237, 267, 235, 236, 268, 239, 240, 269, 270, 271, 272, 273] for applications to inflationary potentials). Our present work is the first concrete proposal to implement a monodromy realization of relaxion models, and explore its implications.

In this Section we briefly review axion monodromy models, explaining the mechanism by which periodic scalars get multi-branched potentials from the introduction of a coupling to a 3-form field. It also serves to fix notation and conventions.

[1] One could try to drop the requirement that $\phi$ is an axion, by declaring it to take values in $\mathbb{R}$ rather than in $\mathbb{S}^1$. This however makes the PQ-symmetry group a ( nonlinearly realized) $\mathbb{R}$, instead of the usual $\mathbb{S}^1 = U(1)$. Noncompact symmetries are again notoriously in conflict with quantum gravity [42].

A more detailed discussion of the monodromy mechanism would take us too far away from the topic and is thus postponed to Appendix D.

As described in [264] (see [275] for related ideas in a different context), an efficient way to describe the introduction of potential terms for axionic scalars is to couple them to a 3-form gauge field. Consider for instance the simplest case, which eventually describes a massive axion. It corresponds to the lagrangian

$$L = -\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}|F_4|^2 + g\phi F_4, \tag{8.2.1}$$

where $F_4 = dC_3$ the field strength of the 3-form. Since the 3-form field has no propagating degrees of freedom in 4d, we can integrate it out via its equation of motion

$$* F_4 = F_0 + g\phi, \tag{8.2.2}$$

leading to an induced scalar potential for the axion

$$V_{KS} = \frac{1}{2}(F_0 + g\phi)^2. \tag{8.2.3}$$

Notice that even if the 3-form in four dimensions does not have propagating degrees of freedom, it can still yield a non-vanishing field strength giving a positive contribution $F_0$ to the vacuum energy. The discrete identification of the scalar is a gauge symmetry which involves a change in $F_0$, as follows

$$\phi \to \phi + 2\pi f \quad , \quad F_0 \to F_0 - 2\pi g f \tag{8.2.4}$$

At the quantum level[2], the vacuum value of the 4-form flux $F_0$ is quantized in units of membrane charge (we will come back to these membranes in Section 8.4)

$$F_0 = n\Lambda_k^2 \quad , \quad n \in \mathbb{Z} \tag{8.2.5}$$

Hence we have the following consistency condition [278]

$$2\pi f g = k\Lambda_k^2 \quad , \quad k \in \mathbb{Z} \tag{8.2.6}$$

which will be important when discussing explicit relaxion monodromy models.

This structure underlies the axion monodromy inflationary models (see e.g. [227, 228, 264, 265, 232, 234]), in which the scalar potential is multivalued with a multibranched structure dictated by the above discrete shift symmetry, akin to the "repeated zone scheme" familiar from solid-state physics [278, 279] (see Figure 8.2 for a qualitative picture). Each branch is labelled by the value of $F_0$. Once a specific

[2]At the classical level, $F_0$ can take an arbitrary constant value implying that the continuous shift of the axion is also a symmetry of the action. However, as emphasized in [276] and in Appendix D the actual value of the 4-form field strength in four dimensions (and not only its shift when crossing a membrane) satisfies Dirac quantization. When embedding the model in string theory, this quantized value indeed corresponds to the integer flux of the magnetic dual in higher dimensions [276, 277]

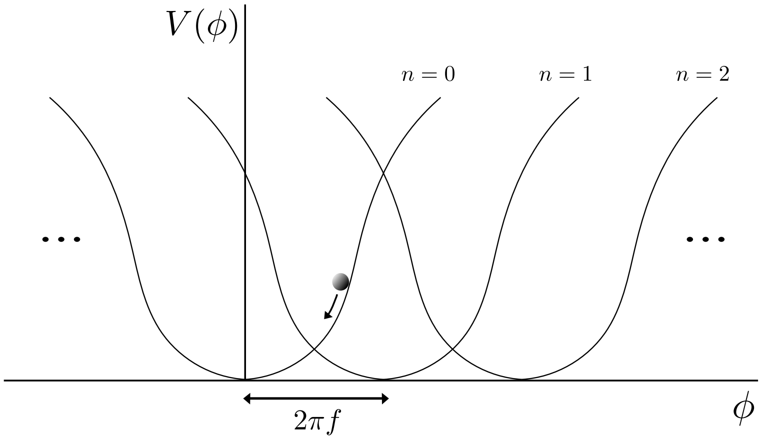


**Figure 8.2:** Multi-branched structure of a typical axion monodromy model.

branch is chosen, one can go up in the potential away from the minimum and travel a distance $\Delta\phi$ larger than the fundamental periodicity $f$. This is specially useful for large field inflationary models in which one needs a transplanckian field excursion for the inflation even if all the scales of the theory remain sub-Planckian. The relation between F-term axion monodromy and a Kaloper-Sorbo (KS) potential like (8.2.1) was explicitly shown in [232], and further generalized in [277] for any axion of a given string compactification.

The above system can also be generalized to multiple axions

$$L = K_{ij} d\phi^i \wedge *d\phi^j - F_4 \wedge *F_4 + C_3 \wedge j \tag{8.2.7}$$

where $j_\mu = \partial_\mu \theta(\phi^i)$ is an external current satisfying $dj = 0$ and $\theta(\phi_i)$ can be a polynomial function mixing the axions. This leads to more generic potentials than (8.2.3) containing also quartic or higher couplings between the axions. The system is then invariant under a set of discrete transformations in the axions combined with some integer shifts on the parameters appearing in $\theta(\phi_i)$ (which in string theory correspond to the internal fluxes of the compactification). The relation (8.2.6) then becomes

$$\delta\theta(\phi) = k\Lambda_k^2 \quad , \quad k \in \mathbb{Z} \tag{8.2.8}$$

where $\delta\theta(\phi)$ is the integer shift that has to be reabsorbed by each 4-form background. This is indeed the situation that naturally arises in string theory flux compactifications [277], in which the full axionic dependence of the scalar potential can be written in terms of 4d couplings of the 4-forms.

The above system of a single axion admits an alternative description in which the scalar is dualized into a 2-form. The 3-form then gets massive by "eating up" the 2-form in a gauge invariant consistent way [232]. We review this (standard) description in Appendix D. The description in terms of 2-forms is valid for the simple monodromic models studied in this Chapter, but has not appeared in the literature for more general monodromic models such as those in [277].

The appealing feature of this mechanism is that the gauge invariance of the 3-form protects the potential from Planck-suppressed operators. More concretely, consider higher-dimensional operators which appear as powers of the gauge invariant field strength $F_4/M^2$, with $M$ the cut-off scale (or the Planck mass in monodromy inflation). After integrating out the 3-form the corrections to the scalar potential will also appear as powers of the leading order potential itself, namely $\delta V \sim \sum_n V_0(V_0/M^4)^n$, so they will be subleading as long as the potential remains below the cut-off scale, even if the field takes large values. Therefore it is a very efficient mechanism to keep a scalar field light, in a way consistent with interactions, and without adding new degrees of freedom or new physics at the EW scale.

This is another motivation to construct a relaxation model by rewriting all the couplings in terms of 3-form fields. The parameter $g$ of relaxation could be safely argued to remain naturally small even for a large field excursion of the axion.

## 8.3 A minimal relaxion monodromy model

As mentioned in Section 8.1, monodromic or multi-branched potentials have been suggested as a way out of certain puzzling features of the naive relaxion models. However, there is actually no explicit relaxion model with a built-in monodromy structure in the literature. We no describe the simplest relaxation model with an axion monodromy structure. We also revise the issue of the smallness of $g$ in the context of monodromy. Clearly, the construction admits many generalizations, and the model in this Section is proposed as a simple illustration. In this sense, we emphasize that the analysis in forthcoming sections is actually meant for general constructions, rather than the precise model in this section.

### 8.3.1 Seesaw-like scales and stability

The scales in axion monodromy show an interesting see-saw structure which has not been pointed out in the literature.

As discussed in the previous section, requiring that an axion potential has a monodromic structure amounts to requiring that it is invariant under a discrete axion shift symmetry which acts non-trivially on the parameters of the potential. In the previous Section we saw in (8.2.6) that

$$g \;=\; k\frac{\Lambda_k^2}{2\pi f}\ . \qquad (8.3.1)$$

So we see that $g$ is quantized in units of $\Lambda_k^2/(2\pi f)$. The structure in (8.3.1) is reminiscent of the seesaw mechanism for neutrino masses (for a review, see e.g. [280]), where the very small neutrino mass arises as a quotient $m_W^2/M_\nu$ between two different mass scales; as long as $M_\nu \gg m_W$ (typically $M_\nu \sim 10^{10} - 10^{16}$ GeV in the neutrino context), the resulting neutrino mass will be much lower than both $m_W$, $M_\nu$.

Similarly, (8.3.1) can explain one of the troubling aspects of the relaxation mechanism, namely the tiny value of $g$. This is typically required to be as small as e.g. $10^{-34}$ GeV, so that, even if relaxation works, we seemingly must introduce a new energy scale far smaller than any other in an unmotivated way. However, the smallness of $g$ is nicely explained by the seesaw mechanism in (8.3.1). For $f \simeq 10^{10}$ GeV one obtains for the flux scale $\Lambda_k$:

$$g = 10^{-34}\ \text{GeV} \longrightarrow \Lambda_k \ \simeq\ \sqrt{gf} \ \simeq\ 10^{-3}\ \text{eV}\ . \tag{8.3.2}$$

So for typical relaxionic values the new fundamental scale is rather $\Lambda_k \simeq 10^{-3}$ eV, and the smallness of $g$ appears as a derived very small quantity determined by the effective see-saw relation $g \simeq \Lambda_k^2/f$. This beautifully follows from the monodromy version of the axion symmetry, via the discrete shift symmetry and the quantization of $F_0$. The new fundamental scale $\Lambda_k$ is much larger than $g$, by about 20 orders of magnitude.

Before moving on we should notice an interesting numerical coincidence in this minimal example. The new scale $\Lambda_k \simeq 10^{-3}$ eV is in the order of magnitude to the observed cosmological dark energy scale $\Lambda_{\text{dark}} \simeq (10^{-3}\text{eV})^4$. It is tantalizing to speculate that both scales are physically related, perhaps through the intimate relation between (free) 4-forms and their contributions to the cosmological constant [281, 282, 283, 276].

In any case, the smallness of $\Lambda_k$ is yet to be explained. A perhaps interesting observation along this line is that, by combining (8.1.2) and (8.3.2), we arrive at

$$\Lambda_k \sim \frac{\Lambda_v^2}{M}, \tag{8.3.3}$$

This resembles a further seesaw between the nonperturbative scale $\Lambda_v$ and the SM cutoff scale $M$. However, in absence of any direct coupling between QCD and $F_4$, (8.3.3) remains accidental. Indeed, (8.1.2) means that the nonperturbative barriers are able to stop the relaxion, and thus constitute a purely phenomenological requirement for the relaxion picture to work.

This see-saw structure can provide an explanation for the originally tiny value of $g$. But one should also address the question of the stability of this parameter, both at the classical level (taking into account non-negligible higher dimensional operators due to the large field excursion of the relaxion) and at the quantum level due to loop corrections. The former makes reference to the infinite tower of non-renormalizable operators which a priori become relevant when the field takes values larger than the cut-off of the theory (as required in relaxation), while the latter refers to the quantum stability of the classical Lagrangian. The argument for which $g$ is technically natural since it is associated to a symmetry breaking is not valid in the context of monodromy, or at least, the underlying protection is more subtle, because indeed the discrete shift symmetry of the axion remains unbroken at the level of the action (so $g$ is not associated to the breaking of the discrete shift symmetry).

However, monodromy provides a new mechanism to guarantee the stability of the effective potential. First, as explained in the previous Section and in Appendix D, the gauge invariance of the 3-form shields the potential against non-renormalizable higher dimensional operators, implying that those should come as powers of the potential itself. Therefore they will remain subleading (due to the original smallness of $g$) even if the field excursion of the relaxion is bigger than the cut-off scale. Let us remark that the stability of the full scalar potential can only be guaranteed if the complete perturbative potential for the axion (including mass terms and interactions) arises from a coupling to a single 3-form field. Therefore, not only $V(g\phi)$ has to be rewritten in terms of a coupling to a 3-form field, but also the axion-Higgs coupling term. This will be the subject of Section 8.3.2.

In addition, the stability of the potential at the quantum level appears as a natural consequence of the previous argument, because all the classical perturbative couplings involving the relaxion $\phi$ are then controlled by $g$. Therefore the quantum corrections will only give rise to a renormalization of the parameter proportional to itself, implying that $g$ is technically natural and will remain small if was originally small.

### 8.3.2 Coupling a multi-branched axion to the SM

Let us start by recalling the minimal relaxion model (8.1.1),

$$V = V(g\phi) + (-M^2 + g\phi)|h|^2 + \Lambda^4 \cos\left(\frac{\phi}{f}\right). \tag{8.3.4}$$

The simplest option for a monodromy invariant version of this coupling one could think of is

$$V \;=\; V_{SM} \;+\; V_{KS} \;-\eta F_4|H|^2 \;+\; V_{cos} \tag{8.3.5}$$

with $\eta$ some constant of order one and

$$V_{KS} = \frac{1}{2}|F_4|^2 - g\phi F_4 \tag{8.3.6}$$

One obtains, after eliminating $F_4$

$$\begin{aligned} V \;&=\; -\mu^2|H|^2 \;+\; \lambda|H|^4 \;+\; (F_0 \;+\; g\phi \;+\; \eta|H|^2)^2 \;+\; V_{cos} = && (8.3.7)\\ &=\; \tilde{\lambda}|H|^4 \;+\; (F_0 \;+\; g\phi)^2 \;+\; 2\eta(-M^2 \;+\; g\phi)|H|^2 \;+\; V_{cos}, && (8.3.8)\end{aligned}$$

where we define

$$M^2 \;=\; \frac{\mu^2}{2\eta} \;-\; F_0\;. \tag{8.3.9}$$

An important point to remark is that $\mu$ is of order the UV cut-off, includes all loops, and is not quantized in general.Then $F_0$ is not required to be large, and no enormous quanta are required. It only has to shift appropriately and will generically be of order $F_0 \simeq \Lambda_k^2$.

The structure just obtained is similar to the relaxion model, with a certain KS-like potential for the axion. Now, as the axion rolls down, it meets first the point at which $\phi = M^2/g$ as usual. But still goes down because the KS potential has not yet reached its minimum. But then eventually the cosine piece enters into the game and stops the relaxation, as usual.

This minimal monodromy version of relaxation improves the original in several respects. The gauge shift symmetry is preserved, even though the axion has a non-trivial quadratic potential. This symmetry also protects the relaxion from Planck suppressed corrections which appear in powers of $F_4^2/M^4$. Furthermore, the minute mass scale $g$ appears as derived from a much larger fundamental scale $\Lambda_k \simeq 10^{-3}$ eV via the relationship $g \simeq \Lambda_k^2/f$.

It is clear that many other models may be constructed by introducing different couplings of the 4-form strength to the Higgs field. Also the values of the mass scales involved, the cut-off $M$ and the 4-form scale $\Lambda_k$ may be very different from the ones in this simple example. As an example, in Section 8.6 we show a model derived from a string setting in which the coupling of $F_4$ to the Higgs system is quite different.

We refrain from entering a detailed model-building search, and instead turn to the interesting question of whether the general idea of a monodromy relaxion is viable, by studying new model-independent constraints. In particular, the multi-branched structure of the axion potential implies the existence of tunneling between vacua corresponding to different axion branches. We have to explore whether this tunneling is sufficiently suppressed so that the relaxion indeed proceeds smoothly through slow roll to reach the point in which it induces a massless Higgs. We study these issues in the next two sections, in particular exploiting the implications of the Weak Gravity Conjecture.

## 8.4 Membrane nucleation and the Weak Gravity Conjecture

Axion monodromy models, in particular those involving a large number of axion windings along a branch, must address the question of possible tunneling transitions between branches (mediated by membrane nucleation) which can reduce the effective field range of the axion (see e.g. [264, 265, 232, 235] for discussion in the axion inflation setup). However a quantitative estimate of the tunneling probability requires information from the UV completion, which determines the tension of the corresponding membrane. As we argue in Appendix D, in theories containing quantum gravity, a version of the WGC can provide useful information on the parametric dependence of the membrane tensions, in a fairly model-independent way, and therefore can yield generic constraints on relaxion models.

### 8.4.1 Membranes and monodromy

In a generic relaxion model, an immediate question is how large a hierarchy we can obtain between $\Delta\phi$, the field range traversed by the relaxion during inflation, and $f$, the relaxion decay constant. Can one really restrict to a single branch and go up the potential to make the field range parametrically large? The answer is no, in general. As is well-known in the monodromy literature [232], in the presence of membranes there are dynamical processes which make the field jump from one branch to a lower one, spoiling the slow rolling [3]. However, as for any non-perturbative tunneling process, the probability for this to happen is in principle exponentially suppressed (see [278, 235] for some discussion in the context of axion monodromy inflation).

Let us describe the process in more detail. During inflation, and the rolling of the relaxion, a bubble may nucleate, bounded by a membrane. The value $F_0$ of the 4-form will jump by $\Lambda_k^2$ upon crossing the membrane. Since the vacuum energy is lower within the bubble than outside, the bubble will expand indefinitely, provided that it is initially large enough so that the pressure associated to the difference in vacuum energies beats the surface tension. The smallest bubble for which this happens is the so-called critical bubble.

Bubbles smaller than the critical radius cost energy to produce, since the surface tension overcomes volume. As a result, they cannot be produced in the vacuum. The critical radius bubble costs no energy, and hence it can be produced by an instanton effect. One can estimate the transition rate for this process in the thin wall approximation. There is a well known expression for this

$$P \sim \exp\left(-B\right), \quad B = \frac{27\pi^2 T^4}{2(\Delta V)^3} \tag{8.4.1}$$

where $T$ is the tension of the membrane which induces a shift on $F_0$ and $\Delta V$ is the variation in the potential energy, in our case

$$\Delta V = V_i - V_f \sim \Lambda_k^2 g\phi. \tag{8.4.2}$$

This formula however neglects gravitational effects. As discussed in [218], it is only valid when the gravitational backreaction of the energy density in the bubble can be ignored. This will be the case whenever the bubble radius[4] $r \sim T/(\Delta V)$ is smaller than the de Sitter radius $H^{-1}$ associated to the energy density of the bubble. In other words, gravitational effects are negligible if

$$q \equiv \frac{1}{rH} \sim \frac{\Delta V}{TH} > 1 \tag{8.4.3}$$

where the variable $q$ parametrizes the importance of the gravitational effects and will be useful later. We will see in Section 8.5.2 that relaxionic models correspond to precisely the opposite regime. For typical parameters, $q = (rH)^{-1} \leq 1$ and gravitational effects are significant.

We therefore need to use the more general expression for vacuum decay which can be found in [218], see Appendix E for details. It turns out that the approximate

[3]There are other processes which may spoil too large windings of the axion, see e.g. [265].

[4]Notice that $r \sim T/(\Delta V)$ is the radius that the bubble would have in flat space. When gravitational effects are important, the expression for the bubble radius is modified, as explained in Appendix E.

expression valid in the relaxation regime is

$$B \approx w(q)\frac{2\pi^2 T}{H^3}. \tag{8.4.4}$$

Here, $w$ is a certain function of $q$ defined in Appendix E and ranging from one to $\approx 0.1$ in the relaxation regime. Notice that unlike (8.4.1), (8.4.4) depends strongly on the Hubble constant $H$, signalling the importance of gravitational effects.

Formulae like (8.4.1) or (8.4.4) are useless without additional information, because they give the tunneling probability in terms of the membrane tension $T$, which cannot be constrained from just effective field theory. It is here where the WGC can be helpful, allowing us to constrain the model even if we do not know the exact UV completion. If the WGC predicts the existence of a membrane with very small $T$, we have $B \ll 1$. Even though we would be outside of the semiclassical approximation inherent to any instanton computation, a small value of $B$ would generically mean that membranes are produced copiously, since their nucleation is not suppressed by exponential effects. Since inflation lasts for so long in relaxation models, we would produce enough membranes so as to completely spoil the slow-roll of the relaxion. Generically, if a strong enough form of the WGC holds, we can use it as a tool to discern which effective field theories might be a priori UV completed when including gravity and which ones would be in the Swampland instead.

### 8.4.2 WGC and membranes

We now briefly discuss the application of the WGC to 3-form fields, particularly in the context of relaxion monodromy models[5].

As discussed in Chapter 6, important subtleties arise when trying to apply the WGC to $p = d - 1$, where the corresponding objects are domain walls, as noted by [147]. Let us particularize to $p = 3$, $d = 4$, for which the relevant object is a membrane in 4d, separating vacua with different value of 4-form flux, $F_4 \to F_4 + \Lambda_k^2$, and (since a three-form field is nondynamical [281, 282, 283, 276, 275, 277]) thus different cosmological constant. In General Relativity, a solution describing a black membrane with nonvanishing cosmological constant $\Lambda$ is [284]

$$ds^2 = -Hdt^2 + H^{-1}dr^2 + r^2(dx^2 + dy^2), \quad H \equiv -\frac{2m_T}{r} - \frac{1}{3}\Lambda r^2. \tag{8.4.5}$$

where $m_T$ is a measure of the brane tension, which should be taken positive. Since $r$ is always timelike, this solution describes a time dependent background which is not asymptotically flat. Furthermore, the solution only has a single coordinate horizon, at $r^3 = -\frac{6m_T}{\Lambda}$, so for $\Lambda > 0$, the only coordinate horizon is behind the singularity at

[5]One could also try to apply this conjecture to axion monodromy models in inflation (see [157] for partial attempts). However, the constrains derived in that case are too weak and do not pose a serious problem for inflation. It is the particular interplay between the scales in relaxation what makes these constraints relevant.

$r = 0$. In a sense, the membrane is always superextremal and hence the validity of its effective field theory description is questionable.

This fits with the naive extrapolation of the extremality condition for dilatonic membranes obtained in [147]. For $p = d - 1$ the extremality condition is given by

$$\left[\frac{\alpha^2}{2} - \frac{d-1}{d-2}\right] T^2 \leq \frac{g^2 q^2}{G_N}. \tag{8.4.6}$$

Here $\alpha$ measures the strength of the dilaton coupling. We see that for $\alpha = 0$ the equation has no solution; the membranes are always superextremal.

From (8.4.6), it seems that a non-trivial dilaton profile can solve the pathologies of the membrane solution (as is the case for domain walls solutions in string theory). Alternatively, consistency of the membrane solution is sensitive to the particular matter coupled to gravity in effective field theory. This sensitivity prevents us from constructing a generic black membrane and studying its decay to formulate a WGC; for some theories it might well be that no nonpathological membranes can be constructed at all.

In string theory, there is another way to go around these issues and derive a WGC for membranes: one may use a chain of dualities, such as $T$ duality, which change the codimension of objects charged under $p$-form fields. This approach was pioneered in [157], which used a chain of dualities to derive a WGC for 0-form fields (producing bounds on instanton actions), which also escapes the naive extrapolation of the $p$-form WGC.

This will also work for $(d-1)$-forms fields coming from RR fields in string theory. In this case, the membranes will be $D$-branes wrapping some cycle of the compactification manifold. If we perform a $T$-duality along some circle direction, the $(d-1)$ form will become a $(d-2)$ form. Charged objects in this case are cosmic strings which, like the domain walls we were considering, cannot have flat asymptotics either. If we perform another $T$-duality along another circle direction, the $(d-2)$ form will turn to a $(d-3)$-form. Charged objects are $(d-3)$ branes, for which an extremality condition and WGC bound can be defined without further issue. Now we can $T$-dualize back. In this way, the original WGC for $(d-3)$ forms can be rewritten in terms of T-dual to give a formulation of WGC for $(d-1)$-forms.

To sum up, although the effective field theory arguments which support the WGC for $p < d - 2$ cannot be applied straightforwardly to domain walls, these difficulties seem due to the high codimension of the charged object preventing a smooth solution within effective field theory. As soon as one includes a (strong enough) dilaton coupling or allows for stringy dualities, $p = d - 1$ does not seem to be very different from the other cases. We conclude that the WGC is plausible (up to order 1 factors) in the 3-form case case as well. Furthermore, we will assume the validity of the Lattice WGC discussed in Chapter 6, so that we have a superextremal membrane of unit charge. As mentioned in Chapter 6, although the black hole arguments above only motivate the mild form of the conjecture, no counterexamples to the strong form are known in string theory. In fact, the membranes predicted by the WGC are present in any stringy model we can think of.

In terms of 3-form data, the gauge coupling $g_3$ is none other than $\Lambda_k^2$, the 4-form flux quantum, which as argued in Section 8.2 equals $2\pi fg$ in monodromic relaxion models. The strong form of the WGC applied to 3-form fields in four dimensions implies that the membrane which shifts the flux $F_0$ by one unit must have a tension lower than

$$T \sim 2\pi fgM_P \tag{8.4.7}$$

where we have restored the Planck mass, $f$ is the decay constant of the axion, and there might be additional $\mathcal{O}(1)$ factors. The tunneling probability associated to membrane nucleation is given by eq. (8.4.4)

$$B \approx 4\pi^3 w(q)\frac{fgM_P}{H^3}. \tag{8.4.8}$$

One can now take a particular effective relaxation model and compute the tunneling probability between different branches. If the tunneling is not suppressed, many such transitions will take place over the exceedingly large number of e-folds that inflation lasts in the original relaxion proposal, spoiling the simple slow roll picture of the relaxion dynamics.

This mechanism will not stop with the appearance of the QCD barriers, since periodic potentials coming from nonperturbative effects have (approximately) the same value at $\phi$ and $\phi + 2\pi f$; only the monodromic part of the energy changes. As a result, $\Delta V$ remains the same and so does the transition rate.

The WGC membrane is a kind of domain wall different from the field-theoretic "bounces" which arise once the QCD barriers switch on (for a study of the implications of these field-theoretic domain walls in the context of axion monodromy inflation, see [285]). These are fully accounted for by low-energy effective field theory, have a tension $T \sim f\Lambda_v^2$ (where $\Lambda_v$ is of order of the strong coupling scale of the gauge group coupled to the relaxion) in relaxionic models, and are already discussed in [251]. Their decay rate [217, 218] can be safely computed within effective field theory and is not an issue for the relaxion proposal over most of its parameter space. These field-theoretic membranes can appear as "holes" on the surface of the 3-form WGC membranes we have been discussing, as explained in Appendix D. On the other hand, the effect of the membranes we consider cannot be computed within effective field theory, and we need additional input like the explicit value of $T$, or the WGC bounds on it.

## 8.5 Constraints on the relaxion

When studying the viability of relaxation in a UV completion using monodromy there are two issues to address. First, the stability of the full scalar potential can only be guaranteed if the complete potential for the relaxion (including also the axion-Higgs coupling term) is rewritten in terms of a coupling to a 3-form field. This has been the subject of Section 8.3 from a pure effective field theory point of view. First attempts to find such a structure in string theory will be given in Section 8.6.

Secondly, if one indeed succeeds in constructing such a model, one should study the tunneling probability by membrane nucleation before making any claim about the effective field range available to relaxation.

A priori it makes no sense to address the second issue without having a complete realization of monodromic relaxation, since one needs the tension of the membrane (and more generally, the action of the instanton) to be able to make any concrete claim. However, it turns out that the typical scales involved in the theory to address the EW hierarchy problem are in general so extreme that we can already draw some conclusions just focusing on the pure relaxion potential $V(g\phi)$. Taking $g$ to be the 3-form coupling, one can apply the formulae in the previous Section to constrain specific relaxion models.

### 8.5.1 Constraints on the relaxion parameter space

The original relaxion model [251] has a parameter space specified by the cutoff $M$, the axion coupling $g$, the axion decay constant $f$, the Hubble scale $H$ during inflation, and the energy scale $\Lambda$ of the gauge group providing the nonperturbative effects which stop the rolling of the potential. If we want the relaxion to also solve the strong CP problem, then $f$ and $\Lambda$ are constrained to have their QCD values. This parameter space is however very constrained. We now review these constraints as established in [251].

