Department of Physics and Astronomy
University of Heidelberg

Master thesis
in Physics
submitted by
Daniel Kläwer
born in Dernbach
2016

SUPER-PLANCKIAN FIELD DISPLACEMENTS AND CONSISTENT QUANTUM GRAVITY

This Master thesis has been carried out by

Daniel Kläwer

at the

Institute for Theoretical Physics

under the supervision of

Dr Eran Palti

ACKNOWLEDGEMENTS

I thank Eran Palti for guidance throughout the whole duration of this thesis project and for always keeping in mind the big picture. I thank the University of Heidelberg and the Institute for Theoretical Physics for providing a superb environment for research to flourish. I thank Arthur Hebecker for his excellent lectures on mechanics which sparked my passion for theoretical physics and for his excellent supervision during my Bachelor thesis. As many other students of quantum field theory and string theory in Heidelberg, I am indebted to Timo Weigand for his great lectures. I would like to thank Fabrizio Rompineve and Lukas Witkowski for useful discussions.

Some of the best ideas often came up during discussions over a good cup of coffee so I would like to thank Philipp Arras, Julius Eckhard, Fabian Klos, Sascha Leonhardt, Jakob Moritz and Moritz Schöne.

I thank my parents, Anja and Rainer, for continuous support of my personal goals. Without them I would be far from where I am now. I also thank my sister Judith, in particular for extensive proofreading.

Finally, I thank my lovely girlfriend Elena for having the patience and calm to follow my path for such a long time. The persistence of her love is the greatest source of power — and she often provided welcome distraction.

ABSTRACT

We study the consistency of super-Planckian scalar field displacements in quantum gravity. The first part of this thesis discusses the cosmology of the early universe and in particular large field inflation, which explains the homogeneous initial conditions. Several models of natural inflation are presented. It is argued that the embedding of these models into a quantum gravity framework can be problematic because of the required super-Planckian scalar field displacements. Concrete obstructions to an implementation in string theory are presented. The second part is devoted to the study of various conjectures on the moduli space of quantum gravity theories which constrain large field displacements. The Weak Gravity Conjecture constrains mass-to-coupling ratios in gauge theories. In string theory these are functions of scalar moduli. We propose a connection between a variant of the Weak Gravity Conjecture and a Swampland Conjecture, which states that as a scalar field displacement asymptotes to infinity an infinite tower of states appears, exponentially light in the displacement. We show that the Weak Gravity Conjecture leads to evidence for this and that the exponential behaviour sets in quickly after the field variation passes the Planck scale. These conjectures can be used to constrain large field inflation models.

Wir betrachten trans-Plancksche skalare Feldauslenkungen in Quantengravitationstheorien. Der erste Teil dieser Arbeit beschäftigt sich mit der Kosmologie des frühen Universums und kosmologischer Inflation, welche die homogenen Anfangsbedingungen erklärt. Wir stellen verschiedene Modelle natürlicher Inflation vor und argumentieren, dass die Einbettung dieser in eine Quantengravitationstheorie aufgrund der benötigten trans-Planckschen Feldauslenkungen problematisch sein kann. Konkret wird dies am Beispiel der Stringtheorie erläutert. Im zweiten Teil beschäftigen wir uns mit verschiedenen Vermutungen über den Moduliraum von Quantengravitation, die große Feldauslenkungen einschränken. Die Weak Gravity Conjecture schränkt Masse-zu-Ladung-Verhältnisse in Eichtheorien ein. Diese sind in Stringtheorie als Funktionen von Skalarfeldern bestimmt. Wir schlagen eine mögliche Verbindung zwischen der Weak Gravity Conjecture und einer Swampland Conjecture vor, die besagt, dass im Limes unendlicher Skalarfeldauslenkungen ein unendlicher Turm von Zuständen in der Theorie asymptotisch exponentiell leicht wird. Wir zeigen mittels der Weak Gravity Conjecture, dass es Anzeichen gibt, dass dieses exponentielle Verhalten schnell für trans-Plancksche Auslenkungen einsetzt. Diese Vermutungen können "large field inflation"-Modelle einschränken.

Contents

List of Illustrations

Figures

Tables

1 Introduction

One hundred and one years after Albert Einstein's discovery of the general theory of relativity, the LIGO collaboration has finally announced the direct detection of gravitational waves on February 11, 2016 — confirming one of its most striking, and for a long time controversial, predictions [1, 2]. The source of the radiation could not have been more dramatic than a binary black hole merger. Even though theorists have long been convinced of the existence of gravitational waves and their field quanta, the gravitons, the indirect discovery of gravitational radiation through the orbital decay of the Hulse-Taylor binary pulsar [3] and now the direct detection of gravitational wave strain from coalescing black holes using a ground-based interferometer should have erased any doubt even among sceptics. Two events (GW150914 and GW151226) of binary black hole mergers have been registered during the first four months of operation of Advanced LIGO leading to the expectation that multiple detections per year should be considered normal, making gravitational wave astronomy possible in the near future. A very different discovery of gravitational radiation was claimed in March 2014 by the BICEP collaboration in the form of primordial gravitational waves [5]. These originate in the very early universe during inflation. Inflation is a phase of approximate de Sitter evolution (exponential expansion) in the early universe curing various fine tuning problems of the standard model of big bang cosmology (ΛCDM). The positive cosmological constant is typically provided by a scalar field slowly rolling down its potential towards a minimum. This inflaton field and the background metric have quantum fluctuations associated to them which also get inflated to macroscopic scale and later become visible as fluctuations in the cosmic microwave background (CMB) temperature. The fluctuations can be classified into tensor and scalar modes. The tensor modes influence the CMB polarisation in a very peculiar manner known as B-modes. Different models of inflation can give different values of the tensor to scalar ratio r. The BICEP discovery of $r = 0.20^{+0.07}_{-0.05}$ turned out to be wrong in the end due to the presence of a foreground signal generated by dust [6]. Nevertheless, future experiments probing the CMB might detect a non-zero r, connecting the mainly theoretically motivated paradigm of cosmological inflation to experiment. Single field inflation models can be broadly divided into small and large field inflation (LFI). In small field inflation models the inflaton travels a sub-Planckian distance and the shape of its potential is sufficiently flat such that slow roll occurs. In large field inflation the potential can be very simple, such as polynomial, because the inflaton starts far out at super-Planckian displacement where the slope of the potential is small relative to its magnitude. This still ensures friction dominated rolling. LFI is phenomenologically very attractive since a future detection of $r \gtrsim 0.01$ implies LFI because of the Lyth bound [8]. The price to pay is that the shape of the potential generically cannot be controlled over super-Planckian distances without extreme fine tuning. An initially flat potential will receive radiative corrections from generic Planck scale physics since at $\Delta\phi \gg M_p$ even Planck suppressed operators will become important. Thus, the potential has to be protected by a powerful symmetry to retain naturalness. A popular choice is to protect the in-