In order for the relaxation mechanism to provide a dynamical solution to the EW hierarchy problem for generic initial conditions of the relaxion $\phi$, inflation must last long enough for $\phi$ to scan the entire range of the Higgs mass. This implies that

$$\Delta\phi \gtrsim M^2/g \tag{8.5.1}$$

leading to a lower bound in the number of efolds

$$N \gtrsim \frac{H^2}{g^2} \tag{8.5.2}$$

about $N_e > 10^{37} - 10^{67}$, depending on the specific details of the model and further constrains on inflation. The relaxion will stop rolling when the barrier of the non-perturbative potential becomes comparable to the slope of the perturbative potential, i.e.

$$gM^2 \sim \frac{\Lambda_v^4}{f} \tag{8.5.3}$$

where $\Lambda_v^4 \equiv \Lambda^4(h = v) = cv^2$ being $c = f_\pi^2 y_u^2$ if $\phi$ is the QCD axion. If we require $M \gg m_W$ (to have a EW hierarchy problem to solve), the relation (8.5.3) implies an upper bound

$$g \ll \frac{\Lambda_v^4}{f m_W^2} \tag{8.5.4}$$

For the minimal field content in which the relaxion is the QCD axion, the above bound reads $g \ll 10^{-16}$ GeV, where we have used $\Lambda_v \sim \Lambda_{QCD} \sim 200$ MeV and $f \sim 10^9$ GeV. For instance, a cutoff $M \sim 10^7$ GeV implies $g \sim 10^{-26}$ GeV. Combining eqs.(8.5.1) and (8.5.3) we get

$$\frac{\Delta\phi}{f} \gg \frac{M^4}{\Lambda_v^4} \sim \frac{M^4}{cv^2} \tag{8.5.5}$$

Therefore a big hierarchy between the EW scale and the cutoff of the theory implies in turn an even bigger hierarchy between the axionic field range and the fundamental periodicity $f$. Notice that this is independent of whether the relaxion is the QCD axion. In terms of the monodromic model, this implies that the axion must travel its fundamental domain an extremely huge number of times, between at least $10^4 - 10^{40}$ times for a cutoff $M \sim 10^4 - 10^{13}$ GeV. It is then reasonable to expect a non-negligible tunneling probability so that it is more efficient to decrease the energy by jumping from one branch to another than by slowly rolling down the potential. In the next section, we explicitly compute this probability by plugging the above constraints in the formulae derived for the transition rate in the previous section. But before that, let us recall that the Hubble constant is also highly constrained in the relaxion models. For the QCD-like barriers to form, we must have

$$H < \Lambda_v. \tag{8.5.6}$$

Also, energy density during inflation must be dominated by the inflaton, rather than the relaxion. As a result

$$H > \frac{M^2}{M_P}. \tag{8.5.7}$$

This already imposes an upper bound on the cutoff given by $M \lesssim (\Lambda_v M_p)^{1/2} \sim 10^9$ GeV for $\Lambda_v$ taking the QCD value. This constraint can even be a bit stronger if we also impose that quantum fluctuactions are subleading with respect to the classical rolling, leading to $M \lesssim 10^7$ GeV .

### 8.5.2 WGC constraints

The above constraints were already discussed in the original relaxion paper. We want to study whether membrane nucleation provides extra constraints on the relaxion parameter space.

If the transition rate is not exponentially suppressed, we can expect a large number of membranes being produced during inflation. If this number is larger than the number of times the relaxion winds its fundamental domain, $\approx 10^4 - 10^{40}$, membranes will efficiently take $\phi$ to 0, thus spoiling the solution to the hierarchy problem. Therefore, in order to have successful relaxation we must have $P \sim \exp(-B) \ll 1$.

Inserting the relations between $f, g, H$ and $M$ found in the previous Section (given by eq.(8.5.3)) into (8.4.8) and using $\phi \sim M^2/g$, we get

$$B \approx 4\pi^3 w(q) q^3 \frac{\Lambda_v^4 M_P}{M^2} \tag{8.5.8}$$

The variable $q$ was defined in (8.4.3) and parametrizes the importance of the gravitational effects. If $q = 1$ we recover the flat-space formula (8.4.1). In particular, we have

$$q = \frac{\Delta V}{HT} = \frac{g\phi}{HM_p} \sim \left(\frac{\phi}{M^2/g}\right)\frac{M^2/M_p}{H} < 1 \tag{8.5.9}$$

which is always smaller than one at the end of relaxation, since for relaxion models satisfying $\phi \sim M^2/g$, the value of $q$ is controlled by the ratio between the density energy for relaxion and inflation. Therefore gravitational effects may play a role suppressing the value of $B$ in (8.5.11) and increasing the tunneling probability. Since $H < \Lambda_v$ (see (8.5.6)) we also have a lower bound for $q$ given by

$$\frac{M^2}{M_p\Lambda_v} < q < 1. \tag{8.5.10}$$

Using (8.5.9), and taking into account $\phi \sim M^2/g$ during relaxation, we may rewrite (8.5.8) as

$$B \approx 4\pi^3 w(q)q^3\left(\frac{\Lambda_v M_P}{M^2}\right)^4. \tag{8.5.11}$$

Requiring a suppressed tunneling probability with $P \sim \exp(-B) \ll 1$, yields a constraint

$$M \leq (4\pi^2 w(q)q^3)^{1/8}\sqrt{\Lambda_v M_P} \sim (w(q)q^3)^{1/8}\ 2.5\cdot 10^9\ GeV \tag{8.5.12}$$

where we have used $\Lambda_v \approx \Lambda_{QCD}$. This constraint for $M$, based only on the tunneling rate, is slightly weaker than those already discussed in [251], unless $q \ll 1$. However, it can be checked that due to the lower bound for $q$ in (8.5.10) it is not possible to get a stronger constraint for the cutoff than $M < 10^9$ GeV.

The bounds above pose no problem for relaxionic models. However, we still have to take into account that in relaxion models the inflation period is extremely large and the number of efolds $N$ enormous. Thus we have to consider that bubbles can form at any moment at any point in the expanding universe. In this connection note first that, as shown in Appendix E, the radius of a bubble changes with time as

$$r(t) = \sqrt{\frac{r_{min}^2}{\cosh^2(\sqrt{\Lambda}t)} + \frac{1}{\Lambda}\tanh^2(\sqrt{\Lambda}t)} \tag{8.5.13}$$

with $\Lambda \approx H^2$ being the associated cosmological constant. This shows that, within a few Hubble times, the bubble reaches the cosmological horizon at $r = H^{-1}$. Intuitively, the bubble expands at almost the speed of light, and it takes a time of order $H^{-1}$ to reach the horizon. This is true for bubbles nucleated at any time; although later bubbles start with exponentially suppressed $r_{min}$ to account for the expansion of spacetime, they always become of cosmological size in one e-fold. Therefore the exponential expansion does not wash out the effects of the bubble.

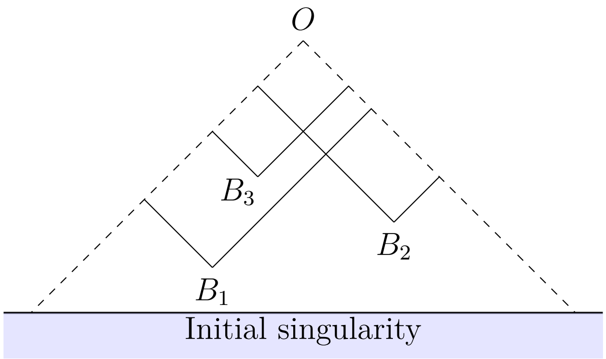


**Figure 8.3:** Conformal diagram showing the past light cone of some observer $O$.Since the bubbles $B_i$ expand so fast, approximately at light speed, we will be within any bubble nucleated in the past light cone.

On the other hand, although bubble nucleation is suppressed, inflation lasts an enormous amount of time in relaxionic models. In principle, many membranes can be produced. The average number of bubbles nucleated in a spacetime region $\mathcal{R}$ of de Sitter space must be

$$N_b \sim \mathrm{Vol}(\mathcal{R}) e^{-B} \tag{8.5.14}$$

where we have not written an equality because of an instanton prefactor whose computation requires an ultraviolet completion of the theory. Typically it is set by some scale of the theory, so we might think of it as a function of $\Lambda, f, g, M \ldots$. Ultimately, the precise value of this factor will not be relevant for our conclusions (as long as it is nonzero of course).

The bottomline of the discussion is that if $\mathcal{R}$ is our past light cone during inflation, then its spacetime volume scales as $\mathrm{Vol}(\mathcal{R}) \sim e^{3N}$ where $N$ is the number of e-folds. Since bubbles grow up to the Hubble radius within one e-fold (and actually start pretty close to it), by the end of inflation we will be within roughly $N_b$ bubbles[6]. This is illustrated in Figure 8.3.

For bubble nucleation not to spoil relaxation, we must require $N_b \ll 1$. Parametrically, this is equivalent to the requirement $N < B$. Since in relaxation $N \sim (H/g)^2$, one has

$$\frac{B}{N} \sim 4\pi^3 w(q) q^5 \frac{\Lambda_v^{12} M_P^6}{f^2 M^{16}} \tag{8.5.15}$$

which sets a new constraint for relaxionic models,

$$M \lesssim \left( \sqrt{4\pi^3 w(q) q^5 \frac{\Lambda_v^6 M_P^3}{f}} \right)^{\frac{1}{8}} \simeq (w(q)q^5)^{1/16}\ 300\ \mathrm{TeV} \tag{8.5.16}$$

[6]A similar argument would not work for the field-theoretic domain walls of tension $T \sim \Lambda_v^2 f$ discussed in the previous section, since these can only appear in the final stages of inflation when the QCD barriers become significant.

where we have used typical QCD values $\Lambda_v \simeq 10^{-1}$ GeV and $f \simeq 10^9$ GeV in the last step. In this case, any hierarchy between the Hubble scale and $M^2/M_P$ (leading to $q$ smaller than one) would scale down this bound and can even rule out any possible hierarchy between the cutoff $M$ and the EW scale.

This constraint is more restrictive than the constraints discussed in the original relaxion papers, and it applies to any relaxation model, independent of the nature of the relaxion or the origin of the scale $\Lambda_v$. Let us remark that the only constraints coming from the relaxion proposal that we have used are: the fact that $\phi \sim M^2/g$ (necessary to get a light effective Higgs mass), the relation (8.5.3) derived from imposing that the non-perturbative barrier eventually stops the running of $\phi$, and the requirement that the energy density is dominated by inflation instead of relaxion, leading to (8.5.7). Hence the result is quite independent of the specific relaxion proposal, as long as it satisfies the three previous conditions and has a built-in monodromy structure.

For instance, the minimal relaxion model described in Section 8.3 includes a new parameter $\eta$ parametrising the coupling of the Higgs field to the Minkowski 4-form. The constraints for relaxation are modified so that the bound in the cut-off becomes

$$M \lesssim \left(\sqrt{4\pi^3 w(q) q^5}\,\frac{\Lambda_v^6 M_P^3}{f}\right)^{\frac{1}{8}} \sqrt{\eta}. \tag{8.5.17}$$

One could a priori think of relaxing the constraint on the cut-off by having an extremely large $\eta$. However this leads the theory out of perturbative control, since $\eta^2$ is a coefficient for the Higgs quartic coupling.

In Figure 8.4 (right) we plot the dependence of the transition rate on the cut-off scale $M$ and the Hubble scale $H$ (which enters in the formulae through the dependence on $q$). In the left Figure we plot the constraints directly in terms of $q$, to show explicetely the relevance of the gravitational effects. The blue and orange regions correspond to the allowed regime with $B > 1$ and $B > N$ respectively. The constraints on the Hubble scale (yielding (8.5.10)) put an additional constraint on the cut-off represented by the green region. This latter constraint is the one already appearing in [251]. Unlike $B > 1$, the constraint $B > N$ motivated above, indeed leads to stronger bounds than those already discussed in [251]. In particular, the bigger is $H$, the stronger are the constraints. Therefore, depending on the exact value of the Hubble constant, cosmological relaxation could still be used to explain the Little hierarchy problem (with $M < 300$ TeV) or instead be totally ruled out.

Let us also finally comment that the minimal version of cosmological relaxation is already ruled out by the strong CP problem. The minimization of the axionic potential implies

$$\theta_{QCD} \sim \langle\phi\rangle \sim gM^2 f^2/\Lambda_v^4 \sim \mathcal{O}(f) \tag{8.5.18}$$

while the experimental constraint imposes $\theta_{QCD} < 10^{-10}$. There are mainly two proposals to solve this problem still keeping the relaxion to be the QCD axion. Either the slope of the perturbative axionic potential decreases dynamically after

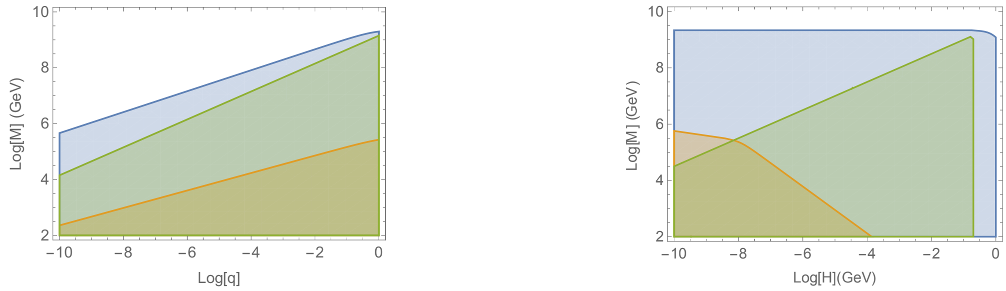


**Figure 8.4:** Plot of constraints $B > 1$ (blue) and $B > N$ (brown) including the $w(q)$ dependence. The blue and brown zones correspond to the regions surviving the tunneling constraints in the parameter space spanned by the cut-off scale $M$ and the parameter $q \simeq M^2/(HM_p)$ (left) or the Hubble scale (right). Only $B > N$ leads to stronger constraints on $M$ than those already coming from eqs.(8.5.6) and (8.5.7), plotted in green.

inflation [251], or the non-perturbative potential induced by QCD instantons is suppressed during relaxation and grow afterwards to force $\theta_{QCD}$ to small values [259]. In the former case, the constraint (8.5.16) is modified to

$$M \lesssim \left(\sqrt{4\pi^3 w(q) q^5}\,\frac{\Lambda_v^6 M_P^3}{f}\,\theta_{QCD}^{\frac{9}{4}}\right)^{\frac{1}{8}} \approx 500\ GeV. \tag{8.5.19}$$

In this case the bound of $M$ is so restrictive that there is no hierarchy problem to solve: Bubble nucleation rules out the original relaxation proposal. We stress however that this bounds only apply to the particular solution of the strong CP problem in [251]; other solutions, such as the one in [259], might still be viable.

The third option is to consider an additional new strong group coupled to the Higgs which stops $\phi$ from rolling. The scale $\Lambda_v$ can be then increased until a few hundred GeV [251], so the tunneling probability is further suppressed and the constraint for the cut-off is relaxed to $M \lesssim 10^8$ GeV, again times the cosmological suppression factor $(w(q)q^5)^{1/16}$. These latter constraints are comparable (perhaps slightly stronger) to the ones obtained in [251].

## 8.6 Monodromy relaxions and string theory

Many of the ingredients of the relaxion models discussed in previous sections are present in string theory. Periodic axion-like fields appear in all string compactifications, and their shift symmetries are typically remnants of gauge invariance of higher dimensional antisymmetric fields (see [10] for review). In particular in Type II orientifolds axions appear from expanding D=10 Ramond-Ramond (RR) antisymmetric fields $C_{MN..Q}$ on harmonics over the 6D compact space, see e.g.[199]. They also appear upon dimensional reduction of Neveu-Schwarz(NS) antisymmetric 2-forms $B_{MN}$. The field strengths of these antisymmetric fields as well as the magnetic fluxes of $F_{MN}$ regular gauge fields are quantized when integrated over the compact space. All these integer quanta are inherent discrete degrees of freedom for any string compactification. This leads to a landscape of string vacua for any given particular compact space, which has been argued to be at the root of the understanding of the smallness of the value of the cosmological constant [276].

As we mentioned above, it has been realized that in large classes of Type IIA/IIB string compactifications the axions may have non-trivial perturbative scalar potentials without spoiling the gauge discrete shift symmetries. This is true as long as not only the axion gets shifted, but also the (quantized flux) parameters of the potential transforms appropriately. The structure is similar to the Kaloper-Sorbo type of potential discussed above generalised to include multiple axions and higher order polynomial interactions. Thus the structure is again that of monodromic axions, whose symmetries are again better described in terms of 4d 3-forms. See [277] for a more general discussion.

So string theory contains two of the required ingredients to construct relaxion models: 1) there are axions and 2) they have the required multi-branched structure so that the axion potential does not spoil the shift symmetry. In this Section we try to take a further step and attempt to build a string construction with the required axion-Higgs couplings appearing in a relaxion model. We use Type IIB orientifolds, which have a rich structure of axions, and exploit D-brane physics to engineer the monodromy in an eventually semirealistic construction. We will just provide an explicit example of the required structure within a toy model, and will not pursue the building a complete model. We will also see the limitations of the approach and the prospects for a more realistic structure.

The setup we consider is a stack of 3 parallel D5-branes wrapping a 2-cycle in a Type IIB orientifold compactification[7] with an orientifold with O3/O7-planes. In this type of compactifications there are axion fields $b_a$ arising from the expansion of the NSNS field $B_{MN}$ on harmonic 2-forms $\omega_a$ which are odd under the orientifold reflection. We choose our axion $\phi$ to be one of this kind, $B_2 = \phi\,\omega_2$. This structure is similar to the one appearing in the original axion monodromy models of string theory in [227, 228].

Let us now describe the realization of the SM gauge and Higgs fields. The initial gauge group associated to the three D5-branes is $U(3)$, and in general there are adjoint scalar fields $\Phi_i$ which parametrize the motion of the D5-branes in the four compact dimensions transverse to the branes. The dynamics of these scalars is described by the Dirac-Born-Infeld (DBI) plus Chern-Simons action. The details of the computation are presented in Appendix F. After some simplifications the DBI action applied to the mentioned system has the structure

$$S_{DBI} = -\mu_5 g_s^{-1}\mathrm{STr}\int d^6\xi\, I, \tag{8.6.1}$$

$$I = \sqrt{\left(1+2\sigma^2\partial_\mu\Phi^i\partial^\mu\bar{\Phi}^i\right)\left(1+\frac{1}{2}g_s^{-1}\mathcal{F}_{ab}\mathcal{F}^{ab}\right)(1-4\sigma^2|[\Phi^1,\Phi^2]|^2)^2}\ , \tag{8.6.2}$$

where $\mu_5 = (\alpha')^3/(2\pi)^5$ is the D5-brane tension, $g_s$ is the string coupling (dilaton)

[7] See Chapters 11 and 12 in [10] for a detailed discussion of this class of string vacua.

and

$$\mathcal{F}_{ab} = \sigma F_{ab} - B_{ab} \ , \tag{8.6.3}$$

where $\sigma = 2\pi\alpha'$ and $(\alpha')^{-1}$ is the string tension. We allow for magnetic flux quanta along the $U(1)$ of the $U(3)$, $F_2 = q\,\omega_2$, and also assume there is an axion zero mode $B_2 = \phi\,\omega_2$. Here $\omega_2$ is an odd 2-form Poincare dual to the 2-cycle $\Sigma_2$ wrapped by the $D5$'s. Expanding this expression to second order in 4d derivatives we obtain

$$\frac{g_s S_{DBI}}{\mu_5} = -\mathrm{STr}\int d^6\xi\left(1-4\sigma^2|[\Phi^1,\Phi^2]|^2\right)\left(1+\sigma^2\partial_\mu\Phi^i\partial^\mu\bar{\Phi}^i+\frac{1}{4}g_s^{-1}\mathcal{F}_{ab}\mathcal{F}^{ab}+\dots\right) \tag{8.6.4}$$

where we have neglected higher order terms in $\mathcal{F}$. Now take the adjoints $\Phi_1$, $\Phi_2$ as

$$\Phi^1 = \begin{pmatrix} 0 & 0 & h^1 \\ 0 & 0 & h^2 \\ (h^1)^* & (h^2)^* & 0 \end{pmatrix}, \quad \Phi^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & m \end{pmatrix},$$

$$[\Phi^1,\Phi^2] = \begin{pmatrix} 0 & 0 & mh^1 \\ 0 & 0 & mh^2 \\ -m(h^1)^* & -m(h^2)^* & 0 \end{pmatrix}. \tag{8.6.5}$$

This correspond to one of the 3 D5-branes getting slightly displaced from the rest, giving rise to a configuration with a gauge group $SU(2)\times U(1)$. Here $m$ is generically of order the string scale, i.e. $m^{-2}\simeq\alpha'$. Now, the potential for the axion arises from

$$\frac{1}{4}g_s^{-1}\mathcal{F}_{ab}\mathcal{F}^{ab} \propto (\sigma F_2 \ - \ B_2)^2 \propto (q-\tilde{g}\phi)^2 \ . \tag{8.6.6}$$

This potential may be understood as arising from a Kaloper-Sorbo structure [271]. The relevant 4-form $F_4$ is the dual of $F_2$ in the D5-brane worldvolume, and has a cross coupling $F_4\wedge B_2$, which upon using the equations of motion for $F_4$ and $B_2 = \tilde{g}\phi\omega$ produces the above expression. On the other hand using the matrices for $\Phi_1, \Phi_2$ above and tracing over gauge indices, we get

$$\mathrm{Tr}|[\Phi^1,\Phi^2]|^2 = -m^2|h|^2 \tag{8.6.7}$$

leading to a scalar potential

$$V_{DBI} \ = \ \mu_5 V_{\Sigma_2} g_s^{-1}\left(1+4\sigma^2 m^2|h|^2\right)\left(1+(q-\tilde{g}\phi)^2\right) \ . \tag{8.6.8}$$

There is a built-in KS symmetry which enforces the dependence on the axion to necessarily appear in powers of $F_4 = (q-\tilde{g}\phi)$. However, that is not the case for the Higgs field, so that a large mass term of order the cut-off is expected to appear at the quantum level, beyond the original DBI contribution to its mass. Including this one obtains

$$V \ = \ -\mu^2|h|^2 \ + \ V_{DBI} \ = \ M^4(q \ - \ \tilde{g}\phi)^2 \ + \ (-\mu^2 \ + \ 4M^4\sigma^2 m^2(1+(q-\tilde{g}\phi)^2)\ |h|^2 \tag{8.6.9}$$

where we have assumed that the quantum correction of the Higgs mass is negative and $M^4 = \mu_5 V_{\Sigma_2} g_s^{-1}$, and have ignored a constant term which is not relevant for the discussion. The Higgs kinetic term is not canonical, and after redefining to the canonically normalized Higgs $\tilde{h}$ one finally gets a potential

$$V = \left(-\frac{\mu^2}{1+(q-\tilde{g}\phi)^2} + 4M^4\sigma^2 m^2\right)|\tilde{h}|^2 + M^4(q - \tilde{g}\phi)^2 . \tag{8.6.10}$$

The structure of this potential is of relaxion type, although not of the minimal class we discuss in the text. At large $\phi$ the Higgs mass-squared is positive, the axion then starts decreasing and at a certain point the Higgs mass vanish, a vev develops and then the non-perturbative QCD (or other) condensate[8] stops the vev. Note that one can obtain the above potential if a 4-form $F_4$ in a KS term couples also to the Higgs field through a term in the action of the form $F_4^2/(1+4\sigma^2 m^2|h|^2)$.

This is an interesting toy model with some similarities with the simplest relaxion model, but still far from a fully realistic realization. The most obvious difficulty in a string embedding of the original relaxion dynamics comes from the required mass scales. The natural scale in the above potential is the string scale, whereas in a relaxion model like that in Section 8.3 the 4-form scales are of order $10^{-3}$ eV. One could perhaps consider models with low string scale, but that would also lower the cut-off scale $M$.

To see this in more detail, let us compute the relaxion couplings $M$, $g$ and $f$ which we used in previous sections in terms of stringy parameters. First, let us fix the axion decay constant. The IIB supergravity part of the action relevant here is [12, 10]

$$-\frac{1}{4\kappa_{10}^2}\int d^{10}x|H_3|^2, \tag{8.6.11}$$

where $\kappa_{10}^2 = \frac{1}{2}(2\pi)^7\alpha'^4$. The $B$ field couples to worldsheet instantons wrapping $\Sigma_2$ [12, 10] as

$$\frac{1}{2\pi\alpha'}\int_{\Sigma_2} B_2 \tag{8.6.12}$$

We now expand $B_2 = \tilde{g}\phi\omega_2$, where we remind the reader that $\omega_2$ is the volume form of $\Sigma_2$. We choose $\tilde{g}$ so that (8.6.11) is canonically normalized, and hence $\tilde{g}^2 V_6/(4\kappa_{10}^2) = \frac{1}{2}$, where $V_6$ is the volume of the compactification manifold. With this choice of $\tilde{g}$, we may read the canonical axion decay constant $f$ from (8.6.12), to have

$$\tilde{g}^2 = \frac{2\kappa_{10}^2}{V_6}, \quad f = \frac{2\pi\alpha'}{\tilde{g}V_{\Sigma_2}} = \sqrt{\frac{V_6}{2\kappa_{10}^2}}\frac{2\pi\alpha'}{V_{\Sigma_2}}. \tag{8.6.13}$$

One may also get the value of $\Lambda_k^2$, the 4-form quantum, using the above and the DBI action. If we momentarily take $F_2 = 0$, we may obtain $g$ from the DBI by

[8] This can arise from non-perturbative gauge dynamics on D-branes to which the axion couples, or from euclidean D-brane instanton effects, see [10] for background.

substituting $B_2 = \tilde{g}\phi\omega_2$ and expanding to second order in $\phi$, from which we get

$$S_{DBI} \approx \frac{1}{2}\frac{\mu_5}{g_s}V_{\Sigma_2}\tilde{g}^2\phi^2, \quad \Rightarrow \quad g^2 = \frac{\mu_5}{g_s}V_{\Sigma_2}\tilde{g}^2 = \frac{2\sqrt{\pi}}{4\pi^2\alpha'}\frac{V_{\Sigma_2}}{V_6 g_s}\kappa_{10}. \tag{8.6.14}$$

From this, we get $\Lambda_k^4 = (2\pi g f)^2$ as

$$\Lambda_k^4 = \frac{4\pi^2\alpha'\sqrt{\pi}}{g_s V_{\Sigma_2}\kappa_{10}} =\sim \frac{M_s^2}{g_s V_{\Sigma_2}}. \tag{8.6.15}$$

This same value of $\Lambda_k^2$ can be obtained in a different way. Allowed values of the flux of the D5 gauge field $F_2$ are $F_2 = \frac{2\pi n}{V_{\Sigma_2}}\omega_2$ for integer $q$, so that $\int_{\Sigma_2} F_2$ is an integer multiple of $2\pi$ (this ensures that the lower D-brane charges induced by the D-brane Chern-Simons terms satisfy the right Dirac quantization condition). Plugging this back into the DBI action, we get an action of the form $\frac{1}{2}\Lambda_k^2 n^2$, with $\Lambda_k$ given by (8.6.15), as expected.

Admittedly, we have no control over $\mu^2$, which is why the model falls short from a stringy realization of relaxation. One possible way out is to locate the relevant branes into warped throats which can exponentially reduce the mass scales $M, \sigma, m$ through a warping factor [286, 287, 288]. This is in fact used in KKLT-like moduli fixing models [289], in which this exponential suppression is also important in order to fine-tune the cosmological constant. It is however not obvious that this proposal does not turn the relaxion proposal into a Randall-Sundrum solution of the hierarchy problem, as the quantum corrections to the Higgs mass $\mu^2$ will also be warped. We speculate that this may not be the case if $\mu^2$ receives contributions from sources outside of the throat; however, we have no explicit implementation of this idea.

The WGC membrane must already be there in string theory. Indeed, this is the case: the membrane is the object coupling to the 4d 3-form, which is the magnetic dual potential of the gauge field on the D5-brane worldvolume. Hence, the WGC membrane is a monopole on the D5-brane worldvolume gauge theory, realized as a D3-brane ending[9] on the D5-brane [292].

An additional improvement to achieve a more realistic model would be to realize the full SM gauge group and matter, and achieve full moduli fixed. This would require a global embedding, in particular with cancellation of RR tadpoles, and hence possibly requiring anti-D5-branes, thus complicating the construction. Note finally that the assumption of a large negative Higgs mass-squared from quantum corrections is crucial to have the relaxation effect, since otherwise the potential would have been positive definite and the Higgs would never reach a zero mass point.

In summary, string theory has in principle many of the ingredients of the simplest relaxion models. However, the detailed realization should address the difficulties of obtaining a viable example, in particular producing the very peculiar structure of mass scales in relaxion models.

[9] One may wonder where the other boundary of the D3-brane is. Actually, the worldvolume $U(1)$ gauge field whose dual has KS coupling must have no Stuckelberg couplings [78], and the latter condition implies that the $U(1)$ resides on a combination of D-branes with a homology relation, which thus defines a chain [290, 291], on which the monopole D-brane actually wraps.

# Part III

# Conclusions & Appendices

# 9 Conclusions

## 9.1 English

The results presented in this thesis are admittedly heterogeneous, but the approaches followed in both parts nevertheless share one basic trait: the focus on gauge symmetries as a tool to construct and understand interesting charged states in the theory.

In Part I, we have introduced a variety of different supercritical string theories, and explained how they are related to the different critical string theories via a process of tachyon condensation. Using these supercritical string theories, we have shown how several discrete symmetries of the 10d string theory can actually be embedded into continuous ones in a supercritical version of the theory, which is broken by a Higgs-like mechanism triggered by the closed string tachyon. The examples include outer automorphisms of the gauge group, discrete isometries of the compactification space, quantum symmetries at orbifolds, etc. By embedding these symmetries into continuous ones we establish beyond doubt their gauge character (as expected in a quantum theory of gravity) and gain a tool to build the topological solitons charged under the original discrete symmetry. We have analyzed in detail these charged massive objects, which in some cases were already known and in others were new.

The possibilities of supercritical string theories with regards to the construction of charged objects are not limited to those related to discrete symmetries. We have constructed the heterotic NS5-brane as a closed-string tachyon soliton, which remarkably can be well understood using the perturbative worldsheet description of the supercritical string. We have checked the consistency of the construction using a variety of arguments both from the spacetime and worldsheet points of view.

In the process of building the NS5 we have learned that heterotic NS charges naturally take values in real K-theory, much like their RR counterparts. This classification of charges is morally the S-dual of the same clasification for type I branes, and allows us to systematically build other NS solitonic objects in a perturbative manner in terms of closed string tachyons. This perturbative worldsheet description is a powerful tool which may be used to study the strongly-coupled dynamics of these objects. It also provides a K-theoretic understanding of the heterotic anomalous Bianchi identity, and thus of the involved charges.

A similar description holds in the type II context, although the solitons constructed are very different. The tachyon profile which gave rise to the heterotic NS5 now gives a singular $\mathbb{C}^2/\mathbb{Z}_2$ orbifold in the critical theory. Again, this is argued from several points of view, using D-brane probes to measure the charge of the object, and also analyzing the nonlinear sigma model of the supercritical theory, showing that in the IR it flows to a $(4,4)$ singular CFT, as expected for an $A_1$ singularity.

We also studied a codimension 8 soliton which corresponds to a fundamental string in the critical theory. To argue this, we had to generalize the well-known $B_2X_8$ coupling of type IIA theory to the critical setup, carefully analyzing the parity anomaly of all the fields in the supercritical theory. The $B_2X_{8+2n}$ coupling of the supercritical theory has an structure which suggests the presence of a real K-theory group classifying the charges of the solitons of the supercritical theory, similar to what happened in the heterotic context. We use the K-theory picture to classify the different solitons which can be obtained, providing explicit tachyon profiles for all of them.

The tachyon condensation picture works in general spacetimes. In fact, it is possible to give a physical meaning to the ambient space commonly employed in GLSM's as a supercritical background; by choosing a suitable tachyon, dynamics is constrained to the critical CY embedded in the ambient space. This allows us to embed discrete isometries of the CY (even nonabelian ones) into continuous ones in the supercritical manifold, providing a picture in which they are broken by the tachyon.

Finally, we speculate about a possible connection between the closed-string tachyon we use and a recent proposal which uses Matrix Factorizations to explain type II T-brane backgrounds.

In part II, we have introduced the Weak Gravity Conjecture as the sharpest tool we currently have to cut out the bad weeds in the Swampland. We have introduced the different forms of the conjecture currently in use, discussed the generic arguments based on black hole evaporation supporting it, the evidence within String Theory, and how the different versions of the conjecture relate to each other.

We have then tried to use WGC-like argument to constraint (non-monodromic) large field inflation models based on axions. Since the WGC for axions is on loose ground, we have focused instead on gravitational instantons of the effective theory of Einstein gravity coupled to the axion, which induce corrections to the axion potential. Morally, these gravitational instantons are to WGC instantons what charged black holes are to ordinary WGC particles. Focusing on them is enough to ascertain the generic form of the corrections to the axion potential. Although the existence of the euclidean instantons is in principle independent of the validity of the WGC, we need to take them outside the regime of validty of effective field theory if we want to place strong constraints on effective field theories: it is in this regime when the WGC instantons become important.

In single-axion models, gravitational instanton effects spoil transplanckian axion decay constant $f$, since the instanton action decreases with $f$. In multiple axion setups, we have considered different scenarios of axion alignment. We have shown that in models of lattice alignment, the parametrically transplanckian axion suffers gravitational instanton effects which jeopardize its transplanckian field range, so in this sense there is not advantage with respect to the single field case. Finally, we have shown that kinetic alignment can be used to achieve moderately transplanckian field ranges, of the form $\sqrt{N} f_{\max}$, where $N$ is the number of axions and $f_{\max}$ the maximum axion periodicity, corresponding to travelling along a diagonal in the lattice. In the effective single-axion model along this transplanckian direction, the gravitational instantons which would spoil the transplanckian range are absent due to a $\mathbb{Z}_N$ discrete gauge symmetry. To be able to exploit this "loophole", however, we need to violate the 0-form strong versions of the WGC. These forms are true in every String Theory example considered so far.

We have also used the WGC to place new constraints on the relaxion proposal to solve the electroweak hierarchy problem. The simplest implementations based on this idea have however a number of problems, including the inconsistent explicit breaking of the axion discrete gauge symmetry, as well as the stability of extremely large trans-Planckian excursions of the relaxion. We have described how one can construct relaxion type models with a monodromy structure built in, which avoids these two problems. The simplest way to achieve this is by coupling a Minkowski 3-form, with a quantized flux $F_4$ coupling both to the relaxion and the Higgs field. The gauge symmetries of the 3-form guarantee the stability of the axion potential under Planck suppressed corrections, similar to the Kaloper-Sorbo structure applied in large field inflation. This allows for large trans-Planckian excursions of the relaxion field. In these constructions the axion shift comes along with shifts of the $F_4$ quanta, so that there is a branched structure of axion potentials. The shift symmetry is in this way consistent with the presence a non-vanishing potential for the relaxion.

An interesting consequence of the monodromy version of relaxion is that the mass parameter $g$ is related to the 4-form flux by $g = F_4/f$, so that one may understand the smallness of the $g$ parameter from a sort of see-saw structure. Thus, e.g., values of the 4-form flux $F_4 \simeq (10^{-3}\text{eV})^2$ and axion decay constants $f \simeq 10^{10}$ GeV give rise to values $g \simeq 10^{-34}$, in the ballpark of simplest relaxion models. The fact that this value for $F_4^2$ is of order of the observed cosmological constant $\Lambda_{dark} \simeq (10^{-3}\text{eV})^4$ is intriguing.

We also describe how in explicit string constructions the main ingredients of the relaxion mechanism are present. There are axions arising from the dimensional reduction of RR and NS antisymmetric tensors, and these axions may have couplings to Higgs doublets with a structure similar to relaxion models. A toy model from the DBI dynamics of D5-branes in Type IIB orientifolds is presented. However the natural scale for the 4-forms in string theory is the string scale $F_4 \simeq M_s^2$, whereas realistic relaxion models have rather very low values, e.g. $F_4 \simeq (10^{-3}\text{eV})^2$. It would be interesting to see whether the presence of strong warping effects or other mechanisms could explain this difference in scales.

We have also studied the constraints on the relaxion parameters coming from membrane nucleation. In particular there are membranes coupling to the 3-forms which induce changes on the 4-form quanta, while also changing the axion branch. These jumps make the slow roll of the relaxion unstable, so that we must impose that the rate for these processes is sufficiently suppressed. To check whether the nucleation of membranes is suppressed or not we need to know the membrane tension $T$. The Weak Gravity Conjecture provides for an upper bound for the tension of order $T \leq gfM_p$. This implies that the transition rate from one relaxion branch to another is indeed suppressed if the cut-off scale is below a certain scale ($M \lesssim 10^9$ GeV in the simplest axion model) which is comparable to other limits. However, taking into account the extremely large period of inflation in which bubbles may be generated in relaxion models, one finds that the upper bound on the cut-off scale becomes much tighter, of order $M \lesssim 300$ TeV. This would imply that the simplest implementations of the relaxion mechanism are able to generate a small hierarchy not much above the EW scale, but not to explain the hierarchy problem itself.

Overall, the work presented in this thesis is aimed at understanding Quantum Gravity from two complementary points of view. On one hand, we have explored aspects of String Theory which are not commonly emphasized, such as the existence of a supercrtitical landscape and its relation to the critical theories via closed-string tachyon condensation. String Theory is so far the only candidate quantum theory of gravity in which we have enough theoretical control so as to perform complicated checks and computations. We should therefore try to understand as much of the theory as we can.

On the other hand, we have used the Weak Gravity Conjecture as a means to weed out the Swampland, constraining effective field theories of both large field inflation and relaxation. We are lucky that such a thing as the Swampland exists, at least in String Theory. The scale of gravity is enormously high, at $10^{19}$ GeV. It is very difficult to imagine that we will be able to access these scales experimentally within our lifetimes, even in an indirect way. Although we may be unexpectedly fortunate, it is possible that until we reach the Planck scale every bit of new physics that shows up can be nicely explained within the realm of effective field theory. In that case, narrrowing the Swampland then constitutes our best shot at checking our ideas about Quantum Gravity and String Theory.

## 9.2 Español

Los resultados presentados en esta tesis son decididamente heterogéneos, pero las metodologías seguidas en ambas partes comparten un rasgo común: Ambas se centran en emplear simetrías gauge como herramienta para construir y entender parte del espectro cargado de la teoría.

En la Parte I, se ha introducido las diferentes teorías supercríticas, explicando su relación con las teorías críticas mediante un proceso de condensación del taquión. Usando estas teorías supercríticas, se ha mostrado cómo varias simetrías discretas de la teoría de cuerdas en diez dimensiones pueden embeberse en simetrías continuas en una versión supercrítica de la teoría, que después se rompe mediante un mecanismo similar al de Higgs iniciado por el taquión de cuerda cerrada. Los ejemplos presentados incluyen automorfismos externos del grupo gauge, isometrías discretas del espacio de compactificación, simetrías cuánticas de orbifolds, etcétera. Al embeber estas simetrías discretas en simetrías continuas, se estable más allá de toda duda el carácter gauge de las primeras (como era de esperar en una teoría cuántica de la gravedad). Al mismo tiempo, se obtiene una herramienta para construir solitones topológicos cargados bajo la simetría discreta original. Estos objetos cargados masivos han sido analizados en detalle. En algunos casos ya eran soluciones conocidas y en otros no.

Las posibilidades de las teorías de cuerdas supercríticas como herramienta para la construcción de objetos cargados no se limitan a ejemplos relacionados con simetrías discretas. También se ha construido en esta tesis la brana NS5 heterótica como un solitón del taquión de cuerda cerrada, el cual puede entenderse utilizando la descripción de hoja de mundo de la cuerda supercrítica. Se ha comprobado la consistencia de esta construcción usando diferentes argumentos, tanto desde el punto de vista del espaciotiempo como desde el punto de vista de la hoja de mundo.

El proceso de construir la NS5 ha revelado que las cargas Neveu-Schwartz heteróticas toman valores de manera natural en una teoría K real, de manera similar a sus contrapartidas Ramond-Ramond en tipo I. Esta clasificación de cargas es, moralmente, S-dual a la clasificación similar para branas de tipo I, y permite la construcción sistemática de otros solitones NS de manera perturbativa, en términos de taquiones de cuerda cerrada. El punto de vista perturbativo es una poderosa herramienta que puede ser usada para analizar las dinámicas a acoplo fuerte de estos objetos, al tiempo que proporciona una nueva perspectiva sobre la identidad de Bianchi anómala de la teoría heterótica, a través de teoría K.

En el contexto de teorías de cuerdas de tipo II se ha encontrado una descripción similar, aunque los solitones construidos son muy diferentes. El perfil del taquión que da lugar a la NS5 en el contexto heterótico constituye un orbifold $\mathbb{C}^2/\mathbb{Z}_2$ en la teoría de tipo II crítica. De nuevo, esto se argumenta desde diferentes puntos de vista, usando D-branas como sondas para medir la carga del objeto, así como analizando el modelo sigma no lineal de la teoría supercrítica, mostrando que en el infrarrojo fluye a una teoría conforme con supersimetría $(4,4)$ y singular, como se espera para una singularidad de tipo $A_1$.

También se estudió un solitón de codimensión 8 que corresponde a una cuerda fundamental en la teoría critica. Para argumentar esto, se generalizó el acoplo $B_2X_8$ de la teoría tipo IIA al contexto supercrítico, mediante un cuidadoso análisis de la anomalía de paridad de todos los campos de la teoría supercrítica. La estructura del acoplo $B_2X_{8+2n}$ de la teoría supercrítica sugiere, de manera similar al caso heterótico, una estructura de teoría K real que clasifica las cargas de los solitones en la teoría supercrítica. Esta perspectiva de teoría K se usa para clasificar los diferentes solitones obtenidos, proporcionando perfiles del taquión explícitos para todos y cada uno de ellos.

El mecanismo de condensación del taquión funciona en espaciotiempos generales, lo que posibilita dar un significado físico al espacio ambiente empleado habitualmente en modelos sigma lineales gaugeados: Al escoger un perfil del taquión adecuado, la dinámica acaba restringida a la variedad de Calabi-Yau crítica embebida en el espacio ambiente. Ello permite embeber isometrías discretas del Calabi-Yau, incluso isometrías no abelianas, en isometrías continuas en el espacio supercrítico, que luego rompe el taquión.

Finalmente, se ha especulado sobre una posible conexión entre el taquión de cuerda cerrada empleado en esta tesis y una propuesta reciente que utiliza factorizaciones de matrices para explicar configuraciones de T-brana en teorías de cuerdas de tipo II.

En la parte II, se ha presentado la Conjetura de la Gravedad Débil (CGD) como la herramienta más afilada de que disponemos actualmente para cortar las malas hierbas del Pantano. Se han presentado las diferentes formas de la conjetura en uso, discutiendo los argumentos genéricos a su favor basados en evaporación de agujeros negros, la evidencia en el contexto de teoría de cuerdas, y cómo las diferentes versiones de la conjetura se relacionan unas con otras.

A continuación, se ha empleado argumentos similares a la CGD para acotar modelos axiónicos no monodrómicos de inflación de campo grande. Dado que la CGD no tiene fundamentos sólidos en el caso de axiones, se ha puesto el énfasis en instantones gravitacionales de la teoría efectiva de gravedad de Einstein acoplada al axión, los cuales inducen correcciones al potencial axiónico. Moralmente, estos instantones son a los instantones predichos por la CGD como los agujeros negros cargados ordinarios a las partículas CGD. El estudio de estos instantones es suficiente para determinar la forma genérica de las correcciones al potencial axiónico. Aunque la existencia de estos instantones es en principio independiente de si la CGD es correcta o no, para acotar significativamente los parámetros de la teoría efectiva del axión es necesario considerarlos fuera del régimen de validez de dicha teoría efectiva.

En modelos de un solo axión, los instantones gravitacionales impiden que la constante de decaimiento $f$ del axión adquiera valores transplanckianos, dado que la acción del instantón decrece al crecer $f$. En modelos con múltiples axiones, se han considerado diferentes posibilidades de alineamiento. En modelos de alineamiento de red, el axión que sería paramétricamente transplanckiano siente efectos de instantón gravitacional que destruyen esta propiedad, por lo que estos modelos no presentan ventaja alguna sobre los modelos de un solo axión. Los modelos de alineamiento cinético pueden conseguir constantes de decaimiento moderadamente transplanckianas, de la forma $\sqrt{N} f_{\max}$, donde $N$ es el numero de axiones y $f_{\max}$ la máxima periodicidad posible para un axión, que corresponde a viajar a lo largo de una diagonal en la red. En la teoría efectiva de un único axión a lo largo de esta dirección transplanckiana, los instantones gravitacionales que impedirían obtener una constante de decaimiento transplanckiana no están presentes debido a una simetría gauge discreta $\mathbb{Z}_N$. Para poder explotar este resquicio en el argumento, es necesario sin embargo violar las versiones fuertes de 0-forma de la CGD. Sin embargo, estas formas se cumplen en todos los ejemplos construidos en teoría de cuerdas hasta la fecha.

También se ha usado la CGD para obtener nuevas cotas en los parámetros de la propuesta relaxiónica para resolver el problema de las jerarquía electrodébil. Las implementaciones más simples de la idea sufren de varios problemas, incluyendo la inconsistencia de romper explícitamente la simetría gauge discreta del axión, así como mantener la estabilidad del potencial a lo largo de la excursión del relaxión, que es increíblemente transplanckiana. Se ha descrito cómo construir modelos relaxiónicos con estructura monodrómica, lo cual evita estos dos problemas. La forma más sencilla de conseguir esto es mediante el acoplo de una 3-forma Minkowskiana con un flujo $F_4$ cuantizado, el cual acopla al mismo tiempo al relaxión y al campo de Higgs. Las simetrías gauge de la 3-forma garantizan la estabilidad del potencial relaxiónico frente a correcciones suprimidas por la escala de Planck, similares a la estructura Kaloper-Sorbo de uso frecuente en modelos monodrómicos de inflación de cambo grande. Ello permite excursiones transplanckianas del relaxión. En estas construcciones, la simetría de desplazamiento del axión viene acompañada de desplazamientos del cuanto de $F_4$, de tal forma que hay una estructura de ramas en el potencial del relaxión. La simetría de desplazamiento es de este modo consistente con la presencia de un potencial no periódico y no nulo para el relaxión.

Una consecuencia interesante de la versión monodrómica del relaxión es que el parámetro con unidades de masa $g$ se relaciona con el flujo de 4-forma según $g = F_4/f$, de tal forma que se puede entender la pequeñez de $g$ mediante un mecanismo de tipo balancín. Así, valores del flujo de 4-forma $F_4 \simeq (10^{-3}\text{eV})^2$ y constantes de decaimiento $f \simeq 10^{10}$ GeV dan lugar a valores $g \simeq 10^{-34}$ GeV, que son típicos de los modelos relaxiónicos más simples. Es intrigante que el valor de $F_4^2$ es del orden de la constante cosmológica observada, $\Lambda_{dark} \simeq (10^{-3}\text{eV})^4$.

Los ingredientes principales de un modelo de relaxión están presentes en teoría de cuerdas: Hay axiones que surgen de la reducción dimensional de tensores antisimétricos tanto en el sector RR como en el sector NS, y dichos axiones pueden tener acoplos a dobletes de Higgs con una estructura similar a la presente en modelos relaxiónicos. Se ha presentado un modelo simplificado que proviene de la dinámica, regida por la acción Dirac-Born-Infeld, de D5-branas en orientifolds de tipo IIB. Sin embargo, la escala natural de las 4-formas en teoría de cuerdas es la escala de la cuerda, $F_4 \simeq M_s^2$, mientras que un modelo relaxiónico realista debe tener valores muy bajos, del orden de $F_4 \simeq (10^{-3}\text{eV})^2$. Sería interesante determinar si esta diferencia en escalas puede explicarse mediante efectos de curvatura fuertes u otros mecanismos.

También se han estudiado las cotas en los parámetros del relaxión provenientes de nucleación de membranas. En particular, existen membranas que se acoplan a las 3-formas, las cuales inducen cambios en el cuanto de 4-forma y cambian la rama del relaxión. Estos saltos vuelven inestable la condición de rodadura lenta del relaxión, por lo que se vuelve necesario imponer que la tasa de producción de dichas burbujas sea lo bastante pequeña. Para determinar esto, es necesario conocer la tensión de las membranas, $T$. La CGD proporciona una cota superior a dicha tensión, del orden $T \leq gfM_p$. Ello implica que la tasa de transición de una rama del potencial a otra está muy suprimida si la escala de corte de la teoría efectiva está por debajo de una cierta escala (igual a $10^9$ GeV en la clase más sencilla de modelos relaxiónicos), comparable a otras cotas ya existentes en dicho parámetro. Sin embargo, se ha de tener en cuenta la enorme duración de la fase inflacionaria en los modelos relaxiónicos, durante la cual la producción de burbujas es posible. Teniendo esto en cuenta, la cota en la escala de corte se vuelve mucho más pequeña, del orden de $M \lesssim 300$ TeV en el modelo más sencillo de relaxión. Ello implicaría que estas implementaciones sencillas del mecanismo del relaxión son capaces de generar una pequeña jerarquía no mucho más allá de la escala electrodébil, pero no pueden resolver el problema de las jerarquías por completo.

De manera general, la finalidad del trabajo presentado en esta tesis es afianzar nuestro entendimiento de la gravedad cuántica desde dos puntos de vista complementarios. Por un lado, se han explorado aspectos de la teoría de cuerdas que no reciben atención frecuentemente, como la existencia de una red de teorías supercríticas conectadas a las teorías críticas y entre sí mediante condensación del taquión de cuerda cerrada. La teoría de cuerdas es, hasta ahora, la única candidata a teoría cuántica de la gravedad en la que hay suficiente control matemático como para realizar cálculos y verificar resultados complicados. Deberíamos por tanto intentar entender tanto de la teoría como podamos.

Por otro lado, hemos usado la Conjetura de la Gravedad Débil como medio para quitar las malas hierbas del Pantano, poniendo cotas a los parámetros de ciertas teorías efectivas tanto de inflación de campo grande como de modelos de relaxión. Tenemos suerte de que el Pantano exista, al menos en teoría de cuerdas. La escala característica de la gravedad es enormemente elevada, $10^{19}$ GeV. Es muy difícil imaginar un experimento que proporcione acceso a esta escala, incluso de manera indirecta, en este siglo. Aunque quizá seamos inusitadamente afortunados, también es posible que toda la nueva física que se descubra hasta la escala de Planck pueda ser explicada en el contexto de teoría efectiva. En tal caso, entender mejor el Pantano constituye nuestra mejor baza para comprobar nuestras ideas sobre gravedad cuántica y teoría de cuerdas.

# A

# Partition functions of supercritical strings

Despite the long history of supercritical strings [3], their detailed construction may be unfamiliar to most readers. In this Appendix we complement the main text references therein (e.g. [4, 293]) by providing their partition functions, which help in visualizing the diverse GSO-like projections. Their construction is standard, with the linear dilaton simply readjusting the zero point energy to reproduce the masses used in the main text.

We start with the supercritical bosonic string in $D$ dimensions, for which we have

$$Z(\tau) = \frac{(4\pi^2\alpha'\tau_2)^{-\frac{D-2}{2}}}{|\eta|^{2(D-2)}} \tag{A.1}$$

For the supercritical type 0A/B in $D = 10 + n$ dimensions, the partition function reads

$$\begin{aligned} Z(\tau) &= \frac{1}{2}\frac{(4\pi^2\alpha'\tau_2)^{-4-\frac{n}{2}}}{|\eta|^{16+2n}|\eta|^{8+n}}\left(\left|\vartheta\begin{bmatrix}0\\0\end{bmatrix}\right|^{8+n} + \left|\vartheta\begin{bmatrix}0\\1/2\end{bmatrix}\right|^{8+n} + \left|\vartheta\begin{bmatrix}1/2\\0\end{bmatrix}\right|^{8+n} \mp \left|\vartheta\begin{bmatrix}1/2\\1/2\end{bmatrix}\right|^{8+n}\right) \\ &= \frac{1}{2}\frac{(4\pi^2\alpha'\tau_2)^{-4-\frac{n}{2}}}{|\eta|^{16+2n}}\left(\left|Z_0^0\right|^{8+n} + \left|Z_0^1\right|^{8+n} + \left|Z_1^0\right|^{8+n} \mp \left|Z_1^1\right|^{8+n}\right) \end{aligned} \tag{A.2}$$

In the last line we have introduced the notation from [12]

$$Z_b^a(\tau) = \frac{\vartheta\begin{bmatrix}a/2\\b/2\end{bmatrix}}{\eta(\tau)} \tag{A.3}$$

Consider now type 0 supercritical theories in $D = 10 + 2k$ dimensions, and orbifold by the $\mathbb{Z}_2$ flipping $k$ extra dimensions, and worldsheet fermions as described in the text. The partition function sums over untwisted and twisted sectors, and reads

$$
\begin{aligned}
Z_0^{\rm orb}(\tau) = \frac{1}{4}(4\pi^2\alpha'\tau_2)^{-4} \mid \eta \mid^{-16} \times \\
\times \Big[ (4\pi^2\alpha'\tau_2)^{-n/2} \mid \eta \mid^{-2n} \left( \left|Z_0^0\right|^8 \left|Z_0^0\right|^n + \left|Z_0^1\right|^8 \left|Z_0^1\right|^n + \left|Z_1^0\right|^8 \left|Z_1^0\right|^n \mp \left|Z_1^1\right|^8 \left|Z_1^1\right|^n \right) \\
+ \mid Z_0^1 \mid^{-n} \left( \left|Z_0^0\right|^8 \left|Z_1^0\right|^n + \left|Z_0^1\right|^8 \left|Z_1^1\right|^n + \left|Z_1^0\right|^8 \left|Z_0^0\right|^n \mp \left|Z_1^1\right|^8 \left|Z_0^1\right|^n \right) \\
+ \mid Z_1^0 \mid^{-n} \left( \left|Z_0^0\right|^8 \left|Z_0^1\right|^n + \left|Z_0^1\right|^8 \left|Z_0^0\right|^n + \left|Z_1^0\right|^8 \left|Z_1^1\right|^n \mp \left|Z_1^1\right|^8 \left|Z_1^0\right|^n \right) \\
+ \mid Z_0^0 \mid^{-n} \left( \left|Z_0^0\right|^8 \left|Z_1^1\right|^n + \left|Z_0^1\right|^8 \left|Z_1^0\right|^n + \left|Z_1^0\right|^8 \left|Z_0^1\right|^n \mp \left|Z_1^1\right|^8 \left|Z_0^0\right|^n \right) \Big]
\end{aligned}
\tag{A.4}
$$

This can be written compactly as

$$
Z_0^{orb.} = \frac{(4\pi^2\alpha'\tau_2)^{-4}}{|\eta(\tau)|^{16}} \frac{1}{4} \sum_{a,b,c,d=0}^{1} Z_{abcd}. \tag{A.5}
$$

by defining generalized partition functions for different sectors

$$
Z_{ab;cd} = |\eta|^{-8} \left| \vartheta \begin{bmatrix} \frac{1}{4}(1+(-)^a) \\ \frac{1}{4}(1+(-)^c) \end{bmatrix} \right|^8 \left| \vartheta \begin{bmatrix} \frac{1}{4}(1+(-)^b) \\ \frac{1}{4}(1+(-)^d) \end{bmatrix}_B \right|^{-n} \left| \vartheta \begin{bmatrix} \frac{1}{4}(1+(-)^{a+b}) \\ \frac{1}{4}(1+(-)^{c+d}) \end{bmatrix} \right|^n \tag{A.6}
$$

Here we have introduced "bosonic" theta functions, which are the partition functions of the orbifolded bosons. They must be replaced by $\eta(\tau)^3(4\pi^2\alpha'\tau_2)^{-1/2}$ in the $b = d = 0$ sector.

Let us move on to the heterotic theories. For the $HO^{(n)}$ theory, the partition function is

$$
\begin{aligned}
Z_{HO^{+(n)}}(\tau) = \frac{1}{4} \frac{(4\pi^2\alpha'\tau_2)^{-4-\frac{n}{2}}}{\mid \eta \mid^{16+2n}} \times \left[ \left(Z_0^0\right)^{16} + \left(Z_0^1\right)^{16} + \left(Z_1^0\right)^{16} + \left(Z_1^1\right)^{16} \right] \\
\times \left[ \left(\bar{Z}_0^0\right)^4 \left|Z_0^0\right|^n - \left(\bar{Z}_0^1\right)^4 \left|Z_0^1\right|^n - \left(\bar{Z}_1^0\right)^4 \left|Z_1^0\right|^n \pm \left(\bar{Z}_1^1\right)^4 \left|Z_1^1\right|^n \right]
\end{aligned}
\tag{A.7}
$$

Note the familiar $SO(32)$ sector in the last factor of the first line. The familiar change of GSO projections in these 32 fermions can be used to directly construct the $E_8 \times E_8$ version of the theory used in chapter 2.