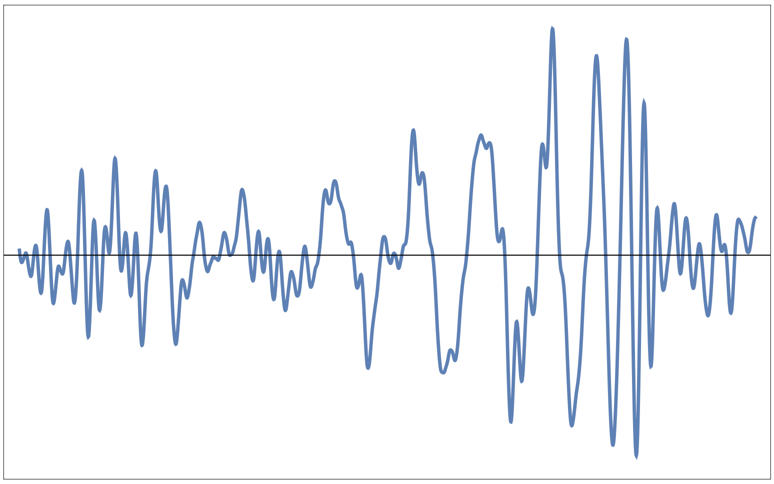

Figure 1.1: The gravitational wave event GW150914 as detected by the LIGO detector in Hanford. Plotted using data from [4].

flaton by a (weakly broken) continuous shift symmetry. This prohibits any explicit dependence of the corrections on ϕ and the only possible ones are derivative interactions which vanish at zero momentum.

Everything would be fine in effective field theory but quantum gravity (QG) puts constraints on this scenario. In fact, general arguments imply that QG does not allow for exact continuous global symmetries. In particular the continuous global shift symmetry of our inflaton candidate is forbidden. The best we can do is to allow for a discrete gauged shift symmetry as it is the case for the phase degree of freedom of a complex scalar. Such a scalar field with a discrete gauged shift symmetry will be dubbed an axion in the following. Even for axions it is to date still unclear if super-Planckian field displacements can be used for inflation. This is because for axions a periodic potential can in fact arise. Typically this potential is generated by instantons and higher order instantons induce higher harmonics of the potential. Even if these higher harmonics are sub-dominant it is still not clear that one can get a fundamental period of super-Planckian length. In fact, the Weak Gravity Conjecture (WGC) [9] applied to axions imposes strong constraints (see for example [10–16]). The WGC in general is a constraint on the tension-to-coupling ratio in p-form gauge theories. It states that there must exist a state of tension T and gauge coupling g such that in Planck units

$$T \lesssim g \,. \tag{1.1}$$

This can then be naturally extended to 0-form potentials, which are the axions [9]. The statement of the WGC is then that there should be an instanton of action S such that

$$S \lesssim M_P/f \,, \tag{1.2}$$

where f is the axion decay constant which characterizes the periodicity of the axion. Demanding a controlled instanton expansion and thus $S \ll 1$ one finds that the axion decay constant and thus moduli space diameter has to be sub-Planckian. We are thus led to the conclusion that an answer

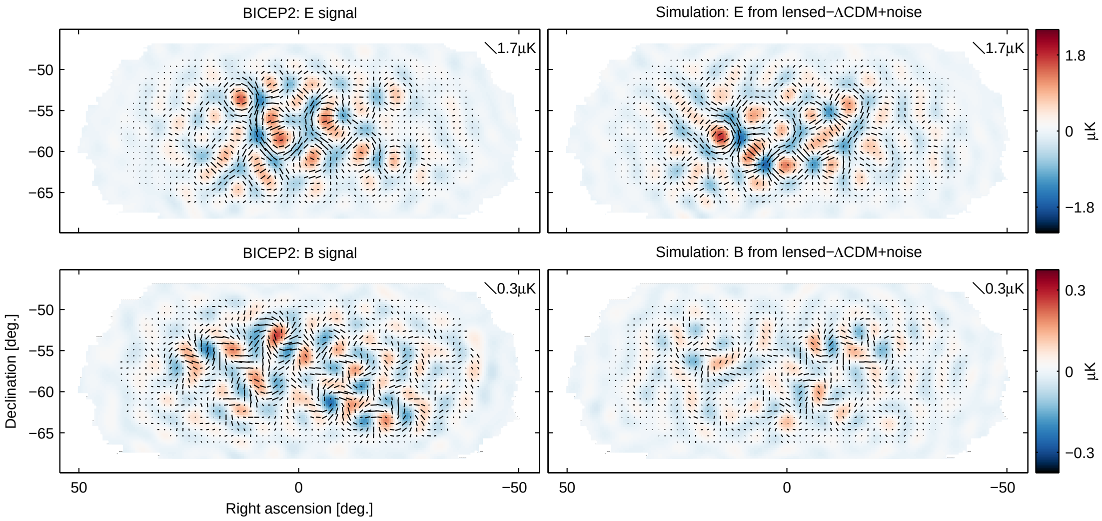

Figure 1.2: Polarisation of the Cosmic Microwave Background radiation as observed by BICEP2. The B-mode signal turned out to be due to foreground dust. Reproduced from [7].

to the question of the consistency of large field displacements in a QG theory is of significant theoretical and phenomenological interest.

In this thesis we will discuss the challenges of large field displacements in time *and* space in consistent QG theories. The first part of the thesis consists of a review of general aspects of LFI which is the main motivation for the second part consisting of a review of various conjectured properties of the moduli space of QG as well as original research [17] aiming to extend and connect two of these conjectures, namely the WGC and a conjecture of Ooguri and Vafa which we will name the Swampland Conjecture (SC, conjecture 2 of [18]).