For the $HO^{+(n)/}$ theory, the partition function is

$$
\begin{aligned}
Z_{HO^{+(n)/}}(\tau) = \frac{1}{2} \frac{(4\pi^2\alpha'\tau_2)^{-4-\frac{n}{2}}}{\mid \eta \mid^{16+2n}} \times \Big[ \left(\bar{Z}_0^0\right)^4 \left(Z_0^0\right)^{16} \left|Z_0^0\right|^n - \left(\bar{Z}_0^1\right)^4 \left(Z_0^1\right)^{16} \left|Z_0^1\right|^n \\
- \left(\bar{Z}_1^0\right)^4 \left(Z_1^0\right)^{16} \left|Z_1^0\right|^n \pm \left(\bar{Z}_1^1\right)^4 \left(Z_1^1\right)^{16} \left|Z_1^1\right|^n \Big].
\end{aligned}
\tag{A.8}
$$

The orbifold can be written compactly using the generalized functions (A.6), as

$$
Z_{HO^{(+n)/}}^{orb} = \frac{(4\pi^2\alpha'\tau_2)^4}{|\eta(\tau)|^{16}} \frac{1}{4} \sum_{a,b,c,d=0}^{1} (-1)^{a+c} Z_{abcd}. \tag{A.9}
$$

# B

# A brief review of $K$-theory and its physical applications

Consider a system of $n$ $D9-\overline{D9}$ branes in either type IIB or type I theories[1]. As discussed in Section 2.2.1, this system contains a tachyon describing the annihilation of the two stacks of branes into the vacuum. As also discussed there, by turning on nontrivial worldvolume fluxes on the branes one can reach lower-dimensional branes after tachyon condensation.

The rationale behind this nonhomogeneous tachyon condensation process is that if there is a charge which may be measured at infinity, it cannot be changed by the process of tachyon condensation. The K-theory proposal in [37] is a systematic classification of the global charges that may be carried by a pair of gauge bundles $(E, F)$, corresponding to the brane and antibrane stacks, respectively, modulo tachyon condensation.

A first observation is that pairs of bundles $(E, F)$ and $(E\oplus H, F\oplus H)$ carry the same charges. The physical interpretation of this operation is the nucleation of extra $m = \text{rank}(H)$ $D9$ branes and extra $m$ $\overline{D9}$ branes, over the brane configuration we started with. Since the extra branes are identical, carrying the same gauge bundle $H$, they carry the same worldvolume fluxes, etc. The two bundles are isomorphic and the bifundamental bundle (of which the tachyon is a section) is trivial. In other words, a configuration $(E\oplus H, F\oplus H)$ is related to $(E, F)$ via tachyon condensation, hence should be considered as equivalent for the aim of classifying charges.

Pairs of unitary bundles $(E, F)$ over a given manifold, modulo the equivalence relation $(E, F) \sim (E \oplus H, F \oplus H)$ form a group, called the $K$-theory group of the manifold. Similarly, $SO$ bundles define the real K-theory group $KO$ of the same manifold. Possible asymptotic charges in type IIB string theory are classified by $K$ theory, whereas in type I they are classified by $KO$ theory. To describe a codimension $p$ defect, the gauge bundles should be nontrivially fibered over the transverse $\mathbb{R}^{9-p}$[2].

[1] An extension of these techniques to IIA is possible, but involves the use of unstable $D9$-branes.

[2] In many of our examples the critical slice is $\mathbb{R}^n$, so strictly speaking there are no non-trivial bundles. As we discuss momentarily, we always impose finite energy at infinity for our bundles, so effectively we have in mind bundles on the compactification of flat space.

Both in open string as well as in supercritical constructions, the tachyon (or its derivative) lives in a bifundamental of $(E, F)$. This means it is is sensitive to the K-theory class: The loci where the tachyon matrix decreases its rank corresponds to the presence of extra degrees of freedom associated to a tachyon soliton. Tachyon profiles describing solitons are in one-to-one correspondence with K-theory classes of the $E$, $F$ bundles. A tachyon matrix $\mathscr{T}$ specifies a $K$-theory class uniquely as follows: Given a tachyon $T$ over some compact[3] base space $X$, one may construct a complex vector bundle $E$ by finding an open cover $\{U_j\}$ of $X$ such that the tachyon has full rank on the overlaps $U_i \cap U_j$ (for instance one in which one of the $U_i$ is a tubular neighbourhood of the $\det T = 0$ locus) and setting the transition functions to be just $T$. From this we can further construct a $K$-theory class representative as $(E, 1)$ where 1 is a trivial vector bundle.

However, one must show that this mathematical correspondence has some physical meaning, for instance showing that the tachyon profile requires the $x$, $y$ bundles to form a representative of the tachyon K-theory class. Physically, the equivalence between topologically nontrivial tachyon profiles and the supercritical $x$, $y$ bundles arises from two assumptions about the theory:

- Tachyons describing a localized soliton must relax to the vacuum far away from the core. In practical terms this means that $\mathcal{T} \to \mathcal{T}_0$ as $r \to \infty$, where $\mathcal{T}_0$ is the vacuum expectation value of the tachyon.

- Configurations of finite action on $\mathbb{R}^k$ can be extended to differentiable configurations on its one-point compactification $\mathbb{S}^k$, modulo gauge transformations. This is an assumption about the form of the action near the vacuum. For instance, in Yang-Mills theory in $k$ dimensions, one must have

$$F_{\mu\nu} \to 0 \tag{B.1}$$

  faster than $r^{-k/2}$ for the action to be finite. Since this implies that the connection is flat at infinity, the configuration can be extended to $\mathbb{S}^k$, possibly in a singular gauge. Similarly, assuming that near the tachyon vev the action takes the form

$$S = \int |\mathcal{D}\mathcal{T}|^2 - V(\mathcal{T}) \tag{B.2}$$

  for some effective potential $V(\mathcal{T})$ with a minimum around $\mathcal{T}_0$, finite action means that $\mathcal{D}\mathcal{T} \to 0$ faster than $r^{-k/2}$.

We should remark that whereas the first assumption is quite reasonable and could be considered as part the definition of a soliton, the second is a strong statement about the action in a regime we do not really control. The supercritical description is valid for small vevs of $\mathcal{T}$ and even though it is possible to ascertain the endpoint of the condensation for several different profiles, we know nothing about the effective action there. This is the case in supercritical tachyon condensation but also in the context of brane-antibrane annihilation in type IIB, where the tachyon

[3]Recall footnote 2.

profiles which yield $D$-branes asymptote to the vacuum expectation value of the tachyon. Although some is known about the effective action of the tachyon in this case (see [294] for a review), it is difficult to argue in a controlled way that the covariant derivative of the IIB tachyon should vanish at infinity to satisfy finite energy conditions.

In any case, once we assume that $\mathcal{D}\mathcal{T} \to 0$, we obtain a link between the tachyon and the supercritical bundles. We will review the argument in the more familiar case of D9-$\overline{\text{D9}}$ tachyon condensation in type IIB this more familiar case first, and then briefly discuss the the closed-string version of the phenomenon.

In type IIB in the presence of extra D9-$\overline{\text{D9}}$ pairs, there is a tachyon $\mathcal{T}_{IIB}$ in the bifundamental of the gauge group. A tachyon vev $|\mathcal{T}_{IIB}| \to \infty$ corresponds to complete annihilation to the type IIB vacuum. Different tachyon profiles describe different configurations in the critical theory after tachyon condensation.

The tachyon asymptotics impose constraints on both D9 and $\overline{\text{D9}}$ connections. For concreteness, consider a single brane-antibrane pair and a tachyon profile of the form $\mathcal{T}_{IIB} = z$. This is well-known to describe a D7-brane sitting at $z = 0$ after condensation. Far from $z = 0$ the tachyon is non-vanishing, indicating complete annihilation. However, there is one extra requirement to be imposed. As discussed above, one should have $\mathcal{D}_\mu\mathcal{T} \to 0$ sufficiently fast. Plugging the tachyon profile we have chosen, this means that

$$\partial_z\mathcal{T} - (A_{D9} - A_{\overline{D9}})\mathcal{T} \to 0 \tag{B.3}$$

or, in other words, that asymptotically

$$A_{D9} - A_{\overline{D9}} \to \frac{1}{z}. \tag{B.4}$$

One may measure the first Chern class of the K-theory class $(E, F)$, where $E$ is the D9 gauge bundle and $F$ the $\overline{\text{D9}}$ gauge bundle, by integrating the Chern-Simons form of the connection (B.4). In this case, one obtains an induced D7-brane charge of 1, as befits a D7-brane after condensation. We thus see that in order to be able to draw the correspondence between a tachyon profile and a particular configuration after condensation it is essential to specify the behaviour at infinity.

We will assume that an analogous discussion holds in supercritical theories. The projection of the covariant derivative $\nabla_\mu\mathscr{T}$ to the critical slice (in supercritical type 0) and the gradient of the tachyon $\nabla_\mu\mathcal{T}^a$ (in $HO^{(+n)/}$) are analogous to the covariant derivative of the tachyon in the IIB example above. We will hypothesize that $\nabla_\mu\mathscr{T} \to 0$ plus terms of order $r^{k/2}$ as we near infinity for tachyon profiles describing solitons of codimension $k$. By expanding the covariant derivative as in (B.3), we get a sum of Christoffel symbols which act as connection coefficients over the $E, F$ bundles. The same arguments as above show that a particular tachyon matrix specifies a pair $(E, F)$ of bundles.

Finally, the real K-theory groups of spheres which are of interest in this thesis have been computed by mathematicians. They are $KO(S^1) = KO(S^2) = KO(S^9) =$

$KO(S^{10}) = \mathbb{Z}_2$, $KO(S^4) = KO(S^8) = \mathbb{Z}$. Thus, in type I theory we can have topological defects of spatial dimension $-1$ (an instanton), $0, 1, 5, 7, 8$. The dimension 1 and 5 objects are the well-known $D1$ and $D5$-branes of type I theory. Indeed, D-brane charge is $\mathbb{Z}$ valued, as the relevant KO groups show. The other defects are all $\mathbb{Z}_2$ objects and thus more exotic. The particle is a spinor of $SO(32)$, and is a perturbative massive state of the fundamental string. It flips sign when circling the $\mathbb{Z}_2$ 7-brane. The 7-brane is in fact the gauge object $S$-dual to the the 7-brane of Section 3.2.2.

This concludes the lightning review of the K-theory tools necessary for this thesis. For more information, the reader is referred to [37] and [295, 296] for a more mathematical introduction.

# C

# The Atiyah-Bott-Saphiro construction in supercritical theories

An essential technical tool in building the tachyon profiles used in Chapters 4 and 5 is the Atiyah-Bott-Saphiro construction [98], which we now review. While the original ABS construction works for complex bundles, it can be employed in the real context with suitable modifications, as described beloww.

## C.1 Real Atiyah-Bott-Saphiro profiles

As discussed in Appendix B, tachyon profiles are in one to one correspondence with K-theory classes. In the $U(n)$ case, a simple representative of the non-trivial homotopy classes is obtained by the Atiyah-Bott-Shapiro (ABS) construction [297]. In this case, the K-theory group is $K(\mathbb{S}^{k-1}) = \mathbb{Z}$ for even $k$, $k = 2p$. The two bundles are taken to describe the chiral spinor bundles $S^{\pm}$ over $\mathbb{R}^k$, namely $U(n)$ bundles with $n = 2^{p-1}$, and the tachyon background generating the K-theory group is

$$\mathscr{T} = \Gamma \cdot \vec{x} \tag{C.1.1}$$

where $\Gamma$ denote the Dirac matrices in Weyl representation, so we indeed have a map $M : S^+ \to S^-$. The $\mathbb{Z}$-valued charge can be identified with the non-trivial Chern classes $\mathrm{tr}\, F^p$, $k = 2p$ along $\mathbb{R}^k$.

We are interested however in generators of the real K-theory groups $KO(S^k)$. Throughout this Section $k = 2p$ will be the dimension of the transverse space to the soliton we wish to construct.

The main tool we will use is the embedding from $U(n)$ to $SO(2n)$ given by treating real and imaginary parts independently (this embedding process sometimes goes under the name "realification"). Given any $n$-dimensional complex vector bundle, it can be regarded as a $2n$-dimensional real vector bundle via the mapping

$$r: \ (z_1, z_2, \ldots) \quad \to \quad (x_1, y_2, x_2, y_2, \ldots), \quad \text{where} \quad z_k \equiv x_k + i y_k. \tag{C.1.2}$$

A mapping $M$ between two complex vector bundles $E, F$, such as the tachyon, induces a mapping $M_{\mathbb{R}} = r_F \circ r_E^* M$ from $r(E)$ to $r(F)$. Given $M$, $M_{\mathbb{R}}$ is just obtained by the replacement

$$1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i \to \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{C.1.3}$$

in the matricial expression for $M$. If the complex bundles are $U(n)$ bundles, their realifications are $SO(2n)$ bundles: At the level of the algebra, one can see that (C.1.3) sends a $n \times n$ anti-hermitian traceless matrix to a $2n \times 2n$ real antisymmetric matrix, so that it indeed defines an algebra embedding from $\mathfrak{u}(n)$ to $\mathfrak{so}(2n)$. Since $U(n)$ is a connected Lie group, the algebra embedding extends to a group embedding via the exponential mapping.

We may now take the tachyon (5.1.8) (which for each point with $|\vec{X}| = 1$ defines an element of $U(2^{p-1})$) and turn the mapping between $U(2^{p-1})$-bundles to a mapping between $SO(2^p)$ bundles over the same base. However, we now have to show that the maps constructed in this way indeed represent nontrivial elements of the $KO$ groups.

To do this, we can now complexify again the bundles (which, for a given real bundle $E_{\mathbb{R}}$, means taking the tensor product with the complex numbers $E_{\mathbb{R}} \otimes \mathbb{C}$). A generic section of the complexification of $r(E)$ is of the form

$$s_{E_x} + is'_{E_x} + s_{E_y} + is'_{E_y} = (s_{E_x} + is'_{E_y}) + i(s'_{E_x} - is_{E_y}) = e + \overline{e}', \tag{C.1.4}$$

where $E_x$ and $E_y$ are the real and imaginary parts of $E$, $s$ and $s'$ are sections of the corresponding bundles, and $e$ and $\overline{e}'$ are sections of $E$ and $\overline{E}$ respectively. Thus the complexification is isomorphic to $E \oplus \overline{E}$.

The complexification of $r(E)$ is trivial if $r(E)$ is or, in other words, $r(E)$ can only be trivial if $E \oplus \overline{E}$ is. The above discussion carries over to pairs of bundles $(E, F)$ and to $K$ theory classes in a straightforward manner, using the map from tachyon profiles to $K$-theory classes described in the main text. If $T$ represents the $K$-theory class $(E, F)$, then $T \oplus \overline{T}$ represents the $K$-theory class $(E \oplus \overline{E}, F \oplus \overline{F})$. Therefore, if the $K$-theory element represented by $T \oplus \overline{T}$ turns out to be nontrivial, then the realification of $T$ describes a nontrivial element in $KO$.

In $k = 4, 8$ dimensions, the ABS tachyon has nontrivial $(k/2)$-th Chern class. Since $c_k(\overline{E}) = (-1)^k c_k(E)$, this means that $T \oplus \overline{T}$ has nontrivial $(k/2)$-th Chern class (in fact, it is twice that of the ABS tachyon). This means that the real tachyons obtained from these indeed describe nontrivial elements of $KO$. Since the $k$-th Pontryagin class of a real bundle is defined as the $(k/2)$-th Chern class of its complexification, we see that the Pontryagin classes of these tachyons are even.

In $k = 2$ dimensions $T \oplus \overline{T}$ is a trivial bundle so the above considerations do not apply, and in $k = 1$ the ABS construction (which strictly speaking is only defined for even dimensions) does not apply. We will consider these cases in detail below. Notice that once we have constructed generators for the $KO$-theory groups from $k = 1$ to 8 we can get generators for any other $KO$-group via Bott periodicity [37].

We now turn to the explicit construction of real tachyons using the above recipe.

For $k = 4$, the ABS construction provides a tachyon, written in terms of $\Gamma$ matrices

$$T = \Gamma^\mu X_\mu \tag{C.1.5}$$

where $\Gamma^\mu$ are $4 \times 4$ $SO(4)$ Dirac matrices. However, as stated in the text, we should regard the above tachyon as a map between chiral spinor bundles $S^\pm$, which are complex 2-dimensional. This means that a more appropriate $2 \times 2$ form of the above tachyon may be obtained. Define the following matrices:

$$\sigma^\mu = (1, i\tau^1, i\tau^2, i\tau^3) \tag{C.1.6}$$
$$\overline{\sigma}^\mu = (\sigma^\mu)^\dagger = (1, -i\tau^1, -i\tau^2, -i\tau^3) \tag{C.1.7}$$

with $\tau^i$ the Pauli matrices. Then, when mapping chiral spinor bundles, (C.1.5) becomes

$$T = \sigma^\mu X_\mu. \tag{C.1.8}$$

Applying the homomorphism (C.1.3), we get the $4 \times 4$ real tachyon

$$T = r(\sigma^\mu) X_\mu = \overline{\eta}_\mu \check{X}^\mu \tag{C.1.9}$$

with $\check{X}^\mu = (X^0, X^3, -X^2, X^1)$ and $\overline{\eta}_\mu$ the set of $4 \times 4$ real matrices[1]

$$\begin{aligned}
\overline{\eta}_0 = r(\sigma^0) = \mathbf{I}_{4\times 4} \quad ; \quad \overline{\eta}_1 = r(\sigma^3) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad ; \\
\overline{\eta}_2 = -r(\sigma^2) = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad ; \quad \overline{\eta}_3 = r(\sigma^1) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{C.1.10}$$

As stated above, this tachyon has Pontryagin class equal to two (when defined over the sphere); it thus describes a two-center Taub-NUT. This is the tachyon profile used in the main text to construct the $A_1$ singular CFT.

For $k = 8$, the Dirac matrices appearing in the expression of the ABS tachyon are 16-dimensional, which translates into 8-dimensional chiral spinor bundles. The homomorphism (C.1.3) turn these to real $16 \times 16$ matrices. In 8 dimensions, one can take a Majorana-Weyl condition of spinors. If the original matrices used in the ABS tachyon are in this representation, then the homomorphism (C.1.3) will yield

[1] This somewhat strange definition of $\check{X}$ and the $\overline{\eta}$ matrices will arise naturally from the construction in Section C.1.1.

two irreducible 8-dimensional blocks, each of which will constitute a Majorana-Weyl representation of the $SO(8)$ Clifford algebra. This means that the soliton constructed via turning the ABS tachyon into an $SO(8)$ tachyon via (C.1.3) describes twice the generator of $KO(S^8)$, in accordance with general considerations above. In this particular case, one could directly take the ABS tachyon with real $8\times 8$ Gamma matrices and that would describe the generator of $KO(S^8)$ [37]:

$$T = X^\mu (\Gamma_{8\times 8})_\mu. \tag{C.1.11}$$

Dirac matrices in $k = 2$ dimensions can be taken as $\tau^1, \tau^2$. The ABS tachyon is then $\tau^1 X^1 + \tau^2 X^2$ which, when restricted to chiral spinors, yields simply $T_{ABS} = X^1 + iX^2 \equiv Z$. The real tachyon is then

$$T = \begin{pmatrix} X_1 & -X_2 \\ X_2 & X_1 \end{pmatrix} = |X| \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{C.1.12}$$

This tachyon has winding one around infinity. This winding cannot be undone with $SO(2)$ gauge transformations, but if embedded in $SO(n)$ for $n > 2$ (in the K-theoretic spirit) the winding can be undone in pairs. To be more specific, the tachyon specifies a closed curve in $SO(2)$, which can be embedded in $SO(n)$. A closed curve in $SO(n)$ is contractible if and only if it lifts to a closed curve in the universal cover $Spin(n)$. A curve of winding 1 lifts to a curve connecting the identity in $Spin(n)$ to minus the identity, whereas a curve of even winding lifts to a closed curve. Thus, $KO(S^2) = \mathbb{Z}_2$ with the generator being precisely (C.1.12), as desired.

For $k = 1$, the ABS construction strictly does not apply, since we are in odd dimension. Nevertheless, one can formally write an ABS tachyon $\Gamma^\mu X_\mu$, with $\Gamma^\mu$ furnishing a representation of the $SO(k)$ Clifford algebra. For $k = 1$, the complex tachyon is just $T = X$. The realification prescription tells us that the real tachyon is $x\mathbf{1}_{2\times 2}$. The determinant of this tachyon is $+1$ everywhere. The two classes of $KO(S^1) = \mathbb{Z}_2$ correspond to orientable vector bundles and non-orientable vector bundles, which is directly measured by the sign of the determinant of the tachyon. Therefore, the tachyon constructed forcing the ABS prescription in this case does not generate $KO(S^1)$.

However, the above discussion also suggests a solution, similar to what happened in the $k = 8$ case above: simply take $T = X$ as the real tachyon. The sign of the determinant indeed changes, so this corresponds to a non-orientable bundle over $S^1$. In a sense the ABS prescription still works, only that like in the $k = 8$ case it generates an order two element of the $KO$ group, which happens to vanish identically in $KO(S^1)$.

### C.1.1 An alternative embedding for the $k = 4$ instanton

In the particular case of $k = 4$, there is another natural route to arrive at the tachyon (C.1.9) in a different way, using the fact that $SO(4)$ can be double covered by $SU(2)\times SU(2)$. Along the route we will develop technology which will be essential for the arguments in Section 5.1.3.

Using the $\sigma$ matrices (C.1.6) we can construct a bijection between 4-vectors $V^\mu$ of $SO(4)$ and bispinors $\mathsf{V}_{\alpha\dot\alpha}$ (we will omit the subindices for conciseness, and call the bivector simply $\mathsf{V}$), given by

$$\mathsf{V} = \sigma^\mu V^\mu \tag{C.1.13}$$

$$V^\mu = \frac{1}{2}\mathrm{tr}\left(\overline{\sigma}^\mu \mathsf{V}\right). \tag{C.1.14}$$

We take a summation convention in which any repeated index, independent of whether it is "up" or "down", is to be summed over. A generic element $(g_1, g_2) = (e^{il_k\tau^k}, e^{ir_k\tau^k})$ of $SU(2)\times SU(2)$ acts on $\mathsf{V}$ as

$$\mathsf{V} \to g_1 \mathsf{V} g_2^{-1} = \mathsf{V} + i(l_k\tau^k\mathsf{V} - r_k\mathsf{V}\tau^k) + \dots . \tag{C.1.15}$$

Acting on the vector representation this is

$$\begin{aligned} V^\mu &\to \frac{1}{2}\mathrm{tr}\left(\overline{\sigma}^\mu g_1 \mathsf{V} g_2^{-1}\right) \\ &= V^\mu + \frac{i}{2}\left[l_i tr\left(\overline{\sigma}^\mu\tau^i\sigma^\nu\right) - r_i \mathrm{tr}\left(\sigma^\nu\tau^i\overline{\sigma}^\mu\right) + \dots\right] V^\nu . \end{aligned} \tag{C.1.16}$$

It is now convenient to introduce the 't Hooft matrices

$$\overline{\eta}_i^{\mu\nu} = \frac{i}{2}\mathrm{tr}\left(\overline{\sigma}^\mu\tau^i\sigma^\nu\right), \tag{C.1.17}$$

$$\eta_i^{\mu\nu} = \frac{i}{2}\mathrm{tr}\left(\sigma^\nu\tau^i\overline{\sigma}^\mu\right). \tag{C.1.18}$$

From the definition it is clear that these are real matrices. Explicit expressions are

$$\overline{\eta}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad ; \quad \overline{\eta}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad ; \quad \overline{\eta}_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{C.1.19}$$

$$\eta_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad ; \quad \eta_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad ; \quad \eta_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{C.1.20}$$

from which we see that these matrices are antisymmetric, and one can also check that $\overline{\eta}_i$ is anti-self-dual, and $\eta_i$ is self-dual. In this way, the mapping of $SU(2)\times SU(2)$ gives a natural construction for the $\eta$ matrices showing up in (C.1.9), while also providing expressions in terms of $\sigma$ matrices which are essential in Section 5.1.3. In terms of these matrices we have

$$V^\mu \to V^\mu + \left(l_i\overline{\eta}_i^{\mu\nu} + r_i\eta_i^{\mu\nu} + \dots\right) V^\nu . \tag{C.1.21}$$

This provides the desired embedding of the $SU(2) \times SU(2)$ element $(l_i\tau^i, r_i\tau^i)$ into the Lie algebra of $SO(4)$, which is generated by the 6 real antisymmetric matrices $\overline{\eta}$, $\eta$. More explicitly, a finite $SO(4)$ rotation $\Lambda$ can be written as

$$\Lambda = e^{l_i\overline{\eta}_i + r_i\eta_i} = e^{l_i\overline{\eta}_i}e^{r_i\eta_i} = e^{r_i\eta_i}e^{l_i\overline{\eta}_i} \tag{C.1.22}$$

using that $[\overline{\eta}_i, \eta_j] = 0$. A last couple of relations that will be useful are

$$[\overline{\eta}_i, \overline{\eta}_j] = -2\varepsilon_{ijk}\overline{\eta}_k \qquad ; \qquad [\eta_i, \eta_j] = -2\varepsilon_{ijk}\eta_k \, . \tag{C.1.23}$$

The ABS $SU(2)$ tachyon determines a representative of the generator of $K(S^4)$. The gauge connection in a representative of the form $(E, 0)$ of this K-theory class must be the famous BPST $SU(2)$ instanton. In regular gauge, this can be written as

$$A_\mu = \frac{x_\nu}{x^2 + \rho^2}\overline{\eta}_i^{\mu\nu}\tau^i \tag{C.1.24}$$

where we have fixed the instanton at the origin, denoted the instanton size by $\rho$ and picked a particular orientation in $SU(2)$. We now need to pick an embedding of this $SU(2)$ into $SO(4)$. Choosing to embed it in the first factor, we identify $l_i = \overline{\eta}_i^{\mu\nu}$. From the discussion above, the self-dual $SO(4)$ instanton in regular gauge is then given by

$$A_\mu = \frac{x_\nu}{x^2 + \rho^2}G^{\mu\nu} \tag{C.1.25}$$

with $G^{\mu\nu}$ a set of 6 $SO(4)$ matrices of the form

$$(G^{\mu\nu})^{\alpha\beta} = \overline{\eta}_i^{\mu\nu}\overline{\eta}_i^{\alpha\beta} \, . \tag{C.1.26}$$

Finite-energy considerations require that the connection must be pure gauge at infinity

$$A_\mu \to U^{-1}\partial_\mu U + \dots \tag{C.1.27}$$

with $U$ a gauge transformation. In the case of the $SU(2)$ instanton we have that [298]

$$U^{-1}\partial_\mu U = -\sigma^{\mu\nu}\frac{x_\nu}{x^2} \tag{C.1.28}$$

with $\sigma^{\mu\nu} = i\overline{\eta}_i^{\mu\nu}\tau^i$ the generator of the Lorentz group in the spinorial representation. The generalization to the vector embedding is straightforward, we have

$$U^{-1}\partial_\mu U = G^{\mu\nu}\frac{x_\nu}{x^2} \, . \tag{C.1.29}$$

In the $SU(2)$ case we can solve for $U$ in (C.1.28), obtaining

$$U = \frac{\sigma^\mu x_\mu}{\sqrt{x^2}} \equiv \hat{x}_\mu\sigma^\mu \, . \tag{C.1.30}$$

What makes this rewriting possible is the following familiar fact. Any $SU(2) = S^3$ element can be specified by three numbers $\lambda_i$, with $U = \exp(i\lambda_i\tau^i)$. But it can also be expressed as a unit quaternion $U = \beta_\mu \sigma^\mu$, as long as $\beta \cdot \beta = 1$. This gives a natural one-to-one map between the 3-sphere at infinity and $SU(2)$ (in the spinor representation).

The generalization of the exponential representation to the vector representation is straightforward

$$U = \exp(i\lambda_i \overline{\eta}_i) \tag{C.1.31}$$

with the $\lambda_i$ determined from $x_\mu$ as before. For the generalization of the quaternion representation, rewrite

$$\mathsf{x} = \hat{x}_\mu \sigma^\mu \,. \tag{C.1.32}$$

This defines an element of a $SU(2) \subset SO(4)$, as before. Acting on any vector $\mathsf{V} = V_\mu \sigma^\mu$ this acts as (in our conventions)

$$\mathsf{V} \to \mathsf{x}\mathsf{V} \tag{C.1.33}$$

which extracting components gives

$$\begin{aligned}(V')^\mu &= \frac{1}{2}\mathrm{tr}\left(\overline{\sigma}^\mu \mathsf{x}\mathsf{V}\right)\\ &= \frac{1}{2} V^\rho \hat{x}^\tau \mathrm{tr}\left(\overline{\sigma}^\mu \sigma^\tau \sigma^\rho\right)\end{aligned} \tag{C.1.34}$$

We read off the explicit mapping in this way:

$$U^{\mu\rho} = \frac{1}{2}\hat{x}^\tau \mathrm{tr}\left(\overline{\sigma}^\mu \sigma^\tau \sigma^\rho\right). \tag{C.1.35}$$

Defining $\eta_0 \equiv 1$, this can be neatly rewritten as

$$U = \hat{x}^\tau \overline{\eta}_\tau \tag{C.1.36}$$

which using $\{\overline{\eta}_i, \overline{\eta}_j\} = -2\delta_{ij}$, $\eta_0^2 = 1$ can be easily seen to lay in $O(4)$ (and explicit computation, or using that the $S^3$ is generated continuously from $\hat{x} = (1,0,0,0)$, shows that it is in $SO(4)$, i.e. $\det U = +1$).

If we want to construct the instanton based on tachyon condensation with a non-trivial tachyon profile, finite-energy considerations impose that $T$ is basically just $U$, given by (C.1.36), with the replacement $\hat{x} \to x$ (accounting for the fact that the actual tachyon vacuum for brane condensation is at $T = \infty$):

$$T = x^\tau \overline{\eta}_\tau \,. \tag{C.1.37}$$

For the supercritical closed string case we choose $X^\mu$ as our base directions, and we view the $SO(4) \times SO(4)$ symmetry as the rotation group acting on the $x, y$ dimensions. By analogy with the gauge instanton case we then take a tachyon profile of the form

$$\mathcal{T} = X^\tau (\overline{\eta}_\tau)_{ab} x^a y^b = \frac{1}{2}\mathrm{tr}\left(\widetilde{\mathsf{x}}\mathsf{X}\mathsf{y}\right) \tag{C.1.38}$$

where $\mathsf{X} = X_\mu \sigma^\mu$, $\mathsf{y} = y_\mu \sigma^\mu$ as before, and in order to display the symmetries of the system more manifestly, we have introduced $\widetilde{\mathsf{x}} = \mathsf{x}^\dagger = \widetilde{x}_\mu \sigma^\mu$, with $\widetilde{x}_\mu = (x_0, -x_1, -x_2, -x_3)$.

**Restriction to a spinor representation.** We have focused in the case of $SO(4)$, so $\mathsf{V}$ is naturally a bispinor. If we are interested just in one $SU(2)$ subgroup, for example the one acting on the left, we can think of $\mathsf{V}$ as a couple of spinors $\mathsf{V}^1$ and $\mathsf{V}^2$, by taking $(\mathsf{V}^i)_\alpha = \mathsf{V}_{\alpha i}$. Under a $SU(2)$ transformation in the left subgroup $\mathsf{V} \to g\mathsf{V}$, which implies $\mathsf{V}^i \to g\mathsf{V}^i$. Looking to the explicit form of the $\sigma^\mu$ matrices, we have $\mathsf{V}^1 = (x_0 + ix_3, ix_1 - x_2)$ and $\mathsf{V}^2 = (ix_1 + x_2, x_0 - ix_3)$. One may now show that the tachyon constructed by embedding an $SU(2)$ instanton into $SO(4)$ comes just from taking the real part of the ABS construction. The $SO(4)$ tachyon profile is

$$\mathcal{T} = X^\tau (\bar{\eta}_\tau)_{ab} x^a y^b = \frac{1}{2}\text{tr}\left[(x^a\bar{\sigma}^a)(X^\tau\sigma^\tau)(y^b\sigma^b)\right] = \frac{1}{2}\text{tr}\left[(x^a\bar{\sigma}^a)T_{ABS}(x^b\sigma^b)\right]. \tag{C.1.39}$$

Now, as above, we may regard the bispinors $(x^b\sigma^b)$ as a pair of $SU(2)$ spinors so that, as a matrix,

$$(y^b\sigma^b) = (\mathsf{V}^1_y \quad \mathsf{V}^2_y), \quad (x^a\bar{\sigma}^a) = (\mathsf{V}^1_x \quad \mathsf{V}^2_x)^\dagger = \begin{pmatrix} (\mathsf{V}^1_x)^\dagger \\ (\mathsf{V}^2_x)^\dagger \end{pmatrix}, \tag{C.1.40}$$

where $\mathsf{V}^i_x, \mathsf{V}^i_y$ are the column vectors defined above. Now expanding the product in (C.1.39), we get

$$(x^a\bar{\sigma}^a)T_{ABS}(y^b\sigma^b) = \begin{pmatrix} (\mathsf{V}^1_x)^\dagger T_{ABS}\mathsf{V}^1_y & (\mathsf{V}^1_x)^\dagger T_{ABS}\mathsf{V}^2_y \\ (\mathsf{V}^2_x)^\dagger T_{ABS}\mathsf{V}^1_y & (\mathsf{V}^2_x)^\dagger T_{ABS}\mathsf{V}^2_y \end{pmatrix}. \tag{C.1.41}$$

Now, since (assuming the $X^\mu$ are real as well)

$$\begin{aligned}\left[(\mathsf{V}^1_x)^\dagger T_{ABS}\mathsf{V}^1_y\right]^* &= (\mathsf{V}^1_x)^T T^*_{ABS}(\mathsf{V}^1_y)^* = ((\mathsf{V}^1_x)^T\tau^2)T_{ABS}(\tau^2(\mathsf{V}^1_y)^*) \\ &= (\mathsf{V}^2_x)^\dagger T_{ABS}\mathsf{V}^2_y\end{aligned} \tag{C.1.42}$$

using that

$$\sigma^*_\mu = \tau^2\sigma_\mu\tau^2 \qquad ; \qquad \bar{\sigma}^*_\mu = \tau^2\bar{\sigma}_\mu\tau^2\,. \tag{C.1.43}$$

and $\tau^2(\mathsf{V}^1)^* = i\mathsf{V}^2$. This means that (C.1.39) can be written as

$$\mathcal{T} = \text{Re}\left((\mathsf{V}^1_x)^\dagger T_{ABS}\mathsf{V}^1_y\right). \tag{C.1.44}$$

where $\mathsf{V}^1_x$ and $\mathsf{V}^1_y$ are unconstrained Weyl spinors. In other words, the real tachyon is just the real part of the complex ABS tachyon.

## C.2 The caloron solution

The ABS construction discussed above, together with its real version, provides us with a tachyon which is used in chapter 4 to describe the heterotic NS5-brane and in chapter 5 to describe the $\mathbb{C}^2/\mathbb{Z}_2$ orbifold. Via deformation we obtain a two-center Taub-NUT solution. The standard multi-center Taub-NUT metric [52] has an asymptotic circle at infinity, whose radius is formally infinite in the $\mathbb{C}^2/\mathbb{Z}_2$ orbifold and its deformations. Nevertheless, the theory at finite radius has very interesting properties which disappear in the infinite radius limit. For instance, the system is T-dual to a configuration of two coincident NS5-branes.

It is interesting to think about what happens to the ABS tachyon when we take one of the dimensions periodic. We still expect a $SO(4)$ instanton to exist in this case, but clearly (C.1.24) will not do, since it is not periodic. This problem was solved by [299], who dubbed the solution they found "caloron". The basic idea is simple: one uses the solution for instantons along a line (see for instance [300]), taken in the limit where there are infinite instantons along the direction to be compactified, with periodicity equal to the compactification radius. The resulting solution has the right periodicity.

In our case, we are directly interested in the tachyon profile, so it will be convenient to write $A = (\partial T)T^{-1}$. The caloron recipe instructs us to sum over an infinite periodic array of solutions of the form

$$A_n = (\partial T_n)T_n^{-1}, \quad T_n = \frac{(\tau - \tau_0 + n)\mathbf{I} + i\vec{\tau}\cdot\vec{X}}{\sqrt{(\tau-\tau_0+n)^2 + |\vec{X}|^2}}, \tag{C.2.1}$$

where $\tau$ is the soon-to-be compact direction and $\vec{\tau}$ the usual vector of Pauli matrices. $\tau_0$ labels the position of the caloron in the compact circle. We are interested in a periodic, pure gauge configuration (except at the locus where the tachyon is not invertible), such that the integral of the Chern-Simons three form around each point of the form $(\tau_0 + n, \vec{0})$ is one. Precisely these properties are satisfied by the configuration

$$A^c = \sum_{n=-\infty}^{\infty}\left(\prod_{k<n} T_k\right) A_n \left(\prod_{k<n} T_k\right)^{-1} = (\partial T^c)(T^c)^{-1}, \quad T^c = \prod_{n=-\infty}^{\infty} T_n. \tag{C.2.2}$$

Thus, the tachyon we are looking for is precisely $T^c$. It is easy to evaluate it by noting from (C.2.1) that, given $\vec{X}$, the matrices $\mathbf{I}, i\vec{\tau}\cdot\vec{X}/|\vec{X}|$ form a representation of the algebra of the complex numbers. This turns the definition of $T^c$ into an infinite product of complex numbers of norm 1, whose argument $\theta^c$ is readily evaluated as follows. Since we are only interested in the argument, we can compute it changing the normalization of the $T_n$. Since

$$T_n = \frac{n+z}{|n+z|}, \quad \text{with} \quad z \equiv \tau - \tau_0 + i|\vec{X}|, \tag{C.2.3}$$

we may write

$$\theta^c = \arg\left[\prod_{n=-\infty}^{\infty}\frac{n+z}{|n+z|}\right] = \arg\left[z\prod_{n=1}^{\infty}\left(\frac{n+z}{n}\right)\left(\frac{-n+z}{n}\right)\right] = \arg\left[z\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right)\right], \tag{C.2.4}$$

where in the last line we have changed the sign of every term in the product. This may introduce a sign in the final result for $T_c$ which is in any case irrelevant, as we will argue below. The product in the last term is the Weierstrass product form of $\sin(\pi z)/(\pi z)$, so that

$$\theta^c = \arg\left[\frac{\sin(\pi z)}{\pi}\right]. \tag{C.2.5}$$

From this, one can construct the periodic caloron tachyon

$$\begin{aligned} T^c = e^{i\theta^c \vec{\tau}\cdot\frac{\vec{X}}{|\vec{X}|}} &= \cos(\theta^c)\,\mathbf{I} + i\sin(\theta^c)\,\vec{\tau}\cdot\frac{\vec{X}}{|\vec{X}|} \\ &= \frac{\tan(\pi(\tau-\tau_0))\,\mathbf{I} + i\tanh\left(\pi|\vec{X}|\right)\vec{\tau}\cdot\frac{\vec{X}}{|\vec{X}|}}{\left[\tanh^2\left(\pi|\vec{X}|\right)+\tan^2(\pi(\tau-\tau_0))\right]^{\frac{1}{2}}} = \frac{T^c_\mu}{|T^c_\mu|}\sigma^\mu. \end{aligned} \tag{C.2.6}$$

with $T^c_\mu = \left\{\tan(\pi(\tau-\tau_0)), \tanh\left(\pi|\vec{X}|\right)\frac{\vec{X}}{|\vec{X}|}\right\}$. In writing the expression in terms of tangents one needs to make a choice of branch cut in the square root in (C.2.6), we have chosen the sign that agrees with (C.2.5) on $|\tau-\tau_0| \leq \frac{1}{2}$. Relatedly, notice that $\tau \to \tau+1$ sends $\theta_c \to \theta_c + \pi$, or equivalently $T^c \to -T^c$. When we view the caloron as a configuration in $SO(4)$ Yang-Mills theory this is fine, since the physical object, the connection $A^c$, is invariant under this transformation. In the context of the supercritical string the tachyon $T^c$ is physical, transforming in a bifundamental of $SO(4)\times SO(4)$, and $T^c \to -T^c$ is simply a gauge transformation in the center of the gauge group.

An interesting point is that on general grounds we expect the endpoint of the condensation (in type 0) with tachyon (C.2.6) to be an ordinary two-center Taub-NUT space. This solution has a $U(1)$ isometry along the compact direction which appears to be broken explicitly (C.2.6). Indeed, (C.2.4) and therefore (C.2.6) have a free parameter $\tau_0$, which corresponds to moving the center of the instanton in the compact direction. Thus, changing $\tau_0$ either corresponds to an irrelevant perturbation of the CFT, or describes some deformation of the ordinary Taub-NUT space. Evidence for the latter comes from noticing that if we separate the two centers of the Taub-NUT space, the theory is no longer singular. In the UV, we could imagine separating the coincident centers of the Pontryagin number 2 solution constructed above into two configurations with Pontryagin number one. Presumably this would correspond to tuning the difference $\tau_0^1 - \tau_0^2$ of the $\mathbb{S}^1$ position of the two Taub-NUT centers away from zero. In the IR, the modulus which desingularizes the CFT is $B$ field on the vanishing 2-cycle of the configuration. The parameter $\tau_0 = \tau_0^1 + \tau_0^2$ of our tachyon profile therefore seems to correspond to turning a $B$-field on the center-of-mass normalizable two-form in the IR.

# D

# Review of axion monodromy

Axion monodromy models are unique in their ability to mix the suppression of unwanted operators unwanted characteristic of axion models and the avoidance of trouble with instantons via a small axion decay constant. This appendix is devoted to a more in-depth review of axion monodromy than that given in Chapter 8, and from a slightly different perspective

We start with the construction of monodromic models, paying special attention to details such as the fate of the discrete axion symmetry.

## D.1 The Kaloper-Sorbo lagrangian

Our starting point is the Kaloper-Sorbo action[264, 278]

$$S_{KS} = \int_M \frac{1}{2}|d\phi|^2 - \frac{1}{2}|F_4|^2 + m\phi F_4 + \int_{\partial M} C_3 \wedge (*F_4 - m\phi), \quad F_4 \equiv dC_3 \tag{D.1.1}$$

where $M$ is spacetime. This system has two gauge symmetries:

$$\phi \sim \phi + 2\pi f, \quad C_3 \sim C_3 + d\lambda_2 \tag{D.1.2}$$

The boundary part of (D.1.1) is composed of two terms. The first term is essential for the variational problem of $F_4$ alone (when $m = 0$) to have a solution with variations of $dC_3$ vanishing at the boundary. Variations not satisfying this would have infinite action and not contribute to the path integral anyway. The second term is necessary to guarantee invariance of the action under the discrete axion shift symmetry.

The discrete gauge symmetry of the axion of (D.1.2) is just telling us that $\phi$ is not valued on a real line but rather on a circle of radius $f$. We find most natural working on the covering space of the circle (therefore taking $\phi$ as a real-valued coordinate) and only imposing gauge equivalence at the end. This will simplify our computations a little bit. Of course, one could just take $\phi$ to be circle-valued from the start and obtain the same results.

Gauge invariance of $C_3$ kills 3 degrees of freedom of $C_3$. This can be seen as follows: $C_3$ has four different polarizations. The two-form gauge parameter $\lambda_2$ has

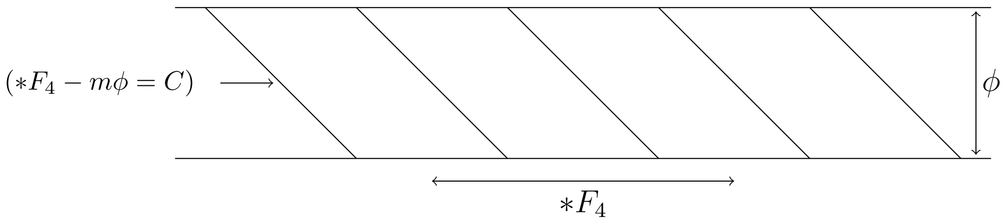


**Figure D.1:** Configuration space of the Kaloper-Sorbo model. It contains two parameters: The (dual of the ) 4-form field strength, which is real-valued, and the periodic scalar $\phi$. The equation of motion for $F_4$ is actually a constraint which forces all the physics to take place on the line $*F_4 - m\phi = C$.

six, although three of these are of the form $\lambda_2 = d\lambda_1$ and therefore do not change $C_3$. Therefore by an appropriate choice of $\lambda_2$ we can kill 3 components of $C_3$.

The fourth component is actually eliminated by the equation of motion of $C_3$. We can find it by plugging an arbitrary variation $C_3 + \delta C_3$ (with $d(\delta C_3) = 0$ at $\partial M$) in (D.1.1) and seeing how the action changes[1]

$$\delta S = \int_M d(*F_4 - m\phi) \wedge \delta C_3. \tag{D.1.3}$$

Therefore the equation of motion is (at least if $H_1(M, \mathbb{Z}) = 0$)

$$*F_4 - m\phi = C \tag{D.1.4}$$

where $C$ is a (so far unspecified) constant. Actually, (D.1.4) is not an equation of motion. It is a constraint, which has to be imposed on every state in order for them to be physical. In fact, it is the 3-form equivalent of Gauss' law of electrodynamics. The constraint removes the final degree of freedom of $C_3$ and we are left with only one degree of freedom, the scalar $\phi$. This is depicted schematically in figure D.1.

Since the equation of motion is just a constraint, finding the effective action for $\phi$ is very simple: one just has to plug back (D.1.4) in the action (D.1.1). Before doing that however, it is illustrative to introduce the linear combinations

$$\xi_\pm = \frac{1}{2}\left(\phi \pm \frac{*F_4}{m}\right). \tag{D.1.5}$$

In terms of these coordinates, (D.1.1) takes the form form

$$S_{KS} = \int_M \frac{1}{2}|d(\xi_+ + \xi_-)|^2 + \frac{m^2}{2}|\xi_+ - \xi_-|^2 - m^2(\xi_+^2 - \xi_-^2) - 2m\int_{\partial M} C_3 \wedge \xi_-. \tag{D.1.6}$$

The constraint (D.1.4) is now simply that $\xi_-$ is a constant which we will call $c$. This allows us to replace the boundary term $-2m\int_{\partial M} C_3 \wedge \xi_-$ by a bulk term $-2m\int_M F_4 \wedge \xi_- = 2m^2 \int_M *(\xi_+ - \xi_-) \wedge \xi_-$, with the final result

$$S_{KS} = \int_M \frac{1}{2}|d\xi_+|^2 - \frac{m^2}{2}|\xi_+ - c|^2. \tag{D.1.7}$$

[1] We remind the reader that in four dimensions with Lorentzian signature, $*dV = 1$, $*1 = -dV$.

We therefore see that the effect of the constraint is to give a quadratic potential for the coordinate $\xi_+$. In monodromy inflation models, the actual inflaton is $\xi_+$, not $\phi$. If we mistook $\xi_+$ by $\phi$, it would seem that the discrete symmetry has been broken. This is clearly impossible as this symmetry is gauged (it is just an artifact created by the fact that $\phi$ is naturally a coordinate on the covering space of the axion circle). In fact, one can compute its action in $\xi_\pm$ coordinates to be to identify

$$(\xi_+, \xi_-) \sim (\xi_+, \xi_-) + \pi f(1,1). \tag{D.1.8}$$

Therefore, the discrete shift symmetry corresponds to shifting both $\xi_+$ and $c\ \pi f$ in the above description. This is not an actual symmetry of the theory with action (D.1.7); a symmetry would act on the fields while leaving the parameters of the lagrangian invariant. This is rather a coordinate change, and in fact one can (in the classical theory) replace the shift by $\pi f$ for any other real number. Note also that, unlike $\phi$, $\xi_+$ is actually real-valued: The discrete shift symmetry identifies the line $\xi_- = c$ with $\xi_- = c + \pi f$ but along a given line points with different $\xi_+$ are identified. This is illustrated in figure D.1.

Once one has implemented the constraint, it does not matter to use $\xi_+$ or $\phi$ (regarded as a coordinate in the covering space) as a coordinate, as it is commonly done in monodromy inflation models; we have introduced $\xi_+$, which is clearly not periodic even before imposing the constraint, to emphasize that the field range is actually $\mathbb{R}$. We can forget about the discrete symmetry the moment we impose the constraint, as depicted in figure D.1.

The classical picture therefore is that the theory is just that of a massive, real-valued scalar field $\xi_+ - c$, with an arbitrary value of $c$. There are not infinitely many different "copies" of the potential; all of them are really the same, being identified by a gauge symmetry. By combining the axion $\phi$ with the (dual of the) 4-form $F_4$ we have removed the constraint posed by the axion discrete symmetry on the potential. It is easy to understand what is going on in figure D.1; The field $\xi_+$ varies along the tilted lines which, thanks to the identification provided by the axion discrete shift symmetry, are actually infinitely long. As we move towards $m = 0$ the lines degenerate to vertical lines; when $m$ vanishes we have infinitely many separate, compact branches.

## D.2 Kaloper-Sorbo protection

It is commonly argued that monodromy protects the potential of $\xi_+$ against corrections. We can see how this happens exactly. In (D.1.1) we have assumed a canonical kinetic term for $F_4$, but this need not be the case. For a more general kinetic term $-K(|F_4|^2)$ the derivation above remains unchanged[2], to yield a real-valued scalar with potential

$$V_{KS}^K = m|\xi_+ - c|(K')^{-1}(m|\xi_+ - c|) - K((K')^{-1}(m|\xi_+ - c|)). \tag{D.2.1}$$

[2] We just replace $*F_4$ by $\wedge * K'(F_0)$ in the boundary term and in the definition of $\xi_\pm$.

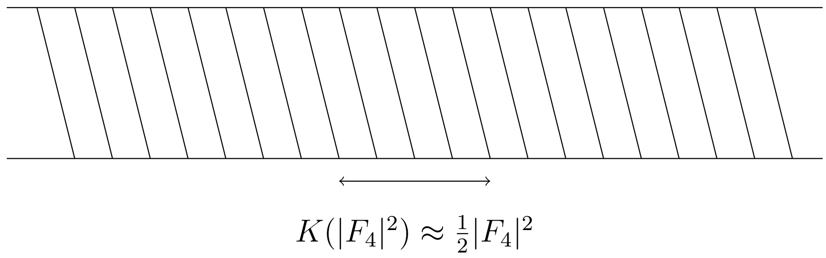


**Figure D.2:** When $m$ is very small, the line allowed by the constraint (D.1.4) is almost vertical. As a result, it winds many times around the axion circle before moving too far along $*F_4$; within the region spanned by the double arrows, the potential for $\xi_+$ can be approximated as quadratic, even if it traverses a transplanckian distance.

Thus, corrections coming from a more general kinetic term are always a function of the original potential $\frac{1}{2}m^2|\xi_+ - c|^2$. If we expand $K$ on a power series where higher order corrections are suppressed by powers of some cutoff $\Lambda$, we have

$$V_{KS}^K = \left(\frac{1}{2} - \sum_{n=2}^{\infty} k_n \left(\frac{m}{\Lambda^2}|\xi_+ - c|\right)^{2(n-1)}\right) m^2|\xi_+ - c|^2. \tag{D.2.2}$$

Higher order terms become relevant not when $\xi_+ - c \approx \Lambda$, as might have been expected, but rather at the much higher value $\xi_+ \sim \frac{\Lambda}{m}\Lambda$. In this way, if $m \ll \Lambda$ and $\Lambda \sim M_P$, monodromy models are able to attain transplanckian field ranges. What is happening physically is shown in figure D.2. As $m \to 0$, one can travel an arbitrarily large distance along $\xi_+$ without straying too far along $F_4$; if the potential comes solely from the nontrivial kinetic term $K(|F_4|^2)$, it is a very good approximation to replace it by the canonical term $\frac{1}{2}|F_4|^2$, and thus by the quadratic potential.

## D.3 The quantum theory

Here we briefly recount how the above picture is affected by quantum effects. In the classical theory, the parameter $c$ is unconstrained, although once it is fixed, it is fixed for good; it is part of the boundary conditions of the theory and to change it would require infinite energy. This changes in the quantum theory in the presence of membranes, which couple to the 3-form field as

$$kq\int_X C_3, \tag{D.3.1}$$

where $k$ is an integer and $X$ is the membrane worldvolume. The membrane with lowest charge, $k = 1$, defines the three-form gauge coupling $q$. Assuming that membrane charge is quantized in units of $q$ is equivalent to saying that the three-form gauge group is compact; it is $U(1)$ rather than $\mathbb{R}$. Noncompact gauge groups generically lead to trouble with quantum gravity [42], arguably even in the case of higher form potentials [301]. We will assume that the membrane gauge group is compact in what follows.

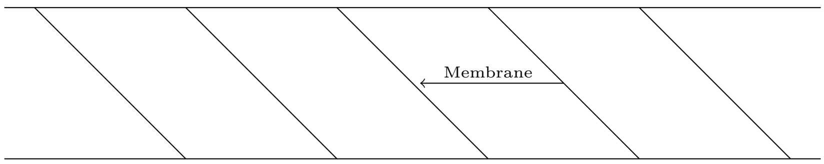


**Figure D.3:** A membrane is an electric source for $C_3$; therefore, the value of $*F_4$ shifts by $q$ as one crosses the membrane. When $n = 1$, this is equivalent to the transition depicted in the figure, from one branch to itself. In inflation one would like the inflaton $\xi_+$ to gently slow roll along the potential; too many of these membrane transitions could spoil this behavior.

If the membrane worldvolume $X$ wraps a nontrivial 3-cycle (for instance, if there is a nontrivial gravitational instanton in the theory, as discussed in [276]), then there is the possibility of taking large gauge transformations for $C_3$

$$C_3 \to C_3 + \frac{2\pi}{q}\Lambda_3, \quad \Lambda_3 \in H_3(M, \mathbb{Z}). \tag{D.3.2}$$

Large transformations are quantized in units of $2\pi/q$ so as to guarantee that (D.3.1) shifts by an integer multiple of $2\pi$, rendering the path integral well-defined. But then we can also perform large gauge transformations on the boundary term of (D.1.1). Since the periodicity of these is already fixed, requiring that the shift in the action is an integer multiple $k$ of $2\pi$ amounts to

$$c = \frac{kq}{2m}, \tag{D.3.3}$$

that is, the value of $c$ is quantized in units of $\frac{q}{2}$. The axion shift symmetry identifies $c \sim c + \pi f$; this imposes an extra condition

$$nq = 2\pi m f, \quad n \in \mathbb{N} \tag{D.3.4}$$

which was heavily used in Chapter 8 to constrain relaxion models. The integer $n$ in (D.3.4) corresponds to a $\mathbb{Z}_n$ discrete symmetry in the theory, precisely the one discussed in [232]. Notice that all the curves $\xi_- = (k + nl)\frac{q}{2m}$, with $l \in \mathbb{Z}$, are all gauge copies of one another; thus, while in the classical theory infinitely many lines were possible, quantum-mechanically only $n$ different branches are allowed; the full theory consists of $n$ copies of the classical theory, or "branches", with membranes mediating transitions among these. The $\mathbb{Z}_n$ symmetry is spontaneously broken by the choice of branch. We will take $n = 1$ from now on, so that only one branch is present. All the effects that could spoil monodromy (bubble nucleation, nonperturbative contributions to the potential) become stronger as $n$ increases; therefore $n = 1$ is the most conservative option.

As is well-known, membranes can mediate tunneling between different values of $\xi_+$ by changing the value of $*F_4$, as depicted schematically in figure D.3. The nucleation rate of these bubbles has been shown in several cases [157, 235] not to be large enough so as to constitute an issue for monodromic inflation: the chances of nucleating a bubble within our horizon during inflation are simply too low. Interestingly, this is not the case for the relaxion scenario (see Chapter 8), which shares many similarities with models of monodromy inflation.