After the introduction chapter 2 reviews the basics of cosmology and LFI in effective field theory (EFT). Particular emphasis is put on models of natural inflation where the inflaton is an axion. These provide very promising models for large field inflation from the EFT point of view. Axions also arise in very large numbers in typical string compactifications — each supersymmetric perturbative string vacuum contains at least one such axion. The problem is that the naive single field natural inflation is basically ruled out in string theory. The required super-Planckian decay constants do not arise from string theory [19] and the (strong) WGC and SC if true rule these out completely. We are then led to study generalisations of the natural inflation paradigm that naively evade these constraints. These are N-flation and various forms of alignment scenarios. A brief comment on axion monodromy is included, which will not be further persued. We then briefly comment on the consistency of these models in EFT.

In chapter 3 we adopt a less general point of view and review the difficulties of obtaining de Sitter vacua in string theory in general and specifically the large axion decay constants required for natural inflation models in controlled string theory constructions. For this we briefly discuss 10d SUGRA as the low energy effective field theory of the superstring and sketch how to obtain

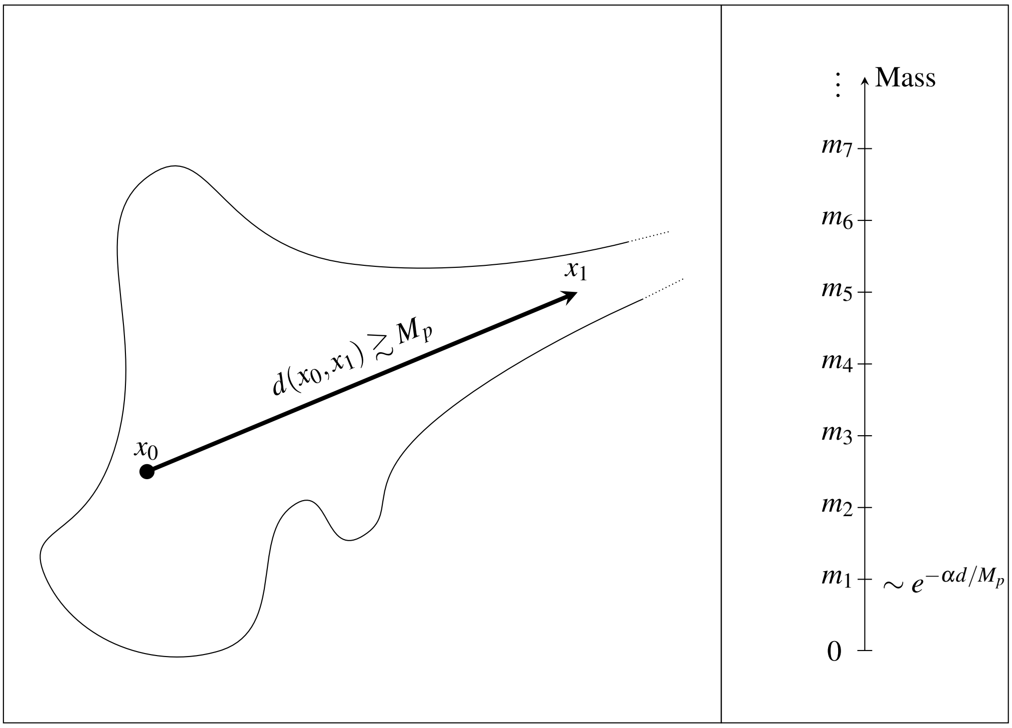

Figure 1.3: The scalar moduli space of a consistent quantum gravity theory. The Swampland Conjecture states that upon displacing over super-Planckian distances from a given effective field theory an infinite tower of exponentially light states must appear.

4d vacua with axions from this. Problems of each of the natural inflation models from 2 are addressed.

In chapter 4 we introduce and study several conjectures on the moduli space of QG. The first one is the swampland proposal [18]. The SC is the conjecture that if infinite distances in moduli space are traversed an infinite tower of states exponentially light in the distance should appear signalling the breakdown of effective field theory. This means that the moduli space of QG has infinite diameter but this cannot be probed in any fixed EFT. We then explain the absence of continuous global symmetries in QG both from general black hole arguments and from string theory. The WGC is presented as a quantitative sharpening of the absence of global symmetries. Electric and magnetic versions are explored and finally we introduce the lattice WGC which predicts an infinite tower of states satisfying the WGC. We ultimately want to relate the lattice WGC tower to the SC tower. In this context we briefly mention the completeness conjecture [20, 21], which demands that every site in the charge lattice has to be occupied. We conclude this chapter in commenting on the consistency of the various models of natural inflation from chapter 2 with the aforementioned conjectures.

In chapter 5 we arrive at the main part of the thesis. We first adopt the point of view that in a QG theory the gauge couplings and masses of WGC states are functions of scalar moduli and therefore the WGC should hold locally in space if these moduli have non-trivial profiles.

This is termed the local WGC. We furthermore introduce the refinement of the SC that the SC behaviour should not only hold for asymptotic distances in moduli space but already set in for finite distances of order the Planck scale. This is supported by an example of monodromy axions in type IIA string theory.

After introducing these ideas, in chapter 6 we present several compelling arguments for the identification of the lattice WGC tower of states with the SC tower of states in the case of a modulus controlling a spatially varying gauge coupling. We first present a bound on the variation of scalar fields in a weakly curved background which will be used to derive the functional form of the gauge kinetic terms in the regime of super-Planckian spatial moduli displacements. In particular we show that for a single spatially varying modulus which traverses a super-Planckian distance in moduli space and controls a gauge coupling the local WGC implies that the gauge coupling drops faster than exponentially in this distance for $\Delta\phi > \mathscr{O}(1)M_p$. We also provide evidence that after reaching a Planck scale displacement, the functional form of the gauge coupling becomes increasingly well described by an exponential, such that in the limit $\Delta\phi \to \infty$ the exponential form becomes exact. The tower of states of the lattice WGC thus has its mass scale set by a function exponential in the modulus displacement and the WGC implies the SC behaviour. We illustrate this behaviour by a monopole example in a weak curvature approximation as well as by a dilatonic, dyonic black hole solution which can be embedded in string theory. We conclude with an analysis for general strongly curved backgrounds, where we find the same behaviour by analysing the trace of the Einstein equations for a generic dilaton-gauge field system.

We use a mostly positive metric convention and define the Planck mass as $M_p^2 = 1/8\pi G$.

2 Large Field Inflation

In this chapter we will motivate the theory of cosmological inflation [22] by highlighting several shortcomings of the standard big bang theory of the early universe. The thermal history of the early universe is briefly reviewed. Then we study several models of natural inflation which are good candidates for large field inflation models from the bottom-up perspective and provide a testing ground for studying large field inflation in quantum gravity.

2.1 Basic Cosmology

Here we briefly review the basic model building ingredients for cosmology. The discussion is based on the book of Weinberg [23] and also [24]. The main assumption used for describing the cosmology of our spacetime is the assumption that at large scales the physics of our universe is homogeneous and isotropic. This is justified by the experimental finding that at scales above $\sim 3\times 10^8$ ly the amount of structure dramatically decreases [25], although there are some isolated structures at scales above $\sim 10^9$ ly such as the Sloan Great Wall. Under these assumptions the metric can be shown to take the FLRW form

$$g = -dt\otimes dt + a^2(t)\gamma(x)\ , \tag{2.1}$$

where γ is the metric of the $t=$ const. spacelike hyperplanes, which is of constant curvature type — positive, negative, or zero. It is sometimes convenient to keep track of all three possibilities by introducing a curvature parameter $K = 0,\pm 1$. The spatial metric can then be written as

$$\gamma = \frac{dr\otimes dr}{1-Kr^2} + r^2 d\Omega^2\ . \tag{2.2}$$

Distances measured with respect to this metric are so-called *comoving* distances since they stay constant during the cosmic expansion. These are to be contrasted with proper distances which include the time-dependent scale factor. The only degree of freedom of the metric, the time evolution of the scale factor $a(t)$ is determined via the Einstein equations by the form of the matter energy-momentum tensor. Homogeneity and isotropy require it to be of the perfect fluid form

$$T = \rho(t)dt\otimes dt + a^2(t)p(t)\gamma\,. \tag{2.3}$$

This must be supplemented with an equation of state $p = w\rho$, where

$$w = \begin{cases} 0 & \text{dust / cold matter}\ , \\ 1/3 & \text{radiation / hot matter}\ , \\ -1 & \text{cosmological constant / vacuum energy}\ . \end{cases} \tag{2.4}$$

Here dust means pressureless non-interacting forms of matter, radiation is any type of highly relativistic matter such that the trace of the energy momentum tensor vanishes and the cosmological constant is any source of energy-momentum proportional to the metric. All these matter constituents are needed to describe the universe we live in since the cosmological constant is measured to be small but positive.

The Einstein equations for this system are known as the *Friedmann equations* and these are conveniently expressed utilising the *Hubble parameter* $H = \dot{a}/a$ as

Friedmann Equations

$$M_p^2 H^2 = \frac{1}{3}\rho - M_p^2 \frac{K}{a^2}\,, \tag{2.5}$$

$$0 = \dot{\rho} + 3H(\rho + p)\,, \tag{2.6}$$

the second of which is just the continuity equation for an isotropic matter distribution. From this one finds the following scaling of the energy densities with the scale factor

$$\rho \propto \begin{cases} a^{-3} & \text{cold matter}\,, \\ a^{-4} & \text{radiation}\,, \\ 1 & \text{cosmological constant}\,. \end{cases} \tag{2.7}$$

During different stages of the cosmological evolution, the constituents of our universe played different roles. This is in part because they dilute with different powers of the scale factor and in part because of their temperature dependent dynamics which is not captured by a simple linear equation of state. These intricacies of the cosmological evolution are discussed in the next section. The continuity equation implies that the Hubble parameter can be expressed as

$$H^2(a) = H_0^2 \left[\Omega_{r,0}\left(\frac{a_0}{a}\right)^4 + \Omega_{m,0}\left(\frac{a_0}{a}\right)^3 + \Omega_{k,0}\left(\frac{a_0}{a}\right)^2 + \Omega_{\Lambda,0}\right]\,, \tag{2.8}$$

where the 0-subscript indicates quantities evaluated at present time and the Ω_i are the densities of the different matter components relative to the *critical density* $\rho_{\text{crit}} = 3M_p^2 H^2$. Finally, $\Omega_{k,0} := -k/(a_0 H_0)^2$. The value of the energy density of the universe relative to the critical density determines the sign of the curvature of the spatial slices. This is easily seen by rewriting (2.5) as

$$k = \frac{8\pi G a^2}{3}(\rho - \rho_{\text{crit}})\,, \tag{2.9}$$

so it follows that

$$\rho \begin{cases} < \rho_{\text{crit}} & \text{negative curvature}\,, \\ = \rho_{\text{crit}} & \text{no curvature}\,, \\ > \rho_{\text{crit}} & \text{positive curvature}\,. \end{cases} \tag{2.10}$$

For the case of a universe dominated by a single matter component one can explicitly integrate the Friedmann equation to get

$$a(t) \propto \begin{cases} t^{2/3} & \text{matter domination ,} \\ t^{1/2} & \text{radiation domination ,} \\ e^{Ht} & \text{vacuum energy domination ,} \end{cases} \tag{2.11}$$

and this can be a useful approximation during different stages of the evolution of our universe.

2.2 A Brief History of the Early Universe

The following discussion is based on the lecture notes of D. Baumann [26]. The evolution of the early universe (2.2) is dictated by the hierarchy of fundamental energy scales in nature (2.1). Right after the big bang the universe was in an extremely hot state and as it expanded it cooled down. For a given particle species of mass m as long as the temperature of the universe is $T > m$ we can effectively treat it as massless and thus it contributes to the energy density as radiation. As the universe cools down due to cosmic expansion and the temperature drops below a given particles mass the contribution to the energy density changes. Another effect is that at low enough temperatures interactions can freeze out. When the interaction rate Γ of a system of particles is smaller than the Hubble rate H, the universe expands fast enough such that the particles will not find each other in order to interact. By inspecting the interactions and masses of the standard model particles one can thus draw conclusions on the matter content of the universe at very early times. A very important event is the electroweak phase transition in which the initially massless gauge bosons of the weak force aquire masses of order 100 GeV. Above this scale every particle in the standard model can be treated as massless and the only dimensionful parameter in the theory is the temperature T. By dimensional analysis alone one can then infer that the interaction rate of electroweak processes scales like

$$\Gamma_{\text{ew}} \sim g_{\text{ew}}^2 T \ . \tag{2.12}$$

This one can compare to the Hubble scale which is again by dimensional analysis

$$H \sim \frac{\sqrt{\rho}}{M_p} \sim \frac{T^2}{M_p} \ . \tag{2.13}$$

Comparing the two, one finds that electroweak interactions are efficient for temperatures

$$10 \times 10^{16}\,\text{GeV} > T > 100\,\text{GeV} \ , \tag{2.14}$$

and we are dealing with a primordial relativistic plasma in thermal equilibrium. Thus the scale factor is approximately described by the radiation behaviour $a \sim t^{1/2}$ and the temperature decreases as $T \sim 1/a$. Below the EW symmetry breaking scale the behaviour of the interaction rate changes to the Sargent rule scaling

$$\Gamma_{\text{ew}} \sim G_F^2 T^5 \ . \tag{2.15}$$

Fundamental Energy Scales	
Planck Scale	10^{19} GeV
Grand Unification	10^{16} GeV
EW scale	100 GeV
QCD scale	150 MeV
Electron Mass	500 keV
H Ionisation Energy	13.6 eV

Table 2.1: List of fundamental energy scales that determine the evolution of the early hot universe.