## D.4 Axion monodromy in the dual 2-form view

The system described in Section D.1 admits an alternative description in which the scalar is dualized into a 2-form. Let us recall the duality in the absence of monodromy. Consider an axion field with action

$$\frac{1}{2}\int d\phi \wedge *d\phi. \tag{D.4.1}$$

As is well-known, the above theory is dual to that of a massless 2-form field $B$, as follows (see e.g. [275]). The path integral for the axion field is equivalent to the path integral of a closed 1-form $v$. This is described by the action

$$S = \int \frac{1}{2} v \wedge *v + B_2 \wedge dv. \tag{D.4.2}$$

where $v$ is an unconstrained 1-form. Integrating out $B$, it acts as a Lagrange multiplier imposing $dv = 0$, and by then setting $v = d\phi$ we recover the original action. If on the other hand we integrate out $v$ we obtain the action

$$S = \frac{1}{2}\int dB_2 \wedge *dB_2. \tag{D.4.3}$$

The Kaloper-Sorbo proposal allows us to extend this duality framework for the case of an axion with a potential. Consider the particular case of an scalar field with a mass term

$$\frac{1}{2}\int d\phi \wedge *d\phi - \frac{1}{2}m^2\phi^2. \tag{D.4.4}$$

To establish the duality [275], notice that the path integral for massive axion action can be rewritten (at least in flat space) as

$$\frac{1}{2}\int d\phi \wedge *d\phi + g\phi F_4 - \frac{1}{2}F_4 \wedge *F_4. \tag{D.4.5}$$

Repeating the argument above for the axion, we then get

$$S = \int \frac{1}{2}\left|dB_2 - gC_3\right|^2 - \frac{1}{2}F_4 \wedge *F_4. \tag{D.4.6}$$

This describes a massless three-form $C$ becoming massive by eating up a massless two-form $B$. Note the gauge invariance

$$C_3 \to C_3 + d\Lambda_2 \quad , \quad B_2 \to B_2 + g\Lambda_2 \tag{D.4.7}$$

The relationship (D.3.4) also arises in the dual formulation. The operator

$$\exp\left(2\pi i f\int_{\partial\Sigma} B_2 - k\Lambda_k^2\int_{\Sigma} C_3\right) \tag{D.4.8}$$

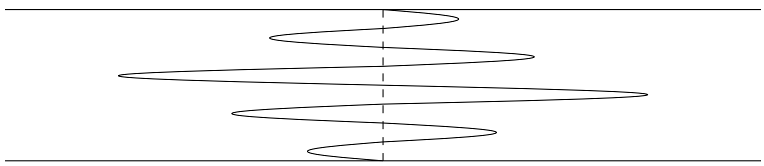

**Figure D.4:** There is no reason for the constraint equation $m = 0$ to be just a straight, vertical line; it can wiggle as one moves around $\phi$. If the wiggles are too large the inflaton will stray too far away on $F_0$ in one period, reaching the region where correcctions to the kinetic term of the 4-form are important and we lose perturbative control.

is gauge-invariant both under large $B$-field gauge transformations and $C$ gauge transformations, provided that $k\Lambda_k^2 = 2\pi g f$. Physically the operator describes the contribution of an euclidean membrane instanton wrapping $\Sigma$ and ending in an worldsheet instanton wrapping $\partial\Sigma$. $k\Lambda_k^2$ is then the membrane charge, which by Dirac quantization is an integer multiple $k$ of the fundamental 3-form coupling, $\Lambda_k^2$, as we saw in Section D.3.

This formulation, in which the strings charged under $B_2$ manifestly have to be boundaries of membranes charged under $C_3$ because of gauge invariance, also allows us to connect the membranes charged under $C_3$ with the field theory bubbles mentioned in Chapter 8. Imagine a flat membrane with a hole in it. One can cross the membrane to achieve a shift of $q$ in $F_0$, or one can cross the hole, which amounts to a shift of $2\pi f$ in the axion. However, as we have discussed, the two are related by the constraint (D.1.4), so that both changes are physically equivalent. It follows that the profile of the axion around the region of the hole is precisely a "kink" field theory solution just like the bubble walls of the field theory bubbles in (8) and discussed in [285].

## D.5 Generalizing the KS lagrangian

In Section D.2, we saw where the Kaloper-Sorbo protection comes from. The constraint

$$K'(F_0) = m\phi \tag{D.5.1}$$

generates almost vertical lines, for $m$ small. Therefore one is protected from dangerous excursions to large values of $F_0$.

However, this only protects the potential as long as one can guarantee that when $m = 0$ the lines are indeed vertical. $m$ is the shift-symmetry breaking parameter; we assume that when $m \to 0$ monodromic effects disappear. The lines defined by the constraint are then forced to be periodic, but nothing prevents them from wiggling dangerously across different values of $F_4$. This is illustrated in figure D.4. This would correspond to a constraint

$$F_0 = \Lambda^2 P'(\phi/f, F_0), \tag{D.5.2}$$

where $\Lambda$ is some energy scale and $P'$ is a periodic function of its first argument, with period $2\pi$. At the lagrangian level, this would correspond to adding a term

$$\Lambda^2 P(\phi/f, F_0) \tag{D.5.3}$$

to (D.1.1), where $P$ is the primitive of $P'$ with respect to $F_0$. Notice that the boundary term must be modified accordingly. Such a term is perfectly fine with all the symmetries of the theory and hence it will generically be there. For practical purposes, the term $\Lambda^2 P'$ just renormalizes $m$ to $m + \Lambda^2 P'$. Via the equation $q = 2\pi m f$, it will also affect the coupling of membranes, which now have a $\phi$-dependent gauge coupling.

Both the appearance of a new scale $\Lambda$ and the fact that the function $P$ must respect the axion shift symmetry suggest that these terms come from novel non-perturbative phenomena. Notice however that, at least in principle, the effects described by (D.5.3) include, but are not restricted to, ordinary corrections to the kinetic term of the 4-form.

Let us now explore what terms show up now in the effective potential of inflation. For simplicity, we will take $P'$ to be independent of $F_0$, something like $P' = \Lambda^2 \cos(\phi/f)$, and we return to canonical kinetic terms for $F_4$, for simplicity. The constraint is now $F_0 = m\phi + \Lambda^2 \cos(\phi/f)$. Plugging back in (D.1.1), we get a potential (for $\Lambda$ small)

$$V = \frac{1}{2}\left(m\phi + \Lambda^2 \cos(\phi/f)\right)^2 \tag{D.5.4}$$

which includes terms $\phi \cos(\phi/f)$ not typically discussed in the monodromic literature. These terms can totally spoil the potential for transplanckian inflation if $\Lambda$ is comparable to $\sqrt{mM_P} \approx 10^{-3} M_P$. We need $\Lambda$ to be much smaller than that.

What kind of effects can generate a nontrivial $P'$? Generically, we expect to have instantons which couple to $\phi$ via terms of the form

$$\frac{i}{f}\int_p \phi. \tag{D.5.5}$$

In the dilute-gas approximation, these will generate a potential for $\phi$. These instanton effects are known not to be an issue for monodromy inflation models because, even though they are not constrained by KS-like arguments, their contributions to the potential are exponentially suppressed. If we use bounds to the instanton action coming from the Weak Gravity Conjecture, this suppression is enormous.

However, this corresponds to a $P$ above with no $F_0$ dependence. More general possibilities should be studied. We need to consider the coupling to instantons to $F_0$ via the $\phi F_0$ term; perhaps bound state of an euclidean membrane, wrapping a pointlike instanton which sits at its center (see figure D.5), if it exists. The membrane couples to $C_3$ as in (D.3.1), while the instanton couples to $\phi$ as (D.5.5). This would-be bound state would couple as

$$iq\int_{S^3} C_3 + \frac{1}{f}\int_p \phi \approx qV_4 F_0 + \frac{\phi}{f}, \tag{D.5.6}$$

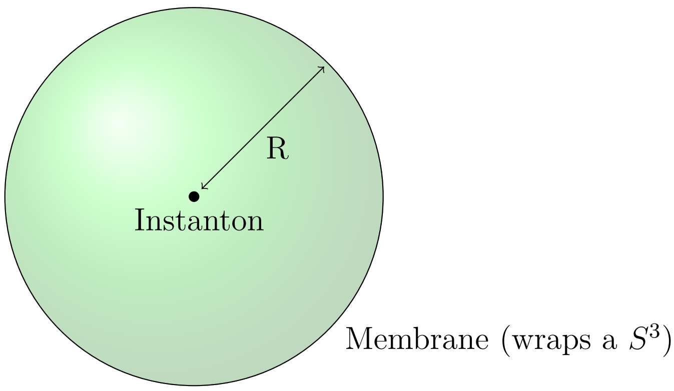


**Figure D.5:** If it were possible to have a bound state of an membrane wrapping a $S^3$ of finite radius centered around a pointlike instanton which couples to $\phi$, we could generate a nontrivial $P'(\phi, F_0)$.

where we have assumed that fluctuations of $F_0$ do not vary significantly within the membrane. In the dilute-gas approximation, this would give a potential

$$\mathrm{Pr}e^{-S} \cos\left(qV_4 F_0 + \frac{\phi}{f}\right), \tag{D.5.7}$$

where Pr is the usual uncomputable instanton prefactor. In any case, we can probably take $f$ of order $S \sim M_P/f$, and thus these effects are suppressed whenever the usual wiggles on top of the Kaloper-Sorbo potential are.

## D.6 Some quirks of monodromy

Here we discuss some more peculiarities of monodromic models.

### D.6.1 Gaps between branches?

One might be tempted to say that the different branches of monodromy (meaning the different disconnected segments in figure D.1) can become disconnected from one another, in analogy to the mass gaps of solid state physics [279]. This comes from the misconception of parametrizing the constraint $F_0 = m\phi + C$ using $\phi$ and thinking that it is somehow still periodic. As argued in Section D.1, it is not.

To make gaps appear in figure D.1, we would have to change the homology class of the 1-cycle wrapped by the constraint. Although changes in topology are common in string theory, it cannot happen in the effective field theory we are describing, since the line described by the constraint would somehow have to “break”. Hence, barring some important stringy effects which are beyond the scope of this discussion, no gap will develop between branches.

### D.6.2 Correction of the KS-potential and modulations

From the above discussion it would seem that there is no clear distinction between KS-protected terms in the potential and those who are not. For instance, a term $m^2\phi^2$ can have corrections, as a term e.g. $\Lambda^2 F_0 \sin(\phi/f)$ can be expanded in series to give a contribution $\Delta m = \frac{\Lambda^2}{f}$.

However, these corrections are an artifact of the modulations. For small $m$, the constraint is

$$F_0 = m\phi + C + \Lambda^2 P'(\phi). \tag{D.6.1}$$

If we run over a transplanckian field range $\Delta\phi$, the effective value of $m$ is

$$m_{\text{eff}} = m + \Lambda^2 P'(\Delta\phi/f) \tag{D.6.2}$$

so that the corrections to $m$ are small whenever the modulations are small and vice-versa. In other words, a large modulation washes out the KS protection of the potential, but if it is small, then KS does a great job in protecting it.

## D.7 An example: Monodromy from Stuckelberg

In this Section we look at monodromy in 4d arising from a Stuckelberg lagrangian in 5d, as an interesting example. The lagrangian is

$$\frac{1}{2}|dC_0 - f_{C_0} e_{5d} A|^2 + \frac{1}{2}|F|^2 \tag{D.7.1}$$

We can get monodromy from this via dimensional reduction on a circle of radius $R$, as follows. Let $dC_0 \equiv \frac{1}{\sqrt{2\pi R}} F_0 d\theta$ and let $f^{-2} \equiv 2\pi R e_{5d}^2$, $A = (2\pi R)^{-1/2}\phi d\theta$. Then upon dimensional reduction we get

$$\frac{1}{2}|d\phi|^2 - \frac{1}{2}\left|F_0 - f_{C_0} e_{5d}\phi\right|^2 \tag{D.7.2}$$

from which we can read the monodromy parameter $m \equiv f_{C_0} e_{5d}$. The gauge symmetry of the higher-dimensional Abelian Higgs model gets mapped to

$$F_0 \to F_0 + 2\pi f_{C_0} e_{5d} f, \quad \phi \to \phi + 2\pi f \tag{D.7.3}$$

from which we see that we have actually arrived at something similar to the $\xi_\pm$ coordinates above. Notice that we have not yet imposed the equations of motion for $F_0$.

Let us look at what nonperturbative effects might arise from weak gravity objects in 5d. The electric WGC gives us a 5-d particle of mass

$$m_{WG5} = e_{5d} M_5 = \frac{M_P}{f}\frac{1}{2\pi R} \tag{D.7.4}$$

which upon integrating over the compact dimension gives an instanton of action $M_P/f$, as predicted directly by weak gravity in 4d. Here we have used the relationship between 5d and 4d Planck masses, $M_5 = \frac{M_P}{\sqrt{2\pi R}}$.

On the other hand, we might have instantons charged under $C_0$. In the Stuckelberg model, these have to spit out a electrically charged particle for consistency. The potential coming from instantonic effects is generated by configurations with a non-zero net instanton charge which asymptote to the vacuum at infinity, but in the Stuckelberg model this can only be the case if each electric particle emitted by an instanton is absorbed by another, with the result that there is no potential for $C_0$ in this case. This makes sense with the fact that, from the point of view of the 4d theory, the five-dimensional $C_0$ is morally a "$(-1)$-form", dual to the 4d 3-form, for which there is no object coupling magnetically.

We can also consider the magnetic WGC for the $U(1)$ in 5d. This tells us that there is a hidden scale $\Lambda \sim \frac{M_P}{f}\frac{1}{2\pi R}$. This is just the mass of the monopole of the massive 4d $U(1)$.

Finally, we should also consider membranes coupled to $C_0$ in 5d, which should become the membranes which change $F_0$ in 4d. The tension of these membranes is again set by the WGC

$$T = f_{C_0} M_5 = m f M_P \tag{D.7.5}$$

which is the standard 3-form WGC.

# E

# Analysis of the bubbles

Here we study in detail several aspects of 4-form membrane and bubble physics which are important for the analysis in Chapter 8. Specifically, we derive the tunneling probability formulae such as (8.4.4), and analyze the nucleation and growth of the bubbles, emphasizing the cosmological effects which take over when the bubble radius is comparable to the de Sitter radius.

## E.1 Coleman-DeLuccia formulae

We are interested in computing the nucleation probability in the thin-wall approximation for a membrane of tension $T$ which transitions from a vacuum with $V_i$ to another with $V_f$, $V_f < V_i$ and both positive. This is $e^{-B}$, and for $B$, there is a slightly complicated formula by Coleman and DeLuccia [218],

$$B = 2\pi^2 T r^3 + \frac{12\pi^2}{\kappa^2}\left\{\frac{1}{V_f}\left[(1-\frac{1}{3}\kappa V_f r^2)^{\frac{3}{2}}-1\right]-\frac{1}{V_i}\left[(1-\frac{1}{3}\kappa V_i r^2)^{\frac{3}{2}}-1\right]\right\}, \tag{E.1.1}$$

where $\kappa \equiv 8\pi G$, and $M_P = G^{-1/2}$. One is supposed to minimize $B$ with respect to $r$, and then the bubble nucleation rate is $\exp[-B(r_{min})]$. The expression simplifies for $V_i = 0$ and it has been used recently in the literature of WGC constraints to large field inflation [157].

However, we are interested in a different case, since for us $V_f = V_i - 2\pi f g^2 \phi_i$ and we assume $V_f - V_i$ small. Note that both $V_f$ and $V_i$ include contributions from the inflationary potential, and are indeed dominated by them. As a result, we will take

$$\frac{\kappa}{3}V_f \approx \frac{\kappa}{3}V_i \approx H^2 \tag{E.1.2}$$

in terms of the Hubble constant during inflation.

Let us now find the extrema of $B(r)$. The condition for stationary point is, discarding the trivial one at $r = 0$,

$$-\sqrt{1-r^2\Lambda_f} + \sqrt{1-r^2\Lambda_i} + \frac{\kappa T}{2} r = 0 \tag{E.1.3}$$

where following Coleman's notation, $\Lambda_i = (\kappa V_i)/3$ and analogously $\Lambda_f$ are the corresponding cosmological constants.

We now solve this equation carefully. Squaring once, we get

$$\gamma r = -\kappa T \sqrt{1-r^2\Lambda_i}, \quad \gamma \equiv \left(\frac{(\kappa T)^2}{4} + \Lambda_f - \Lambda_i\right). \tag{E.1.4}$$

Existence of nontrivial solutions now depends crucially on $\gamma < 0$. If this is not the case, the would-be bubble would have a radius so large it would extend beyond the cosmological horizon. Fortunately, in relaxation we have $\Lambda_f - \Lambda_i = -\frac{\kappa}{3} 2\pi f g^2 \phi_0$ with $\phi_0 = M^2/g$. For this to overcome the WGC membrane tension $(\kappa T)^2 \sim (fg/M_P)^2$, we must have $\phi_0 > f$, the original vev of the relaxion must be larger than the fundamental decay constant, which is always satisfied in any reasonable relaxionic model.

One may wonder what happens for more generic monodromy models. For instance, [228] provide examples of monodromic models which have a linear potential at large fields. For a generic potential, the requirement that bubbles can form is that the change $\Delta V$ in the potential upon crossing of a membrane $\Delta V \gtrsim (fg)^2$. Parametrizing a linear potential as $X^2 g\phi$, for some mass scale $X$, we get a condition $\phi_0 > f\frac{M}{X}$, which is only troublesome if $X \ll M$. However, both for the original relaxation proposal and the monodromic models in Chapter 8, we take $X = M$. To do otherwise means introducing a new mass scale, $X$, very different from $M$. In a sense, this shifts the problem from explaining the hierarchy between $m_W$ and $M$ to explaining the hierarchy between $X$ and $M$.

We now may take squares in (E.1.4) to obtain

$$r_{min}^2 = \frac{1}{\left(\frac{\gamma}{\kappa T}\right)^2 + \Lambda_i} \tag{E.1.5}$$

Plugging back in (E.1.1) we get our exact expression

$$B = 2\pi^2 \left( \frac{2\left(\left(\frac{1}{\frac{\kappa^2 T^2 \Lambda_f}{\tilde{\gamma}^2}+1}\right)^{3/2} - 1\right)}{\kappa \Lambda_f} + \frac{T}{\left(\Lambda_i + \frac{\gamma^2}{\kappa^2 T^2}\right)^{3/2}} - \frac{2\left(\left(\frac{1}{\frac{\kappa^2 T^2 \Lambda_i}{\gamma^2}+1}\right)^{3/2} - 1\right)}{\kappa \Lambda_i} \right) \tag{E.1.6}$$

where we have defined

$$\tilde{\gamma} \equiv \left(\frac{(\kappa T)^2}{4} + \Lambda_i - \Lambda_f\right). \tag{E.1.7}$$

When $\Lambda_i \to 0$, it reduces to the transition-to-Minkowski limit

$$B \to \frac{216\pi T^4}{V(4V + 3\kappa^2 T^2)^2}. \tag{E.1.8}$$

We are instead interested in a different limit in which the difference $\Delta\Lambda = \Lambda_i - \Lambda_f$ is very small compared with either $\Lambda_i, \Lambda_f$, while also being very large compared to $(\kappa T)^2$. The parameter

$$p \equiv \frac{\kappa T}{\sqrt{\Lambda}} = \frac{fg}{M_P H} \lesssim \frac{fg}{M^2} = \frac{f}{\phi_0} \tag{E.1.9}$$

(where we have used $H \gtrsim M^2/M_p$, and renamed $\Lambda_i \equiv \Lambda$ as our reference cosmological constant) is very small in relaxionic models, so it is a nice variable to expand $B$. We also introduce

$$q \equiv \frac{1}{\sqrt{\Lambda}} \left( \frac{\Delta\Lambda}{\kappa T} \right) \sim \frac{g\phi}{M_P H} \lesssim \frac{\phi}{\phi_0}. \tag{E.1.10}$$

Unlike $p$, $q$ can be of order 1 or larger, but its introduction simplifies formulae[1]. $p$ is precisely the inverse of the parameter $A$ in [157].

After expansion for small $p$, we get

$$\kappa B = w(q)\frac{2\pi^2 p}{\Lambda} + \mathcal{O}(p^2), \quad w(q) \equiv \frac{1+2q^2}{\sqrt{1+q^2}} - 2q. \tag{E.1.11}$$

The function $w(q)$ goes from 1 at $q = 0$ to $\approx 0.1$ at $q = 1$. Substituting $p$, when studying membranes in relaxion scenarios, it is a good approximation to take

$$B \approx \frac{2\pi^2 T}{H^3} w(q) \tag{E.1.12}$$

which is precisely (8.4.4). Notice that the radius of the bubble is always large. From (E.1.5),

$$r^2_{min} = \frac{1}{\Lambda} \frac{1}{1 + \left(\frac{p}{4} - q\right)^2}, \tag{E.1.13}$$

which is very close to the de Sitter radius $(\Lambda)^{-1/2}$ for small $p$ and $p/q$.

Substitution of typical relaxionic values, while also using $fgM^2 \sim \Lambda_v^4$, $M_P H \gtrsim M^2$, yields

$$B \lesssim 2\pi^2 w(q) \left(\frac{\Lambda_v^4 M_p^4}{m_W^8}\right) \left(\frac{\Lambda_v}{\Lambda_{QCD}}\right)^4 \left(\frac{m_W}{M}\right)^8 \sim 10^{57} w(q) \left(\frac{\Lambda_v}{\Lambda_{QCD}}\right)^4 \left(\frac{m_W}{M}\right)^8. \tag{E.1.14}$$

[1] There is a numeric factor of 1/3 in (E.1.10), which we have ommitted. This does not affect our results, which are all order-of-magnitude estimates.

where we have used $\Lambda_{QCD} \sim 0.2\ GeV$, $M_P \sim 10^{19}\ GeV$, $m_W \sim 10^2\ GeV$. This is another form of equation (8.5.11), which highlights the enormous scales involved.

A final but important comment is in order. The results of [218] which we used as starting point rely on the thin-wall approximation. As discussed in that reference, this means that the thickness of the membrane $L$ should be much lower than the de Sitter radius $\Lambda^{-1/2} = H^{-1}$. Here we are at a loss: although the WGC gives the value of the tension of the membrane, it does not provide a value for $L$.

We have no proof that the bubbles we consider satisfy the thin-wall approximation. However, we can give a plausibility argument. Generically, we expect the effective field theory to be a valid description of the physics up to scales of order the cutoff $M^2$. If the membrane is thicker than $M^{-2}$, it should arise as a soliton of the effective field theory. However, there are no such solitons [2]. So either the EFT we were using is incomplete, or we can trust the thin-wall approximation.

For the stringy models similar to the one in Section 8.6 we can be a little more explicit, since the membrane is generically a $D$-brane wrapping some cycle or chain in the compactification manifold. Then the membrane thickness will generically be of order $\sqrt{\alpha'} g_s$ [302, 303]. In any successful stringy embedding of relaxation, this must be much smaller than the de Sitter radius to claim control of the theory (otherwise the tower of excited string states and winding modes stretching around de Sitter space become light during relaxation). Otherwise we would have e.g. low tension fundamental strings which should be included in the effective field theory during inflation.

In other words, if the thin-wall approximation for the WGC membrane does not hold, we likely cannot trust field theory for scales of order $\sim H$ anyway. Hence we expect to be able to trust the formulae of this Appendix in the relaxionic context, if the relaxion proposal indeed admits an UV completion.

## E.2 Bubble growth, energy balance, and cosmological effects

The bubble solution is obtained in the so-called hyperbolic coordinates [304]. One may change to global coordinates, more natural in the inflationary context, by recalling the definition of de Sitter space as the hyperboloid

$$\Lambda^{-2} = w^2 - \tau^2 + x^2 + y^2 + z^2 = w^2 - \tau^2 + \rho^2 \tag{E.2.1}$$

as done in [218]. These coordinates are related to the usual ones by

$$\Lambda\rho^2 = -\sinh^2(\sqrt{\Lambda}t) + \cosh^2(\sqrt{\Lambda}t)\Lambda r^2 \tag{E.2.2}$$

where $r$ is a dimensionful coordinate such that the metric on the 3-sphere of $t = 0$ is just the flat metric in $\mathbb{R}^4$ restricted to the locus $\Lambda^{-2} = w^2 + r^2$. The worldvolume

[2] Barring the field-theoretic bubbles which arise when the wiggles in the relaxion potential become large; as discussed in the main text, these are not the bubbles we are concerned with.

of the membrane boundary is given by $\rho = r_{min}$, where $r_{min}$ is given by (E.1.5). Therefore, the bubble is indeed a ball in global coordinates, whose radius expands according to

$$r(t) = \sqrt{\frac{r_{min}^2}{\cosh^2(\sqrt{\Lambda}t)} + \frac{1}{\Lambda}\tanh^2(\sqrt{\Lambda}t)}. \tag{E.2.3}$$

For small enough $r_{min}$, we recover the Minkowski limit, but for late times the membranes reach the cosmological horizon. At this point they are frozen by the exponential expansion of de Sitter space. Similarly, membranes initially larger than the de Sitter radius contract instead of expanding. This may give some insight as to why no bubbles exist for $\gamma > 0$; for $\gamma = 0$, $r_{min}$ is precisely $(\Lambda)^{-1/2}$, the de Sitter radius. Bubbles with $\gamma > 0$ should therefore have an initial radius larger than this, and therefore would initially contract. But the critical bubble is precisely the smallest one which can expand; hence no critical bubbles exist for $\gamma > 0$.

The expression (E.1.5) tells us that gravitational effects make the critical bubble radius smaller than it would have been otherwise: The flat-space expression

$$r^2_{\text{min,flat space}} = \left(\frac{T}{\Delta\Lambda}\right)^2 > \frac{1}{\left(\frac{\gamma}{\kappa T}\right)^2 + \Lambda_i} = r^2_{min}. \tag{E.2.4}$$

The flat-space expression can be obtained straightforwardly by demanding that the total energy of the bubble vanishes, since in tunneling (or in any other transition) energy is conserved. In the thin-wall approximation one can take

$$E \approx 4\pi T r^2 - \frac{4}{3}\pi r^3 = 0, \tag{E.2.5}$$

which results in (E.2.4).

A similar argument holds when gravitational effects are taken into account. This is because in asymptotically de Sitter spacetimes the (ADM) mass is well-defined, thanks to the presence of a timelike Killing vector (within the horizon). Consider a spherical region of radius $r$ at a constant time slice in de Sitter. Initially, we have empty de Sitter space with cosmological constant $\Lambda_i$. The ADM mass of this region is given by [140]

$$GM = \int_0^r 4\pi r^2 \rho(r')dr' = \frac{1}{2}\Lambda_i r^3. \tag{E.2.6}$$

This expression already takes into account the gravitational self-energy of the system (which precisely cancels the effect of the warping of the geometry within the bubble). Now, this region is replaced by a thin-wall bubble whose cosmological constant is $\Lambda_i$. We have a contribution $\frac{4}{3}\pi\Lambda_i r^3$ to the energy, analogously to (E.2.6). However, this is not the end of the story. The bubble is bounded by a membrane, which self-gravitates. If the tension of the membrane is $T$, close to it the energy density is $T_{00} = T\delta(l)$, where $l$ is the normalized normal coordinate to the membrane. We have

$\delta(l) = \delta(r)(g_{rr})^{1/2}$. Since $(g_{rr})^{1/2}$ jumps precisely at the location of the membrane, its value can be taken as the arithmetic mean of the values at both sides of the membrane[3] . As a result, the total ADM mass of the bubble is

$$GM = \frac{1}{2}\Lambda_f r^3 + \frac{\kappa T}{4} r^2 \left( \sqrt{1 - \Lambda_f r^2} + \sqrt{1 - \Lambda_i r^2} \right). \tag{E.2.7}$$

Demanding equality with (E.2.6) yields precisely (E.1.5). Since (E.2.7) is a full expression for the energy of the membrane including gravitational backreaction, we may expand it for small $\Lambda_i, \Lambda_f$ to see the effect of the first corrections:

$$GM \approx \frac{4}{3}\pi \Lambda_f r^3 + 4\pi^2 T r^2 - 2\pi T \left( \frac{\Lambda_i + \Lambda_f}{2} \right) r^4 + \mathcal{O}(\Lambda^2). \tag{E.2.8}$$

The $r^4$ term is precisely the Newtonian gravitational interaction energy between a membrane of tension $T$ and radius $r$ with a ball of energy density $\frac{1}{2}(V_i + V_f)$. We now have a clear picture of what is happening: The membrane feels the gravitational field of the bubble interior, which pushes it inwards[4]. The gravitational field sourced by the energy density outside of the bubble is zero via Gauss' law. For large bubbles, this gravitational attraction is strong enough so as to diminish the radius of the critical bubble significantly.

[3] There is some ambiguity here, as the product $\delta(r)(g_{rr})^{1/2}$ involves a product of two distributions which must be conveniently regularized. The arithmetic prescription used here comes from using exactly the same smoothing for the membrane $\delta$ as the one used for the change in vacuum energy.

[4] We are neglecting the gravitational field sourced by the membrane itself, but it would be easy to include. All we would have to do is to replace the $\sqrt{1 - \Lambda_i r^2}$ term above by $\sqrt{1 - \Lambda_f r^2 - 2m/r}$, where $m = 4\pi T r^2$ is the ADM mass of the membrane alone. Nevertheless, in relaxion models with WGC membranes, this contribution is negligible, again of order $f/\phi_0$.