Event	Time	Temperature
Inflation	?	-
Reheating	?	?
EW Phase Transition	2×10^{-11} s	100 GeV
QCD scale	2×10^{-5} s	150 MeV
Neutrino Decoupling	1 s	1 MeV
e^+e^- annihilation	6 s	500 keV
Nucleosynthesis	3 min	100 keV
$\rho_{\text{matter}} = \rho_{\text{radiation}}$	6×10^{4} yr	0.75 eV
Photon decoupling	3.8×10^{5} yr	0.25 eV
$\rho_{\text{matter}} = \rho_{\text{vac}}$	9×10^{9} yr	0.33 meV
This Thesis	1.38×10^{10} yr	0.24 meV

Table 2.2: Thermal History of our universe and related events. Adapted from [26].

Comparison of this to the Hubble scale leads to the conclusion that below $T \lesssim 1\,\text{MeV}$ the electroweak interactions freeze out. In particular neutrinos, which interact only through the weak force, decouple at this temperature. Before the neutrinos decouple we reach several mass scales of standard model particles. Since we are still in thermal equilibrium massive particles get an exponential Boltzmann suppression factor for $T < m$. Thus one can ignore massive particles at temperatures below their mass and the universe is still described well by a relativistic gas of equation of state

$$\rho = \frac{\pi^2}{30} g_*(T) T^4 \,, \tag{2.16}$$

where g_* is the effective number of relativistic degrees of freedom at a given temperature [26]. At $T \simeq 150\,\text{MeV}$ the QCD phase transition occurs and the quarks are bound into heavy hadrons such that the only remaining relativistic particles are pions, muons, electrons, neutrinos and photons in decreasing mass order. When the neutrinos finally decouple at $T \simeq 1\,\text{MeV}$, corresponding to around 1 s after the big bang, the relativistic matter content goes out of equilibrium and (2.16) stops being a good description of the equation of state of the universe. From this point until today the neutrinos preserve a relativistic Fermi-Dirac distribution with temperature decaying as $T \sim 1/a$. The temperature subsequently hits the electron mass and it becomes efficient for electron-positron pairs to annihilate. This leads to a heating of the photons relative to the neutrinos. After sufficiently many deuterium nuclei were formed the fusion to heavier nuclei, mainly helium, started at $T \simeq 100\,\mathrm{keV}$. This is termed *big bang nucleosynthesis*. While the temperature was still above $T \gtrsim 1\,\mathrm{eV}$ the universe consisted of light nuclei and electrons Compton scattering with photons to form a hot plasma. The electrons trying to recombine with the nuclei to form atoms were knocked out by high energy photons until the temperature reached $T \simeq 0.3\,\mathrm{eV}$. Finally at $T \simeq 0.25\,\mathrm{eV}$, or $t \simeq 380\,000\,\mathrm{yr}$, the photons interaction rate fell below the expansion rate of the universe and photons decoupled to form what is now known as the *cosmic microwave background* (2.3). At this point the non-relativistic matter content of the universe already dominated for a long time and the scale factor was well described by $a \sim t^{\frac{2}{3}}$ until after around $t \simeq 9 \times 10^9\,\mathrm{yr}$ the energy density provided by relativistic and non-relativistic matter were both sufficiently diluted such that the cosmological constant — or dark energy began dominating.

2.3 Fine-tuning Problems

The story of scales told in the last section is beautiful and simple. Nevertheless, in recent decades some problems with this simple model cosmology were realised. Our most accurate probe of the early universe is the cosmic microwave background (CMB). There are two main observables in the CMB — its local temperature and polarisation. The CMB temperature is extremely isotropic, which poses an immediate problem. The time between the hot big bang and CMB decoupling is about 380 000 years, whereas our observation of isotropy occurs around 14 Gyrs after the decoupling. If we trace back two CMB photons coming from opposite directions in the sky these have to come from two patches which have never been in causal contact since the big bang. This uncaused correlation between the different patches of the CMB sky is a huge fine-tuning problem of the standard big bang cosmology. We have to assume that the initial conditions of our universe were fine-tuned to be as isotropic as observed. This is the *horizon problem*.

The situation is illustrated in figure (2.3). To discuss issues of causality it is convenient to introduce a *conformal time* variable τ such that the spacetime metric becomes conformal to Minkowski space. This is achieved by the definition

$$\tau = \int \frac{dt}{a(t)} \,. \tag{2.17}$$

In the coordinate system defined by (τ, r) light rays propagate along 45° lines. While in Minkowski space we can detect events from arbitrarily large distances if these happened far enough in the past this is not necessarily true in a big bang cosmology since there is a beginning of time. This is termed the *particle horizon* since particles beyond this horizon will never get in causal contact with us. Similarly we may not be able to send out signals to infinite radial distance in the future because the expansion of the universe might be fast enough to outperform the finite speed of light. This is termed the *event horizon*, since events beyond this horizon will never be in causal contact to us. These are quantified in terms of horizon radii as follows. We define the

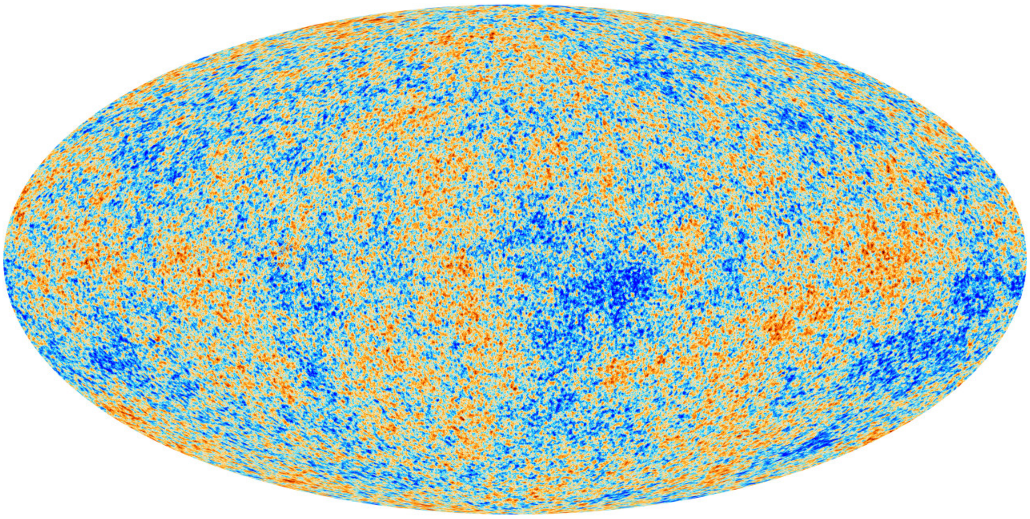

Figure 2.3: Temperature fluctuations in the Cosmic Microwave Background. The pictured fluctuations are of order 10^{-4}. Reproduced from [27].

infinitesimal comoving distance element as

$$dR = \frac{dr}{\sqrt{1-kr^2}} \,. \tag{2.18}$$

Then the comoving radius of the particle horizon at time t is given by

$$R_{\text{ph}}(t) = \int_0^R dR = \int_0^t d\tau = \int_0^t \frac{dt'}{a(t')} \,. \tag{2.19}$$

The proper horizon distance is then just

$$d_{\text{ph}}(t) = a(t) R_{\text{ph}}(t) \,. \tag{2.20}$$

If we assume based on the scaling (2.7) that at early times the universe was dominated by radiation one can easily see that the integral converges and there is such a particle horizon. In a completely analogous fashion the comoving radius and proper distance of the event horizon is

$$\begin{aligned} R_{\text{eh}}(t) &= \int_t^\infty \frac{dt'}{a(t')} \,, \\ d_{\text{eh}}(t) &= a(t) R_{\text{eh}}(t) \,. \end{aligned} \tag{2.21}$$

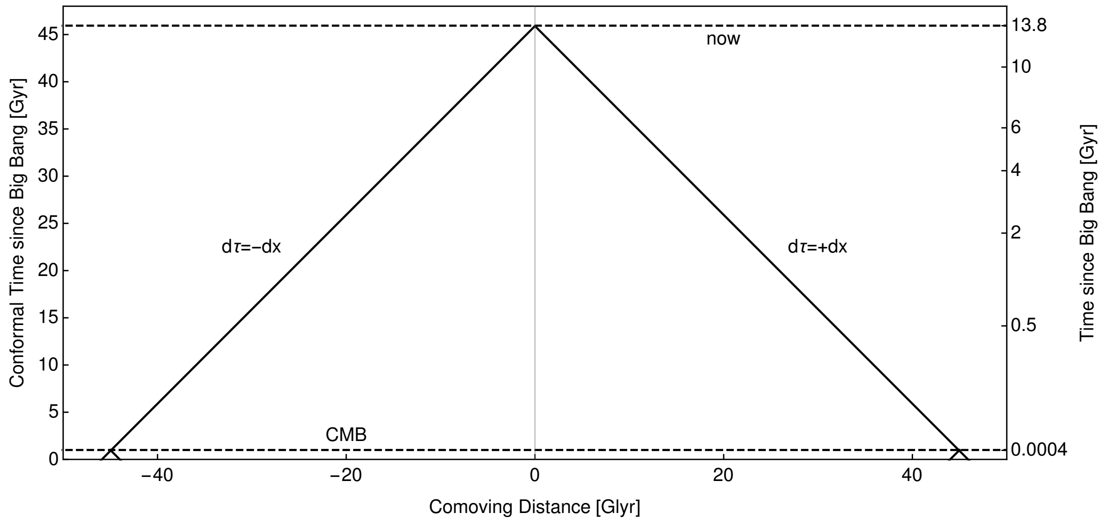

Figure 2.4: Two CMB photons travelling in opposite directions and the past light cones of their creation. It is obvious that these could have never been in causal contact, assuming radiation domination and thus a big bang in the early universe. Under this assumption, the CMB sky would be composed of order $\sim 10^4$ causally disconnected patches.

For cold and relativistic matter the integral now diverges so there is no event horizon, but if there is a source of vacuum energy this will eventually dominate and the then finite horizon distance will asymptote to $d_{\rm eh} \to 1/H$. While a possible event horizon of our universe is ultimately only asymptotically relevant, the finite size of the particle horizon in the conventional big bang cosmology is an immediate problem since it is in conflict with the CMB observations as described above. We can also approach the problem from a different angle. The change in proper distance with time for a particle at comoving distance $R(t)$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} d(R,t) = \underset{\text{A}}{\dot{a}R} + \underset{\text{B}}{a\dot{R}} \,, \tag{2.22}$$

the term A describing the velocity of recession of the particle due to the expansion of the universe while term B is the particles own motion. For a particle at rest it is then clear that the velocity of recession reaches the speed of light at the *comoving Hubble radius*

$$R_H = \frac{1}{aH} \,. \tag{2.23}$$

It descibes a sort of instantaneous measure of the horizon size. In an accelerating universe the comoving Hubble radius shrinks since particles which are initially subluminally receeding will eventually surpass the speed of light. In a decelerating universe the opposite effect takes place and the comoving Hubble radius expands. One can rewrite the particle horizon radius as an integral over the comoving Hubble radius as

$$R_{\rm ph}(t) = \int_0^t \frac{dt'}{a(t')} = \int_{\ln 0}^{\ln a(t)} \frac{d\ln a}{aH} \,. \tag{2.24}$$

Since in the standard big bang cosmology we have that the comoving Hubble radius grows with time the integral is dominated from the upper boundary and we are again led to the conclusion that the distance that particles could have travelled between the big bang and the CMB decoupling is much smaller than it was between then and today. So to resolve the horizon problem we need to add conformal time before the CMB is generated or equivalently add a period of shrinking comoving Hubble radius to the pre CMB evolution.