# F

# DBI D5-potential for the axion

Here, we give the details of the computation of the potential (8.6.2) in Section 8.6. The effective action for the microscopic fields of a system of D5-branes in the 10d Einstein frame is given by the Dirac-Born-Infeld (DBI) + Chern-Simons (CS) actions

$$\frac{S}{\mu_5} = -g_s^{-1}\mathrm{STr}\int d^6\xi\,\sqrt{-\det\left(P[E_{MN}+E_{Mi}(Q^{-1}-\delta)^{ij}E_{jN}]+2\pi\alpha' F_{MN}\right)det(Q^i_j)} + \mathrm{STr}\int P[C_6+C_4\wedge\mathcal{F}_2] \quad \text{(F.1)}$$

where

$$E_{MN} = g_s^{1/2}G_{MN} - B_{MN} \quad ; \quad Q^i_j = \delta^i_j + i2\pi\alpha'[\Phi^i,\Phi^k]E_{kj} \quad \text{(F.2)}$$
$$\sigma = 2\pi\alpha' \quad ; \quad \mu_5 = (2\pi)^5\alpha'^{-3}g_s^{-1} \quad ; \quad \mathcal{F}_2 = 2\pi\alpha' F_2 - B_2 \quad \text{(F.3)}$$

$P[\cdot]$ denotes the pullback of the 10d background onto the D5-brane worldvolume and 'STr' is the symmetrised trace over gauge indices. The indices $M, N$ denote the directions extended by the D5-brane while $i, j$ denote the transverse directions. In the absence of NS and RR fluxes the Chern-Simons action plays no role in the discussion.

We are interested in the scalar potential for the position moduli of the D5's, since the Higgs field will later appear as off-diagonal fluctuations of the adjoint field parametrizing the position of a stack of D5-branes. Therefore we will neglect all the terms involving the 4d gauge bosons and the Wilson lines. We are going to assume for the moment no warping, diagonal Minkowski and compact metric and no mixed Minkowski-internal tensors. We also consider vanishing 3-form $G_3$ fluxes but allow for an open string background given by the magnetic worldvolume field strength

$$F_2 = q\omega_2\ , \quad \text{(F.4)}$$

where $\omega_2$ is the orientifold-odd volume form of the 2-cycle $\Sigma_2$ wrapped by the D5-brane. Even if there is no B-field induced by $G_3$ on the brane, we can still have a coupling of the D5 position moduli to the axion coming from dimensionally reducing $B_2$ in the same 2-cycle

$$B_2 = \phi\omega_2 \ , \tag{F.5}$$

as we will see in the following.

Neglecting derivative couplings, the determinant in the DBI action can be factorized between Minkowski and the internal space as follows

$$\det(P[E_{MN}] + \sigma F_{MN}) = g_s^2 \det\left(\eta_{\mu\nu} + 2\sigma^2 \partial_\mu \Phi \partial_\nu \bar{\Phi}\right) \cdot \det\left(g_{ab} + g_s^{-1/2}\mathcal{F}_{ab}\right) \tag{F.6}$$

where $\mu, \nu$ label the 4d non-compact directions and $a, b$ the internal D7-brane dimensions, and

$$\mathcal{F}_{ab} = \sigma F_{ab} - B_{ab} \ . \tag{F.7}$$

Then, using the matrix identity

$$\begin{aligned} \det(1+\varepsilon M) &= 1 + \varepsilon \,\mathrm{tr}\ M - \varepsilon^2 \left[\frac{1}{2}\mathrm{tr}\ M^2 - \frac{1}{2}(\mathrm{tr}\ M)^2\right] \\ &+ \varepsilon^3 \left[\frac{1}{3}\mathrm{tr}\ M^3 - \frac{1}{2}(\mathrm{tr}\ M)(\mathrm{tr}\ M^2) + \frac{1}{6}(\mathrm{tr}\ M)^3\right] \\ &- \varepsilon^4 \left[\frac{1}{4}\mathrm{tr}\ M^4 - \frac{1}{8}(\mathrm{tr}\ M^2)^2 - \frac{1}{3}(\mathrm{tr}\ M)(\mathrm{tr}\ M^3) \right. \\ &+ \left. \frac{1}{4}(\mathrm{tr}\ M)^2(\mathrm{tr}\ M^2) + \frac{1}{24}(\mathrm{tr}\ M)^4\right] \end{aligned} \tag{F.8}$$

we obtain on one hand that

$$\begin{aligned} &-\det\left(\eta_{\mu\nu} + 2\sigma^2 \partial_\mu \Phi \partial_\nu \bar{\Phi}\right) \det\left(g_{ab} + g_s^{-1/2}\mathcal{F}_{ab}\right) \\ &= \left(1 + 2\sigma^2 \partial_\mu \Phi \partial^\mu \bar{\Phi}\right)\left(1 + \frac{1}{2} g_s^{-1} \mathcal{F}_{ab}\mathcal{F}^{ab}\right), \end{aligned} \tag{F.9}$$

where we have neglected terms with more than two derivatives in Minkowski. On the other hand we have that

$$det(Q^m_n) = g_s^2 \det\left(\delta^m_n + i\sigma[\phi^m, \phi^p]\delta_{pn}\right) = g_s^2 \left(1 + \frac{1}{2}\sigma^2[\phi^m, \phi^n][\phi_n, \phi_m] + \dots\right) \tag{F.10}$$

In a supersymmetric configuration with vanishing D-terms $[\Phi, \Phi^*] = 0$ we get

$$Tr([\phi^m, \phi^n]^2) = [\phi^m, \phi^n][\phi_n, \phi_m] = -4|[\Phi^1, \Phi^2]|^2 \tag{F.11}$$

and the quartic terms combine with the quadratic terms to complete a perfect square (see e.g. [240, 305]) . To do the computation more explicit we took the transverse space to the branes to be $T^4$, and $\Phi_1, \Phi_2$ the two adjoints parametrizing the position in the torus.

Putting everything together we find that the relevant part of the DBI action is given by

$$S_{DBI} = -\mu_5 g_s^{-1} \mathrm{STr} \int d^6\xi \, I, \tag{F.12}$$

$$I = \sqrt{\left(1 + 2\sigma^2 \partial_\mu \Phi^i \partial^\mu \bar{\Phi}^i\right)\left(1 + \frac{1}{2} g_s^{-1} \mathcal{F}_{ab}\mathcal{F}^{ab}\right)(1 - 4\sigma^2 |[\Phi^1, \Phi^2]|^2)^2} \ , \tag{F.13}$$

which are precisely equations (8.6.1) and (8.6.2).

# Bibliography


[1] R. Percacci, *Asymptotic Safety*, arXiv:0709.3851.

[2] A. Font, F. Quevedo, and S. Theisen, *A comment on continuous spin representations of the Poincare group and perturbative string theory*, *Fortsch. Phys.* **62** (2014) 975–980, [arXiv:1302.4771].

[3] A. H. Chamseddine, *A Study of noncritical strings in arbitrary dimensions*, *Nucl.Phys.* **B368** (1992) 98–120.

[4] S. Hellerman, *On the landscape of superstring theory in $D > 10$*, hep-th/0405041.

[5] S. Hellerman and I. Swanson, *Dimension-changing exact solutions of string theory*, *JHEP* **0709** (2007) 096, [hep-th/0612051].

[6] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string.* Cambridge University Press, 1998.

[7] N. Seiberg, *Notes on quantum Liouville theory and quantum gravity*, *Prog.Theor.Phys.Suppl.* **102** (1990) 319–349.

[8] A. M. Polyakov, *Quantum Geometry of Bosonic Strings*, *Phys.Lett.* **B103** (1981) 207–210.

[9] M. Matone, *Uniformization theory and 2-D gravity. 1. Liouville action and intersection numbers*, *Int.J.Mod.Phys.* **A10** (1995) 289–336, [hep-th/9306150].

[10] L. E. Ibañez and A. M. Uranga, *String theory and particle physics: An introduction to string phenomenology.* Cambridge University Press, 2012.

[11] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory, volume 1: Introduction.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, anniversary ed., 1987.

[12] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond.* Cambridge University Press, 1998.

[13] M. Berasaluce-González, M. Montero, A. Retolaza, and A. M. Uranga, *Discrete gauge symmetries from (closed string) tachyon condensation*, *JHEP* **1311** (2013) 144, [arXiv:1305.6788].

[14] J. H. Schwarz and E. Witten, *Anomaly analysis of brane - anti-brane systems*, *JHEP* **0103** (2001) 032, [hep-th/0103099].

[15] E. G. Gimon and J. Polchinski, *Consistency conditions for orientifolds and d manifolds*, *Phys.Rev.* **D54** (1996) 1667–1676, [hep-th/9601038].

[16] E. R. Sharpe, *Discrete torsion*, *Phys.Rev.* **D68** (2003) 126003, [hep-th/0008154].

[17] M. Billo, B. Craps, and F. Roose, *Anomalous couplings of type 0 D-branes*, hep-th/9908030.

[18] A. A. Tseytlin, *On dilaton dependence of type II superstring action*, *Class.Quant.Grav.* **13** (1996) L81–L85, [hep-th/9601109].

[19] D. Polyakov, *RR: Dilaton interaction in a type IIB superstring*, *Nucl.Phys.* **B468** (1996) 155–162, [hep-th/9512028].

[20] S. Hellerman and I. Swanson, *Charting the landscape of supercritical string theory*, *Phys.Rev.Lett.* **99** (2007) 171601, [arXiv:0705.0980].

[21] T. Banks and L. Susskind, *Brane - anti-brane forces*, hep-th/9511194.

[22] O. Bergman, K. Hori, and P. Yi, *Confinement on the brane*, *Nucl.Phys.* **B580** (2000) 289–310, [hep-th/0002223].

[23] A. Sen, *Tachyon condensation on the brane anti-brane system*, *JHEP* **9808** (1998) 012, [hep-th/9805170].

[24] M. Headrick, S. Minwalla, and T. Takayanagi, *Closed string tachyon condensation: An Overview*, *Class.Quant.Grav.* **21** (2004) S1539–S1565, [hep-th/0405064].

[25] K. Ohmori, *A Review on tachyon condensation in open string field theories*, hep-th/0102085.

[26] M. Schnabl, *Analytic solution for tachyon condensation in open string field theory*, *Adv.Theor.Math.Phys.* **10** (2006) 433–501, [hep-th/0511286].

[27] S. Hellerman and M. Schnabl, *Light-like tachyon condensation in Open String Field Theory*, *JHEP* **1304** (2013) 005, [arXiv:0803.1184].

[28] J. A. Harvey and P. Kraus, *D-Branes as unstable lumps in bosonic open string field theory*, *JHEP* **0004** (2000) 012, [hep-th/0002117].

[29] N. Berkovits, *SuperPoincare invariant superstring field theory*, *Nucl.Phys.* **B450** (1995) 90–102, [hep-th/9503099].

[30] N. Berkovits, A. Sen, and B. Zwiebach, *Tachyon condensation in superstring field theory*, *Nucl.Phys.* **B587** (2000) 147–178, [hep-th/0002211].

[31] H. Sonoda and B. Zwiebach, *Covariant closed string theory cannot be cubic*, *Nucl.Phys.* **B336** (1990) 185.

[32] A. Adams, J. Polchinski, and E. Silverstein, *Don't panic! Closed string tachyons in ALE space-times*, *JHEP* **0110** (2001) 029, [hep-th/0108075].

[33] F. Quevedo, S. Krippendorf, and O. Schlotterer, *Cambridge Lectures on Supersymmetry and Extra Dimensions*, arXiv:1011.1491.

[34] B. Zumino, *Supersymmetric sigma models in two-dimensions*, 1997.

[35] S. Hellerman and I. Swanson, *Cosmological unification of string theories*, *JHEP* **0807** (2008) 022, [hep-th/0612116].

[36] L. Alvarez-Gaume and E. Witten, *Gravitational Anomalies*, *Nucl.Phys.* **B234** (1984) 269.

[37] E. Witten, *D-branes and K theory*, *JHEP* **9812** (1998) 019, [hep-th/9810188].

[38] T. Banks and L. J. Dixon, *Constraints on String Vacua with Space-Time Supersymmetry*, *Nucl.Phys.* **B307** (1988) 93–108.

[39] L. Abbott and M. B. Wise, *Wormholes and Global Symmetries*, *Nucl.Phys.* **B325** (1989) 687.

[40] S. R. Coleman and K.-M. Lee, *Wormholes made without massless matter fields*, *Nucl.Phys.* **B329** (1990) 387.

[41] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, *Gravity and global symmetries*, *Phys.Rev.* **D52** (1995) 912–935, [hep-th/9502069].

[42] T. Banks and N. Seiberg, *Symmetries and Strings in Field Theory and Gravity*, *Phys.Rev.* **D83** (2011) 084019, [arXiv:1011.5120].

[43] L. Susskind, *Trouble for remnants*, hep-th/9501106.

[44] S. Hawking, *Black hole explosions*, *Nature* **248** (1974) 30–31.

[45] G. Dvali and C. Gómez, *Black Hole's 1/N Hair*, *Phys.Lett.* **B719** (2013) 419–423, [arXiv:1203.6575].

[46] G. Dvali and C. Gómez, *Black Holes as Critical Point of Quantum Phase Transition*, *Eur.Phys.J.* **C74** (2014) 2752, [arXiv:1207.4059].

[47] G. James and M. Liebeck, *Representations and Characters of Finite Groups*. Cambridge University Press, 1993.

[48] W. G. Unruh, *Notes on black hole evaporation*, *Phys. Rev.* **D14** (1976) 870.

[49] C. Burgess, J. Conlon, L.-Y. Hung, C. Kom, A. Maharana, *et. al.*, *Continuous Global Symmetries and Hyperweak Interactions in String Compactifications*, *JHEP* **0807** (2008) 073, [arXiv:0805.4037].

[50] D. Regalado, *Flavour hierarchies and unification in String Theory.* PhD thesis, Instituto de Física Teórica, 2014.

[51] F. Denef, *Les Houches Lectures on Constructing String Vacua*, arXiv:0803.1194.

[52] T. Ortín, *Gravity and strings.* Cambridge University Press, 1st ed., 2007.

[53] E. Witten, *An SU(2) Anomaly*, *Phys.Lett.* **B117** (1982) 324–328.

[54] E. Witten, *Global gravitational anomalies*, *Commun.Math.Phys.* **100** (1985) 197.

[55] L. Alvarez-Gaume, *An introduction to anomalies*, 1985.

[56] A. Hebecker, S. C. Kraus, D. Lust, S. Steinfurt, and T. Weigand, *Fluxbrane Inflation*, *Nucl.Phys.* **B854** (2012) 509–551, [arXiv:1104.5016].

[57] M. S. Costa and M. Gutperle, *The Kaluza-Klein Melvin solution in M theory*, *JHEP* **0103** (2001) 027, [hep-th/0012072].

[58] P. Saffin, *Gravitating fluxbranes*, *Phys.Rev.* **D64** (2001) 024014, [gr-qc/0104014].

[59] M. Gutperle and A. Strominger, *Fluxbranes in string theory*, *JHEP* **0106** (2001) 035, [hep-th/0104136].

[60] A. M. Uranga, *Wrapped fluxbranes*, hep-th/0108196.

[61] M. Dine, R. G. Leigh, and D. A. MacIntire, *Of CP and other gauge symmetries in string theory*, *Phys.Rev.Lett.* **69** (1992) 2030–2032, [hep-th/9205011].

[62] S. Chaudhuri, G. Hockney, and J. D. Lykken, *Maximally supersymmetric string theories in $D < 10$*, *Phys.Rev.Lett.* **75** (1995) 2264–2267, [hep-th/9505054].

[63] S. Chaudhuri and J. Polchinski, *Moduli space of CHL strings*, *Phys.Rev.* **D52** (1995) 7168–7173, [hep-th/9506048].

[64] L. E. Ibañez, H. P. Nilles, and F. Quevedo, *Reducing the Rank of the Gauge Group in Orbifold Compactifications of the Heterotic String*, *Phys.Lett.* **B192** (1987) 332.

[65] G. Aldazabal, A. Font, L. E. Ibañez, and A. Uranga, *Building GUTs from strings*, *Nucl.Phys.* **B465** (1996) 34–70, [hep-th/9508033].

[66] G. Aldazabal, A. Font, L. E. Ibañez, and A. Uranga, *String GUTs*, *Nucl.Phys.* **B452** (1995) 3–44, [hep-th/9410206].

[67] A. Dabholkar and C. Hull, *Duality twists, orbifolds, and fluxes*, *JHEP* **0309** (2003) 054, [hep-th/0210209].

[68] T. Weigand, *Lectures on F-theory compactifications and model building*, *Class.Quant.Grav.* **27** (2010) 214004, [arXiv:1009.3497].

[69] A. Sen, *F theory and orientifolds*, *Nucl.Phys.* **B475** (1996) 562–578, [hep-th/9605150].

[70] K. Dasgupta and S. Mukhi, *F theory at constant coupling*, *Phys.Lett.* **B385** (1996) 125–131, [hep-th/9606044].

[71] S. Cecotti, *A geometric introduction to* F*-theory*. SISSA, 2010.

[72] Hübsch, Tristan, *Calabi-Yau mainfolds: A Bestiary for physicists*. Wspc, 1992.

[73] A. Strominger and E. Witten, *New Manifolds for Superstring Compactification*, *Commun.Math.Phys.* **101** (1985) 341.

[74] K.-w. Choi, D. B. Kaplan, and A. E. Nelson, *Is CP a gauge symmetry?*, *Nucl.Phys.* **B391** (1993) 515–530, [hep-ph/9205202].

[75] M. Berasaluce-Gonzalez, L. E. Ibañez, P. Soler, and A. M. Uranga, *Discrete gauge symmetries in D-brane models*, *JHEP* **1112** (2011) 113, [arXiv:1106.4169].

[76] L. Ibañez, A. Schellekens, and A. Uranga, *Discrete Gauge Symmetries in Discrete MSSM-like Orientifolds*, *Nucl.Phys.* **B865** (2012) 509–540, [arXiv:1205.5364].

[77] M. Berasaluce-Gonzalez, P. Camara, F. Marchesano, D. Regalado, and A. Uranga, *Nonabelian discrete gauge symmetries in 4d string models*, *JHEP* **1209** (2012) 059, [arXiv:1206.2383].

[78] M. Berasaluce-Gonzalez, P. Camara, F. Marchesano, and A. Uranga, *Zp charged branes in flux compactifications*, *JHEP* **1304** (2013) 138, [arXiv:1211.5317].

[79] P. Anastasopoulos, M. Cvetič, R. Richter, and P. K. Vaudrevange, *String Constraints on Discrete Symmetries in MSSM Type II Quivers*, *JHEP* **1303** (2013) 011, [arXiv:1211.1017].

[80] G. Honecker and W. Staessens, *To Tilt or Not To Tilt: Discrete Gauge Symmetries in Global Intersecting D-Brane Models*, *JHEP* **1310** (2013) 146, [arXiv:1303.4415].

[81] F. Marchesano, D. Regalado, and L. Vazquez-Mercado, *Discrete flavor symmetries in D-brane models*, *JHEP* **1309** (2013) 028, [arXiv:1306.1284].

[82] E. Witten, *Baryons and branes in anti-de Sitter space*, *JHEP* **9807** (1998) 006, [hep-th/9805112].

[83] D. Bailin and A. Love, *Orbifold compactifications of string theory*, *Phys.Rept.* **315** (1999) 285–408.

[84] J. Casas and C. Muñoz, *Yukawa Couplings in SU(3) X SU(2) X U(1)-y Orbifold Models*, *Phys.Lett.* **B212** (1988) 343.

[85] T. Kobayashi, H. P. Nilles, F. Ploger, S. Raby, and M. Ratz, *Stringy origin of nonabelian discrete flavor symmetries*, *Nucl.Phys.* **B768** (2007) 135–156, [hep-ph/0611020].

[86] S. Gukov, M. Rangamani, and E. Witten, *Dibaryons, strings and branes in AdS orbifold models*, *JHEP* **9812** (1998) 025, [hep-th/9811048].

[87] B. A. Burrington, J. T. Liu, and L. A. Pando Zayas, *Finite Heisenberg groups in quiver gauge theories*, *Nucl.Phys.* **B747** (2006) 436–454, [hep-th/0602094].

[88] K. Hori and C. Vafa, *Mirror symmetry*, hep-th/0002222.

[89] A. Hanany and A. Iqbal, *Quiver theories from D6 branes via mirror symmetry*, *JHEP* **0204** (2002) 009, [hep-th/0108137].

[90] A. Giveon, M. Porrati, and E. Rabinovici, *Target space duality in string theory*, *Phys.Rept.* **244** (1994) 77–202, [hep-th/9401139].

[91] J. Shelton, W. Taylor, and B. Wecht, *Nongeometric flux compactifications*, *JHEP* **0510** (2005) 085, [hep-th/0508133].

[92] E. Witten, *Phase transitions in M theory and F theory*, *Nucl.Phys.* **B471** (1996) 195–216, [hep-th/9603150].

[93] I. García-Etxebarria, M. Montero, and A. Uranga, *Heterotic NS5-branes from closed string tachyon condensation*, arXiv:1405.0009.

[94] J. Callan, Curtis G., J. A. Harvey, and A. Strominger, *World sheet approach to heterotic instantons and solitons*, *Nucl.Phys.* **B359** (1991) 611–634.

[95] E. Gava, K. Narain, and M. Sarmadi, *Little string theories in heterotic backgrounds*, *Nucl.Phys.* **B626** (2002) 3–25, [hep-th/0112200].

[96] J. Callan, Curtis G., J. A. Harvey, and A. Strominger, *Supersymmetric string solitons*, hep-th/9112030.

[97] E. Witten, *Small instantons in string theory*, *Nucl.Phys.* **B460** (1996) 541–559, [hep-th/9511030].

[98] M. Atiyah, R. Bott, and A. Shapiro, *Clifford modules*, *Topology* **3** (1964) S3–S38.

[99] L. Alvarez-Gaume and P. H. Ginsparg, *The Structure of Gauge and Gravitational Anomalies*, *Annals Phys.* **161** (1985) 423.

[100] S. Sugimoto, *Anomaly cancellations in type I D-9 - anti-D-9 system and the USp(32) string theory*, *Prog.Theor.Phys.* **102** (1999) 685–699, [hep-th/9905159].

[101] P. S. Aspinwall, *K3 surfaces and string duality*, hep-th/9611137.

[102] M. B. Green, J. H. Schwarz, and P. C. West, *Anomaly Free Chiral Theories in Six-Dimensions*, *Nucl.Phys.* **B254** (1985) 327–348.

[103] E. Witten, *Sigma models and the ADHM construction of instantons*, *J.Geom.Phys.* **15** (1995) 215–226, [hep-th/9410052].

[104] X.-G. Wen and E. Witten, *Electric and Magnetic Charges in Superstring Models*, *Nucl.Phys.* **B261** (1985) 651.

[105] G. W. Moore and E. Witten, *Self-duality, Ramond-Ramond fields, and K theory*, *JHEP* **0005** (2000) 032, [hep-th/9912279].

[106] A. M. Uranga, *D-brane probes, RR tadpole cancellation and K theory charge*, *Nucl.Phys.* **B598** (2001) 225–246, [hep-th/0011048].

[107] S. Hellerman and X. Liu, *Dynamical dimension change in supercritical string theory*, hep-th/0409071.

[108] O. Bergman and M. R. Gaberdiel, *A Nonsupersymmetric open string theory and S duality*, *Nucl.Phys.* **B499** (1997) 183–204, [hep-th/9701137].

[109] O. Bergman and M. R. Gaberdiel, *Dualities of type 0 strings*, *JHEP* **9907** (1999) 022, [hep-th/9906055].

[110] M. B. Green, J. A. Harvey, and G. W. Moore, *I-brane inflow and anomalous couplings on d-branes*, *Class.Quant.Grav.* **14** (1997) 47–52, [hep-th/9605033].

[111] Y.-K. E. Cheung and Z. Yin, *Anomalies, branes, and currents*, *Nucl.Phys.* **B517** (1998) 69–91, [hep-th/9710206].

[112] R. Minasian and G. W. Moore, *K theory and Ramond-Ramond charge*, *JHEP* **9711** (1997) 002, [hep-th/9710230].

[113] E. Witten, *Some comments on string dynamics*, hep-th/9507121.

[114] C. Vafa and E. Witten, *A One loop test of string duality*, *Nucl.Phys.* **B447** (1995) 261–270, [hep-th/9505053].

[115] E. Witten, *Five-branes and M theory on an orbifold*, *Nucl.Phys.* **B463** (1996) 383–397, [hep-th/9512219].

[116] S. Forte, *Explicit Construction of Anomalies*, *Nucl.Phys.* **B288** (1987) 252.

[117] M. Nakahara, *Geometry, Topology and Physics, Second Edition*. Graduate student series in physics. Taylor & Francis, 2003.

[118] J. T. Liu and R. Minasian, *Higher-derivative couplings in string theory: dualities and the B-field*, *Nucl.Phys.* **B874** (2013) 413–470, [arXiv:1304.3137].

[119] S. Gukov, *K theory, reality, and orientifolds*, *Commun.Math.Phys.* **210** (2000) 621–639, [hep-th/9901042].

[120] E. Witten, *Phases of N=2 theories in two-dimensions*, *Nucl.Phys.* **B403** (1993) 159–222, [hep-th/9301042].

[121] C. Doran, B. Greene, and S. Judes, *Families of quintic Calabi-Yau 3-folds with discrete symmetries*, *Commun.Math.Phys.* **280** (2008) 675–725, [hep-th/0701206].

[122] T. Hübsch and A. Rahman, *On the geometry and homology of certain simple stratified varieties*, *Journal of Geometry and Physics* **53** (Jan., 2005) 31–48, [math/0210394].

[123] A. Strominger, *Massless black holes and conifolds in string theory*, *Nucl.Phys.* **B451** (1995) 96–108, [hep-th/9504090].

[124] A. Klemm, P. Mayr, and C. Vafa, *BPS states of exceptional noncritical strings*, *Nucl.Phys.Proc.Suppl.* **58** (1997) 177, [hep-th/9607139].

[125] G. Aldazabal, A. Font, L. E. Ibanez, and A. Uranga, *New branches of string compactifications and their F theory duals*, *Nucl.Phys.* **B492** (1997) 119–151, [hep-th/9607121].

[126] V. Braun and D. R. Morrison, *F-theory on Genus-One Fibrations*, *JHEP* **1408** (2014) 132, [arXiv:1401.7844].

[127] D. R. Morrison and W. Taylor, *Sections, multisections, and U(1) fields in F-theory*, arXiv:1404.1527.

[128] L. B. Anderson, I. GarcÃa-Etxebarria, T. W. Grimm, and J. Keitel, *Physics of F-theory compactifications without section*, *JHEP* **1412** (2014) 156, [arXiv:1406.5180].

[129] D. Klevers, D. K. Mayorga Pena, P.-K. Oehlmann, H. Piragua, and J. Reuter, *F-Theory on all Toric Hypersurface Fibrations and its Higgs Branches*, *JHEP* **1501** (2015) 142, [arXiv:1408.4808].

[130] I. GarcÃa-Etxebarria, T. W. Grimm, and J. Keitel, *Yukawas and discrete symmetries in F-theory compactifications without section*, *JHEP* **1411** (2014) 125, [arXiv:1408.6448].

[131] C. Mayrhofer, E. Palti, O. Till, and T. Weigand, *Discrete Gauge Symmetries by Higgsing in four-dimensional F-Theory Compactifications*, *JHEP* **1412** (2014) 068, [arXiv:1408.6831].

[132] C. Mayrhofer, E. Palti, O. Till, and T. Weigand, *On Discrete Symmetries and Torsion Homology in F-Theory*, arXiv:1410.7814.

[133] M. Cvetič, R. Donagi, D. Klevers, H. Piragua, and M. Poretschkin, *F-Theory Vacua with $Z_3$ Gauge Symmetry*, arXiv:1502.0695.

[134] T. W. Grimm, T. G. Pugh, and D. Regalado, *Non-Abelian discrete gauge symmetries in F-theory*, arXiv:1504.0627.

[135] A. Collinucci and R. Savelli, *F-theory on singular spaces*, arXiv:1410.4867.

[136] A. Collinucci and R. Savelli, *T-branes as branes within branes*, arXiv:1410.4178.

[137] R. Donagi and M. Wijnholt, *Gluing Branes, I*, *JHEP* **1305** (2013) 068, [arXiv:1104.2610].