Another problem with the standard big bang picture is the *flatness problem*. Our universe is measured to be flat to extremely high accuracy. The strongest constraint available on the curvature of its spatial slices is given by the Planck collaboration [28] as

$$\Omega_K = 0.000 \pm 0.005 \qquad @\ 95\%\ \mathrm{CL}\ . \tag{2.25}$$

This means that either our universe is exactly flat (K=0) or the initial conditions at the beginning of the universe were so extremely fine-tuned such that the energy density as of today is still very close to the critical density even if it departs from it over time according to (2.8). To explain the observed degree of flatness one can estimate from the scaling of the curvature parameter that it would have been smaller than $\sim 10^{-19}$ at the time of electron-positron annihilation.

Grand Unified Theories usually predict the existence of magnetic monopoles. These monopoles exist if the GUT gauge group G satisfies the topological criterion [29]

$$\pi_2\left(G/SU(3)\times SU(2)\times U(1)\right) \neq \{1\}\ . \tag{2.26}$$

This is for example the case for any simple G. If these monopoles indeed exist they should be pair-created in the very early universe above the GUT scale and have to be sufficiently diluted to be compatible with their non-detection.

The last problem we will discuss is the explanation of the CMB fluctuations. Even if the CMB is extremely isotropic there still has to be some natural explanation for the tiny deviations from the average temperature, which is not provided by the ΛCDM model.

2.4 Inflation as a Solution

Inflation is a natural solution to the aforementioned fine-tuning problems. The paradigm of inflation assumes that before the hot big bang there was a period of approximately exponential expansion of the universe. Such an exponential and hence accelerated expansion can be driven by a positive cosmological constant. For an exponential scale factor the metric takes the form

$$ds^2 = -dt^2 + e^{2Ht}\left(dr^2 + r^2 d\Omega^2\right)\ . \tag{2.27}$$

While this looks inherently time dependent it can actually be cast into a manifestly static form by the coordinate transformation

$$\tilde{t} = t - \frac{1}{2H}\ln\left(-\frac{1}{H^2} + r^2 e^{2Ht}\right)\ , \quad \tilde{r} = re^{Ht}\ , \tag{2.28}$$

upon which the metric then takes the de Sitter form

$$ds^2 = -\left(1 - H^2\tilde{r}^2 d\tilde{t}^2 + \frac{d\tilde{r}^2}{1-H^2\tilde{r}^2} + \tilde{r}^2 d\Omega^2\right)\ . \tag{2.29}$$

This obviously admits a timelike Killing vector and thus cannot represent an evolving universe. In order to do cosmology we have to spontaneously break the time translation symmetry and promote the cosmological constant to a dynamical field. This is the basic idea of inflation. We seek for an effective field theory which produces an approximately constant energy density over some time which then decays, initiating the hot big bang. If the expansion occurs over a sufficient number of e-foldings the CMB photons observable to us could have all originated from one single causal patch. Thus the horizon problem is resolved because during such a phase of quasi de Sitter evolution the comoving Hubble radius shrinks rapidly as $\sim e^{-Ht}$.

In order to solve the flatness and horizon problems one can show that the amount of expansion during inflation must satisfy

$$e^{\mathcal{N}_e} > \frac{a_I H_I}{a_0 H_0} \,, \tag{2.30}$$

where the subscripts $I, 0$ denote the values at the end of inflation and the present ones. If one assumes that the values for the scale factor and Hubble scale at the end of inflation are roughly the same as during the radiation dominated era then one can relate this to the energy density at the beginning of radiation domination. The requirement that this should be at most of order the Planck scale then leads to [23]

$$\mathcal{N}_e \gtrsim 68 \,. \tag{2.31}$$

If the energy density was smaller at the beginning of radiation domination then a lower number of e-folds is required. One can then also show that this is sufficient to resolve the monopole problem [23].

In the most simple models of inflation the required vacuum energy is provided by the potential energy of a minimally coupled scalar field. The corresponding action is given by

$$S = \int \left[\frac{M_p^2}{2} \star R - \tfrac{1}{2} d\phi \wedge \star d\phi - \star V(\phi) \right] \,, \tag{2.32}$$

with equations of motion for homogeneous ϕ given by

$$3M_p^2 H^2 = \tfrac{1}{2}\dot{\phi}^2 + V(\phi) \,, \quad \ddot{\phi} + 3H\dot{\phi} + \partial_\phi V(\phi) = 0 \,, \tag{2.33}$$

so evidently the expansion of the universe provides a *Hubble friction* term for the scalar field. What are the conditions for inflation? We have seen that we necessarily need a shrinking comoving Hubble radius. This can be seen to imply that

$$\varepsilon_H := -\frac{\dot{H}}{H^2} < 1 \,. \tag{2.34}$$

Furthermore, we will require that the relative rate of change in ε is small compared to the Hubble scale H

$$\eta_H := \left| \frac{\dot{\varepsilon}}{H\varepsilon} \right| < 1 \,, \tag{2.35}$$

which is necessary for inflation to last over multiple Hubble times hence a large number of e-foldings $\mathcal{N}_e$. We will assume that both so-called *Hubble slow roll parameters* ε_H, η_H are sufficiently small such that the dynamics is approximately descibed by an exponential scale factor over an appropriate number of e-foldings $\mathcal{N}_e$. These conditions are very general conditions on inflationary models and valid beyond the simple model of a single scalar field. For the scalar field one can see that smallness of ε_H, η_H is ensured if the *potential slow roll parameters*

Slow Roll Conditions

$$\varepsilon := \frac{M_p^2}{2}\left(\frac{V'}{V}\right)^2 \overset{!}{\ll} 1\,, \quad \eta := M_p^2\frac{|V''|}{V} \overset{!}{\ll} 1\,, \tag{2.36}$$

are much smaller than one. In this regime they are then related to the Hubble slow roll parameters via $\varepsilon \approx \varepsilon_H$ and $\eta \approx 2\varepsilon_H - \eta_H/2$. In this case we have an overdamped motion of ϕ in its potential. One can show that the equations of motion approximately reduce to

$$3M_p^2H^2 = V(\phi)\,, \quad 3H\dot{\phi} + \partial_\phi V(\phi) = 0\,. \tag{2.37}$$

In particular we have a quite simple expression for the Hubble scale H in terms of the scalar potential V. By (2.37), number of e-folds during inflation starting from a given time t_{start} up to a given time t_{end} can be calculated in a given model as

$$\mathcal{N}_e = \int_{t_{\text{start}}}^{t_{\text{end}}} H(t')dt' \simeq \int_{\phi_{\text{end}}}^{\phi_{\text{start}}} \frac{d\phi}{M_p}\frac{1}{\sqrt{2\varepsilon}}\,. \tag{2.38}$$

Another important feature of inflation is that it actually predicts fluctuations in the CMB. These are the quantum fluctuations of the inflaton and graviton which are generated during inflation and then blown up to macroscopic size. The magnitude of these fluctuations can be calculated by QFT in curved spacetime techniques. The modes of a scalar field satisfy the *Mukhanov-Sasaki* equation

$$\ddot{v}_k + 3H\dot{v}_k + \frac{k^2}{a^2}v_k = 0\,. \tag{2.39}$$

At sufficiently early times the oscillation frequency $\omega_k = k^2/a^2$ of a given mode of momentum k is well above the Hubble scale. This means that the friction term can be neglected and we have free oscillations. As the comoving Hubble radius shrinks, at some point the friction starts to dominate and the mode is frozen out. This happens when the physical momentum of the mode crosses the Hubble scale. This *horizon crossing* happens when $k = aH$. These frozen fluctuations are then imprinted into the CMB later on.