[138] M. Herbst, K. Hori, and D. Page, *Phases Of N=2 Theories In 1+1 Dimensions With Boundary*, arXiv:0803.2045.

[139] 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].

[140] R. M. Wald, *General Relativity.* University Of Chicago Press, 1984.

[141] R. Penrose, *The question of cosmic censorship*, *J. Astrophys. Astron.* **20** (1999) 233–248.

[142] S. W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199–220. [,167(1975)].

[143] S. R. Coleman, J. Preskill, and F. Wilczek, *Quantum hair on black holes*, *Nucl. Phys.* **B378** (1992) 175–246, [hep-th/9201059].

[144] J. H. Traschen, *An Introduction to black hole evaporation*, in *Mathematical methods in physics. Proceedings, Winter School, Londrina, Brazil, August 17-26, 1999*, 1999. gr-qc/0010055.

[145] A. M. Polyakov, *Particle Spectrum in the Quantum Field Theory*, *JETP Lett.* **20** (1974) 194–195. [Pisma Zh. Eksp. Teor. Fiz.20,430(1974)].

[146] G. 't Hooft, *Magnetic Monopoles in Unified Gauge Theories*, *Nucl. Phys.* **B79** (1974) 276–284.

[147] B. Heidenreich, M. Reece, and T. Rudelius, *Sharpening the Weak Gravity Conjecture with Dimensional Reduction*, *JHEP* **02** (2016) 140, [arXiv:1509.0637].

[148] G. T. Horowitz and A. Strominger, *Black strings and P-branes*, *Nucl. Phys.* **B360** (1991) 197–209.

[149] G. W. Gibbons, *Antigravitating Black Hole Solitons with Scalar Hair in N=4 Supergravity*, *Nucl. Phys.* **B207** (1982) 337–349.

[150] R. C. Myers and M. J. Perry, *Black Holes in Higher Dimensional Space-Times*, *Annals Phys.* **172** (1986) 304.

[151] G. W. Gibbons and K.-i. Maeda, *Black Holes and Membranes in Higher Dimensional Theories with Dilaton Fields*, *Nucl. Phys.* **B298** (1988) 741–775.

[152] D. Garfinkle, G. T. Horowitz, and A. Strominger, *Charged black holes in string theory*, *Phys. Rev.* **D43** (1991) 3140. [Erratum: Phys. Rev.D45,3888(1992)].

[153] J. X. Lu, *ADM masses for black strings and p-branes*, *Phys. Lett.* **B313** (1993) 29–34, [hep-th/9304159].

[154] M. J. Duff and J. X. Lu, *Black and super p-branes in diverse dimensions*, *Nucl. Phys.* **B416** (1994) 301–334, [hep-th/9306052].

[155] M. J. Duff, H. Lu, and C. N. Pope, *The Black branes of M theory*, *Phys. Lett.* **B382** (1996) 73–80, [hep-th/9604052].

[156] C. Cheung and G. N. Remmen, *Naturalness and the Weak Gravity Conjecture*, *Phys. Rev. Lett.* **113** (2014) 051601, [arXiv:1402.2287].

[157] J. Brown, W. Cottrell, G. Shiu, and P. Soler, *Fencing in the Swampland: Quantum Gravity Constraints on Large Field Inflation*, *JHEP* **10** (2015) 023, [arXiv:1503.0478].

[158] B. Heidenreich, M. Reece, and T. Rudelius, *Weak Gravity Strongly Constrains Large-Field Axion Inflation*, *JHEP* **12** (2015) 108, [arXiv:1506.0344].

[159] R. G. Hulet, E. S. Hilfer, and D. Kleppner, *Inhibited Spontaneous Emission by a Rydberg Atom*, *Phys. Rev. Lett.* **55** (1985) 2137–2140.

[160] Y. Nakayama and Y. Nomura, *Weak gravity conjecture in the AdS/CFT correspondence*, *Phys. Rev.* **D92** (2015), no. 12 126006, [arXiv:1509.0164].

[161] D. Harlow, *Wormholes, Emergent Gauge Fields, and the Weak Gravity Conjecture*, *JHEP* **01** (2016) 122, [arXiv:1510.0791].

[162] J. E. Kim, H. P. Nilles, and M. Peloso, *Completing natural inflation*, *JCAP* **0501** (2005) 005, [hep-ph/0409138].

[163] M. Montero, A. M. Uranga, and I. Valenzuela, *Transplanckian axions!?*, *JHEP* **08** (2015) 032, [arXiv:1503.0388].

[164] A. de la Fuente, P. Saraswat, and R. Sundrum, *Natural Inflation and Quantum Gravity*, *Phys. Rev. Lett.* **114** (2015), no. 15 151303, [arXiv:1412.3457].

[165] T. Rudelius, *On the Possibility of Large Axion Moduli Spaces*, arXiv:1409.5793.

[166] T. Rudelius, *Constraints on Axion Inflation from the Weak Gravity Conjecture*, arXiv:1503.0079.

[167] T. C. Bachlechner, C. Long, and L. McAllister, *Planckian Axions and the Weak Gravity Conjecture*, arXiv:1503.0785.

[168] A. Hebecker, P. Mangat, F. Rompineve, and L. T. Witkowski, *Winding out of the Swamp: Evading the Weak Gravity Conjecture with F-term Winding Inflation?*, *Phys. Lett.* **B748** (2015) 455–462, [arXiv:1503.0791].

[169] J. Brown, W. Cottrell, G. Shiu, and P. Soler, *On Axionic Field Ranges, Loopholes and the Weak Gravity Conjecture*, *JHEP* **04** (2016) 017, [arXiv:1504.0065].

[170] D. Junghans, *Large-Field Inflation with Multiple Axions and the Weak Gravity Conjecture*, *JHEP* **02** (2016) 128, [arXiv:1504.0356].

[171] E. Palti, *On Natural Inflation and Moduli Stabilisation in String Theory*, *JHEP* **10** (2015) 188, [arXiv:1508.0000].

[172] K. Kooner, S. Parameswaran, and I. Zavala, *Warping the Weak Gravity Conjecture*, arXiv:1509.0704.

[173] R. Kappl, H. P. Nilles, and M. W. Winkler, *Modulated Natural Inflation*, arXiv:1511.0556.

[174] K. Choi and H. Kim, *Aligned natural inflation with modulations*, arXiv:1511.0720.

[175] D. Baumann and L. McAllister, *Inflation and String Theory*, arXiv:1404.2601.

[176] **BICEP2 Collaboration** Collaboration, P. Ade *et. al.*, *Detection of B-Mode Polarization at Degree Angular Scales by BICEP2*, *Phys.Rev.Lett.* **112** (2014), no. 24 241101, [arXiv:1403.3985].

[177] **BICEP2 Collaboration, Planck Collaboration** Collaboration, P. Ade *et. al.*, *A Joint Analysis of BICEP2/Keck Array and Planck Data*, *Phys.Rev.Lett.* (2015) [arXiv:1502.0061].

[178] D. H. Lyth, *What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?*, *Phys.Rev.Lett.* **78** (1997) 1861–1863, [hep-ph/9606387].

[179] M. Reece, *Notes on inflation and the lyth bound*, March, 2014.

[180] K. Freese, J. A. Frieman, and A. V. Olinto, *Natural inflation with pseudo - Nambu-Goldstone bosons*, *Phys.Rev.Lett.* **65** (1990) 3233–3236.

[181] D. Croon and V. Sanz, *Saving Natural Inflation*, *JCAP* **1502** (2015), no. 02 008, [arXiv:1411.7809].

[182] 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].

[183] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, *N-flation*, *JCAP* **0808** (2008) 003, [hep-th/0507205].

[184] K. Choi, H. Kim, and S. Yun, *Natural inflation with multiple sub-Planckian axions*, *Phys.Rev.* **D90** (2014), no. 2 023545, [arXiv:1404.6209].

[185] T. Higaki and F. Takahashi, *Natural and Multi-Natural Inflation in Axion Landscape*, *JHEP* **1407** (2014) 074, [arXiv:1404.6923].

[186] S. H. H. Tye and S. S. C. Wong, *Helical Inflation and Cosmic Strings*, arXiv:1404.6988.

[187] R. Kappl, S. Krippendorf, and H. P. Nilles, *Aligned Natural Inflation: Monodromies of two Axions*, *Phys.Lett.* **B737** (2014) 124–128, [arXiv:1404.7127].

[188] T. C. Bachlechner, M. Dias, J. Frazer, and L. McAllister, *Chaotic inflation with kinetic alignment of axion fields*, *Phys.Rev.* **D91** (2015), no. 2 023520, [arXiv:1404.7496].

[189] I. Ben-Dayan, F. G. Pedro, and A. Westphal, *Hierarchical Axion Inflation*, *Phys.Rev.Lett.* **113** (2014), no. 26 261301, [arXiv:1404.7773].

[190] T. Higaki and F. Takahashi, *Axion Landscape and Natural Inflation*, arXiv:1409.8409.

[191] T. C. Bachlechner, C. Long, and L. McAllister, *Planckian Axions in String Theory*, arXiv:1412.1093.

[192] C. Burgess and D. Roest, *Inflation by Alignment*, arXiv:1412.1614.

[193] X. Gao, T. Li, and P. Shukla, *Combining Universal and Odd RR Axions for Aligned Natural Inflation*, *JCAP* **1410** (2014), no. 10 048, [arXiv:1406.0341].

[194] R. Blumenhagen, M. Cvetic, and T. Weigand, *Spacetime instanton corrections in 4D string vacua: The Seesaw mechanism for D-Brane models*, *Nucl.Phys.* **B771** (2007) 113–142, [hep-th/0609191].

[195] L. Ibanez and A. Uranga, *Neutrino Majorana Masses from String Theory Instanton Effects*, *JHEP* **0703** (2007) 052, [hep-th/0609213].

[196] B. Florea, S. Kachru, J. McGreevy, and N. Saulina, *Stringy Instantons and Quiver Gauge Theories*, *JHEP* **0705** (2007) 024, [hep-th/0610003].

[197] R. Blumenhagen, M. Cvetic, S. Kachru, and T. Weigand, *D-Brane Instantons in Type II Orientifolds*, *Ann.Rev.Nucl.Part.Sci.* **59** (2009) 269–296, [arXiv:0902.3251].

[198] L. E. Ibáñez and A. M. Uranga, *String theory and particle physics: An introduction to string phenomenology.* Cambridge University Press, 2012.

[199] T. W. Grimm, *Axion inflation in type II string theory*, *Phys.Rev.* **D77** (2008) 126007, [arXiv:0710.3883].

[200] I. Ben-Dayan, F. G. Pedro, and A. Westphal, *Towards Natural Inflation in String Theory*, arXiv:1407.2562.

[201] C. Long, L. McAllister, and P. McGuirk, *Aligned Natural Inflation in String Theory*, *Phys.Rev.* **D90** (2014), no. 2 023501, [arXiv:1404.7852].

[202] G. Shiu, W. Staessens, and F. Ye, *Widening the Axion Window via Kinetic and Stückelberg Mixings*, arXiv:1503.0101.

[203] R. Kappl, H. P. Nilles, and M. W. Winkler, *Natural Inflation and Low Energy Supersymmetry*, arXiv:1503.0177.

[204] G. Shiu, W. Staessens, and F. Ye, *Large Field Inflation from Axion Mixing*, arXiv:1503.0296.

[205] I. Garcia-Etxebarria, F. Marchesano, and A. M. Uranga, *Non-perturbative F-terms across lines of BPS stability*, *JHEP* **0807** (2008) 028, [arXiv:0805.0713].

[206] A. M. Uranga, *D-brane instantons and the effective field theory of flux compactifications*, *JHEP* **0901** (2009) 048, [arXiv:0808.2918].

[207] Z. Kenton and S. Thomas, *D-brane Potentials in the Warped Resolved Conifold and Natural Inflation*, *JHEP* **1502** (2015) 127, [arXiv:1409.1221].

[208] S. W. Hawking, *The path-integral approach to quantum gravity*, *Advanced Series in Astrophysics and Cosmology* **8** (June, 1993) 163–206.

[209] N. Arkani-Hamed, J. Orgera, and J. Polchinski, *Euclidean wormholes in string theory*, *JHEP* **0712** (2007) 018, [arXiv:0705.2768].

[210] S. B. Giddings and A. Strominger, *Axion Induced Topology Change in Quantum Gravity and String Theory*, *Nucl.Phys.* **B306** (1988) 890.

[211] A. Strominger, *Vacuum topology and incoherence in quantum gravity*, *Phys. Rev. Lett.* **52** (May, 1984) 1733–1736.

[212] S. R. Coleman, *Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence*, *Nucl.Phys.* **B307** (1988) 867.

[213] S. B. Giddings and A. Strominger, *Loss of Incoherence and Determination of Coupling Constants in Quantum Gravity*, *Nucl.Phys.* **B307** (1988) 854.

[214] D. Gross, M. Perry, and L. Yaffe, *Instability of Flat Space at Finite Temperature*, *Phys.Rev.* **D25** (1982) 330–355.

[215] M. S. Volkov and A. Wipf, *Black hole pair creation in de Sitter space: A Complete one loop analysis*, *Nucl.Phys.* **B582** (2000) 313–362, [hep-th/0003081].

[216] E. Witten, *Instability of the Kaluza-Klein Vacuum*, *Nucl.Phys.* **B195** (1982) 481.

[217] S. R. Coleman, *The Fate of the False Vacuum. 1. Semiclassical Theory*, *Phys. Rev.* **D15** (1977) 2929–2936. [Erratum: Phys. Rev.D16,1248(1977)].

[218] S. R. Coleman and F. De Luccia, *Gravitational Effects on and of Vacuum Decay*, *Phys. Rev.* **D21** (1980) 3305.

[219] V. Rubakov and O. Y. Shvedov, *A Negative mode about Euclidean wormhole*, *Phys.Lett.* **B383** (1996) 258–261, [gr-qc/9604038].

[220] V. Rubakov and O. Y. Shvedov, *Instability of space-time due to Euclidean wormholes*, gr-qc/9608065.

[221] T. W. Grimm, *Axion Inflation in F-theory*, *Phys.Lett.* **B739** (2014) 201–208, [arXiv:1404.4268].

[222] N. Kaloper, *Stringy cosmic strings and axion cohomology*, *Phys.Rev.* **D47** (1993) 2403–2410, [hep-th/9206079].

[223] M. Aryal and A. E. Everett, *Gravitational effects of global strings*, *Phys. Rev. D* **33** (Jan, 1986) 333–337.

[224] D. Harari and A. P. Polychronakos, *A Global String With an Event Horizon*, *Phys.Lett.* **B240** (1990) 55.

[225] R. Gregory, *GLOBAL STRING SINGULARITIES*, *Phys.Lett.* **B215** (1988) 663.

[226] C. Long, L. McAllister, and J. Stout, *Systematics of Axion Inflation in Calabi-Yau Hypersurfaces*, arXiv:1603.0125.

[227] E. Silverstein and A. Westphal, *Monodromy in the CMB: Gravity Waves and String Inflation*, *Phys.Rev.* **D78** (2008) 106003, [arXiv:0803.3085].

[228] L. McAllister, E. Silverstein, and A. Westphal, *Gravity Waves and Linear Inflation from Axion Monodromy*, *Phys.Rev.* **D82** (2010) 046003, [arXiv:0808.0706].

[229] M. Berg, E. Pajer, and S. Sjors, *Dante's Inferno*, *Phys.Rev.* **D81** (2010) 103535, [arXiv:0912.1341].

[230] R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu, *Oscillations in the CMB from Axion Monodromy Inflation*, *JCAP* **1006** (2010) 009, [arXiv:0907.2916].

[231] E. Palti and T. Weigand, *Towards large r from [p, q]-inflation*, *JHEP* **1404** (2014) 155, [arXiv:1403.7507].

[232] F. Marchesano, G. Shiu, and A. M. Uranga, *F-term Axion Monodromy Inflation*, *JHEP* **1409** (2014) 184, [arXiv:1404.3040].

[233] R. Blumenhagen and E. Plauschinn, *Towards Universal Axion Inflation and Reheating in String Theory*, *Phys.Lett.* **B736** (2014) 482–487, [arXiv:1404.3542].

[234] L. McAllister, E. Silverstein, A. Westphal, and T. Wrase, *The Powers of Monodromy*, *JHEP* **1409** (2014) 123, [arXiv:1405.3652].

[235] S. Franco, D. Galloni, A. Retolaza, and A. Uranga, *On axion monodromy inflation in warped throats*, *JHEP* **1502** (2015) 086, [arXiv:1405.7044].

[236] R. Blumenhagen, D. Herschmann, and E. Plauschinn, *The Challenge of Realizing F-term Axion Monodromy Inflation in String Theory*, *JHEP* **1501** (2015) 007, [arXiv:1409.7075].

[237] A. Hebecker, S. C. Kraus, and L. T. Witkowski, *D7-Brane Chaotic Inflation*, *Phys.Lett.* **B737** (2014) 16–22, [arXiv:1404.3711].

[238] L. E. IbÃ¡nez and I. Valenzuela, *The inflaton as an MSSM Higgs and open string modulus monodromy inflation*, *Phys.Lett.* **B736** (2014) 226–230, [arXiv:1404.5235].

[239] A. Hebecker, P. Mangat, F. Rompineve, and L. T. Witkowski, *Tuning and Backreaction in F-term Axion Monodromy Inflation*, arXiv:1411.2032.

[240] L. E. Ibanez, F. Marchesano, and I. Valenzuela, *Higgs-otic Inflation and String Theory*, *JHEP* **1501** (2015) 128, [arXiv:1411.5380].

[241] A.-K. Kashani-Poor and A. Tomasiello, *A stringy test of flux-induced isometry gauging*, *Nucl.Phys.* **B728** (2005) 135–147, [hep-th/0505208].

[242] E. Witten, *Nonperturbative superpotentials in string theory*, *Nucl.Phys.* **B474** (1996) 343–360, [hep-th/9604030].

[243] C. Beasley and E. Witten, *New instanton effects in supersymmetric QCD*, *JHEP* **0501** (2005) 056, [hep-th/0409149].

[244] M. Billo, L. Ferro, M. Frau, F. Fucito, A. Lerda, *et. al.*, *Flux interactions on D-branes and instantons*, *JHEP* **0810** (2008) 112, [arXiv:0807.1666].

[245] M. Billo, L. Ferro, M. Frau, F. Fucito, A. Lerda, *et. al.*, *Non-perturbative effective interactions from fluxes*, *JHEP* **0812** (2008) 102, [arXiv:0807.4098].

[246] J. Anglin and R. C. Myers, *Wormholes and supersymmetry*, *Phys.Rev.* **D46** (1992) 2469–2479, [hep-th/9206072].

[247] G. W. Gibbons, M. B. Green, and M. J. Perry, *Instantons and seven-branes in type IIB superstring theory*, *Phys.Lett.* **B370** (1996) 37–44, [hep-th/9511080].

[248] E. Bergshoeff and K. Behrndt, *D - instantons and asymptotic geometries*, *Class.Quant.Grav.* **15** (1998) 1801–1813, [hep-th/9803090].

[249] E. Bergshoeff, A. Collinucci, U. Gran, D. Roest, and S. Vandoren, *Non-extremal D-instantons*, *JHEP* **0410** (2004) 031, [hep-th/0406038].

[250] E. Bergshoeff, A. Collinucci, U. Gran, D. Roest, and S. Vandoren, *Non-extremal instantons and wormholes in string theory*, *Fortsch.Phys.* **53** (2005) 990–996, [hep-th/0412183].

[251] P. W. Graham, D. E. Kaplan, and S. Rajendran, *Cosmological Relaxation of the Electroweak Scale*, arXiv:1504.0755.

[252] L. E. Ibanez, M. Montero, A. Uranga, and I. Valenzuela, *Relaxion Monodromy and the Weak Gravity Conjecture*, arXiv:1512.0002.

[253] H. K. Dreiner, *Hide and seek with supersymmetry*, in *Hidden symmetries and Higgs phenomena. Proceedings, Summer School, Zuoz, Switzerland, August 16-22, 1998*, 1998. hep-ph/9902347.

[254] J. R. Espinosa, C. Grojean, G. Panico, A. Pomarol, O. Pujolas, and G. Servant, *Cosmological Higgs-Axion Interplay for a Naturally Small Electroweak Scale*, arXiv:1506.0921.

[255] E. Hardy, *Electroweak relaxation from finite temperature*, *JHEP* **11** (2015) 077, [arXiv:1507.0752].

[256] S. P. Patil and P. Schwaller, *Relaxing the Electroweak Scale: the Role of Broken dS Symmetry*, arXiv:1507.0864.

[257] J. Jaeckel, V. M. Mehta, and L. T. Witkowski, *Musings on cosmological relaxation and the hierarchy problem*, arXiv:1508.0332.

[258] R. S. Gupta, Z. Komargodski, G. Perez, and L. Ubaldi, *Is the Relaxion an Axion?*, arXiv:1509.0004.

[259] B. Batell, G. F. Giudice, and M. McCullough, *Natural Heavy Supersymmetry*, arXiv:1509.0083.

[260] L. Marzola and M. Raidal, *Natural Relaxation*, arXiv:1510.0071.

[261] K. Choi and S. H. Im, *Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry*, arXiv:1511.0013.

[262] D. E. Kaplan and R. Rattazzi, *A Clockwork Axion*, arXiv:1511.0182.

[263] S. Di Chiara, K. Kannike, L. Marzola, A. Racioppi, M. Raidal, and C. Spethmann, *Relaxion Cosmology and the Price of Fine-Tuning*, arXiv:1511.0285.

[264] N. Kaloper and L. Sorbo, *A Natural Framework for Chaotic Inflation*, *Phys.Rev.Lett.* **102** (2009) 121301, [arXiv:0811.1989].

[265] N. Kaloper and A. Lawrence, *Natural chaotic inflation and ultraviolet sensitivity*, *Phys.Rev.* **D90** (2014), no. 2 023506, [arXiv:1404.2912].

[266] L. E. Ibáñez and I. Valenzuela, *BICEP2, the Higgs Mass and the SUSY-breaking Scale*, *Phys. Lett.* **B734** (2014) 354–357, [arXiv:1403.6081].

[267] M. Arends, A. Hebecker, K. Heimpel, S. C. Kraus, D. Lust, C. Mayrhofer, C. Schick, and T. Weigand, *D7-Brane Moduli Space in Axion Monodromy and Fluxbrane Inflation*, *Fortsch. Phys.* **62** (2014) 647–702, [arXiv:1405.0283].

[268] H. Hayashi, R. Matsuda, and T. Watari, *Issues in Complex Structure Moduli Inflation*, arXiv:1410.7522.

[269] I. GarcÃa-Etxebarria, T. W. Grimm, and I. Valenzuela, *Special Points of Inflation in Flux Compactifications*, arXiv:1412.5537.

[270] R. Blumenhagen, A. Font, M. Fuchs, D. Herschmann, E. Plauschinn, Y. Sekiguchi, and F. Wolf, *A Flux-Scaling Scenario for High-Scale Moduli Stabilization in String Theory*, *Nucl. Phys.* **B897** (2015) 500–554, [arXiv:1503.0763].

[271] A. Retolaza, A. M. Uranga, and A. Westphal, *Bifid Throats for Axion Monodromy Inflation*, *JHEP* **07** (2015) 099, [arXiv:1504.0210].

[272] D. Escobar, A. Landete, F. Marchesano, and D. Regalado, *Large field inflation from D-branes*, arXiv:1505.0787.

[273] S. Bielleman, L. E. Ibáñez, F. G. Pedro, and I. Valenzuela, *Multifield Dynamics in Higgs-Otic Inflation*, arXiv:1505.0022.

[274] R. Blumenhagen, C. Damian, A. Font, D. Herschmann, and R. Sun, *The Flux-Scaling Scenario: De Sitter Uplift and Axion Inflation*, arXiv:1510.0152.

[275] G. Dvali, *Three-form gauging of axion symmetries and gravity*, hep-th/0507215.

[276] R. Bousso and J. Polchinski, *Quantization of four form fluxes and dynamical neutralization of the cosmological constant*, *JHEP* **06** (2000) 006, [hep-th/0004134].

[277] S. Bielleman, L. E. Ibáñez, and I. Valenzuela, *Minkowski 3-forms, Flux String Vacua, Axion Stability and Naturalness*, arXiv:1507.0679.

[278] N. Kaloper, A. Lawrence, and L. Sorbo, *An Ignoble Approach to Large Field Inflation*, *JCAP* **1103** (2011) 023, [arXiv:1101.0026].

[279] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Cengage Learning, 1976.

[280] S. F. King, *Neutrino mass models*, *Rept. Prog. Phys.* **67** (2004) 107–158, [hep-ph/0310204].

[281] J. D. Brown and C. Teitelboim, *Neutralization of the Cosmological Constant by Membrane Creation*, *Nucl. Phys.* **B297** (1988) 787–836.

[282] M. J. Duncan and L. G. Jensen, *Four Forms and the Vanishing of the Cosmological Constant*, *Nucl. Phys.* **B336** (1990) 100.

[283] J. L. Feng, J. March-Russell, S. Sethi, and F. Wilczek, *Saltatory relaxation of the cosmological constant*, *Nucl. Phys.* **B602** (2001) 307–328, [hep-th/0005276].

[284] J. B. Griffiths and J. Podolský, *Exact Space-Times in Einstein's General Relativity*. Cambridge Monographs on Mathematical Physics, 2009.

[285] A. Hebecker, F. Rompineve, and A. Westphal, *Axion Monodromy and the Weak Gravity Conjecture*, *JHEP* **04** (2016) 157, [arXiv:1512.0376].

[286] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370–3373, [hep-ph/9905221].

[287] S. B. Giddings, S. Kachru, and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev.* **D66** (2002) 106006, [hep-th/0105097].

[288] 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].

[289] 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].

[290] P. G. Camara, L. E. Ibanez, and F. Marchesano, *RR photons*, *JHEP* **09** (2011) 110, [arXiv:1106.0060].

[291] F. Marchesano, D. Regalado, and G. Zoccarato, *U(1) mixing and D-brane linear equivalence*, *JHEP* **08** (2014) 157, [arXiv:1406.2729].

[292] A. Hanany and E. Witten, *Type IIB superstrings, BPS monopoles, and three-dimensional gauge dynamics*, *Nucl. Phys.* **B492** (1997) 152–190, [hep-th/9611230].

[293] S. Hellerman and I. Swanson, *Cosmological solutions of supercritical string theory*, *Phys.Rev.* **D77** (2008) 126011, [hep-th/0611317].

[294] A. Sen, *Tachyon dynamics in open string theory*, *Int.J.Mod.Phys.* **A20** (2005) 5513–5656, [hep-th/0410103].

[295] M. Atiyah and D. Anderson, *K-theory*. Advanced Books Classics Series. Westview Press, 1994.

[296] M. Karoubi, *K-theory. An elementary introduction*, *ArXiv Mathematics e-prints* (Feb., 2006) [math/0602082].

[297] M. Atiyah, R. Bott, and A. Shapiro, *Clifford modules*, *Topology* **3** (1964) S3–S38.

[298] S. Vandoren and P. van Nieuwenhuizen, *Lectures on instantons*, arXiv:0802.1862.

[299] B. J. Harrington and H. K. Shepard, *Periodic Euclidean Solutions and the Finite Temperature Yang-Mills Gas*, *Phys.Rev.* **D17** (1978) 2122.

[300] E. Witten, *Some Exact Multi - Instanton Solutions of Classical Yang-Mills Theory*, *Phys.Rev.Lett.* **38** (1977) 121–124.

[301] D. Gaiotto, A. Kapustin, N. Seiberg, and B. Willett, *Generalized Global Symmetries*, *JHEP* **02** (2015) 172, [arXiv:1412.5148].

[302] S. H. Shenker, *Another length scale in string theory?*, hep-th/9509132.

[303] M. R. Douglas, D. N. Kabat, P. Pouliot, and S. H. Shenker, *D-branes and short distances in string theory*, *Nucl. Phys.* **B485** (1997) 85–127, [hep-th/9608024].

[304] M. Spradlin, A. Strominger, and A. Volovich, *Les Houches lectures on de Sitter space*, in *Unity from duality: Gravity, gauge theory and strings. Proceedings, NATO Advanced Study Institute, Euro Summer School, 76th session, Les Houches, France, July 30-August 31, 2001*, pp. 423–453, 2001. hep-th/0110007.

[305] T. W. Grimm, T.-W. Ha, A. Klemm, and D. Klevers, *The D5-brane effective action and superpotential in N=1 compactifications*, *Nucl. Phys.* **B816** (2009) 139–184, [arXiv:0811.2996].