One can derive the following dimensionless fluctuation power spectra [24] [1]

$$\Delta_s^2 = \frac{1}{24\pi^2}\frac{1}{\varepsilon}\frac{V}{M_p^4}\,, \quad \Delta_t^2 = \frac{2}{3\pi^2}\frac{V}{M_p^4}\,. \tag{2.40}$$

Here s,t indicate scalar and tensor fluctuations. There are two characteristic quantities which descibe the basic properties of these spectra. The first is the tensor to scalar ratio $r = \Delta_t^2/\Delta_s^2$

[1] CMB observations are experimentally probing the CMB spectrum around a given fiducial *pivot scale* k_*. The spectra are understood to be evaluated at the point in field space at which the pivot scale modes cross the horizon.

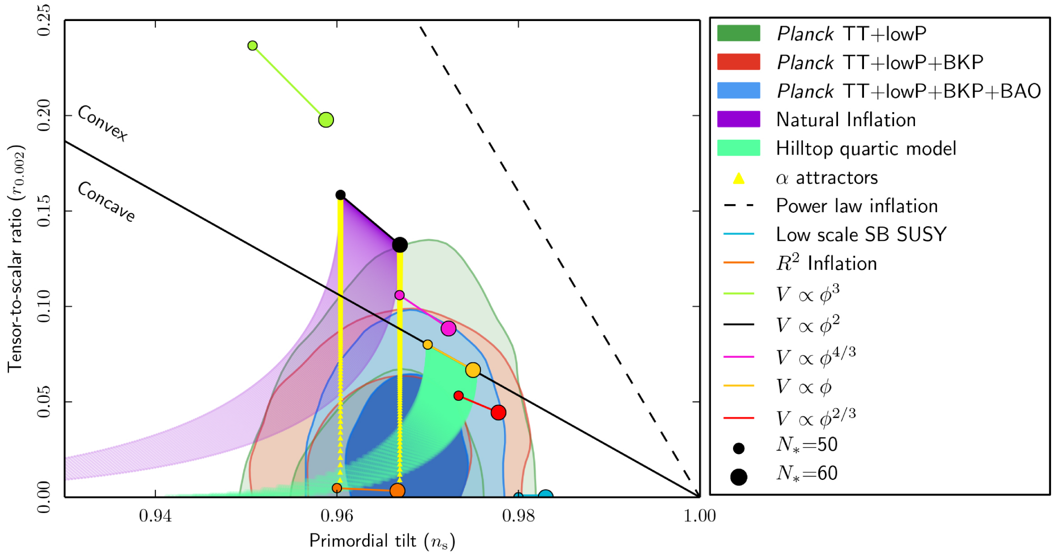

Figure 2.5: Exclusion plot comparing several inflation models consistency with Planck CMB observations in the $r-n_s$ plane. Models with convex polynomial potentials are disfavoured. Reproduced from [30].

which is simply the relative magnitude of these types of fluctuations. The second is the spectral index which describes the deviation from a scale invariant[2] spectrum around the pivot scale

$$\Delta_s^2 \simeq \mathscr{A}_s \left(\frac{k}{k_*}\right)^{n_s-1} . \tag{2.41}$$

They are given by

$$n_s = 1 + 2\eta - 6\varepsilon , \quad r = 16\varepsilon . \tag{2.42}$$

To extract the basic inflationary parameters from a given model of inflation one can compute the slow roll parameters, then calculate the number of e-folds $\mathscr{N}_*$ remaining when the inflaton passed the pivot scale at ϕ_* via (2.38), invert to extract the relevant quantities (2.40,2.41) in terms of $\mathscr{N}_*$ and finally impose a certain number of e-folds (for example $\mathscr{N}_* = 68$) to ensure the solution of the fine tuning problems. Comparing the computed values for different models with the observed CMB data leads to exclusion plots in the $r-n_s$ plane as in figure (2.5). A very simple class of models is given by the family of monomial potentials

$$V(\phi) = M^{4-\alpha}\phi^{\alpha} . \tag{2.43}$$

[2] Inflation predicts a nearly scale invariant power spectrum.

Quantity	Measured Value	Confidence Level
$r_{0.002}$	< 0.114	95%
n_s	0.9655 ± 0.0062	68%
$10^9 \mathscr{A}_s$	$2.198^{+0.076}_{-0.085}$	68%

Table 2.6: Observational results from CMB measurements by the Planck satellite (2015 TT+lowP, TT+lensing for r [28]).

The basic inflationary parameters are

$$\begin{array}{c} \varepsilon = \frac{1}{2}\alpha^2 \left(\frac{M_p}{\phi}\right)^2 \\ \mathscr{N}_e \simeq \frac{1}{2\alpha}\frac{\phi^2_{\text{start}}}{M_p^2} \end{array} \qquad \eta = \alpha|\alpha-1|\left(\frac{M_p}{\phi}\right)^2 \quad . \tag{2.44}$$

In this case the smallness of the slow roll parameters amounts to the requirement that $\phi \gtrsim \alpha M_p$, that we either need super-Planckian field displacements to begin with or an extremely flat potential. This illustrates two different philosophies in inflationary model building. For generic potentials one possible way of enforcing the slow roll speed limits (2.36) is to displace the scalar field far enough from its minimum such that the size of the potential dominates its slope — this is the idea of *large field inflation*. Another way is to arrange for a situation where the slope of the potential is itself sufficiently flat. In this case one does not need large field values and can get along with *small field inflation*. Since the general topic of this thesis is the analysis of super-Planckian field displacements we will focus on large field inflation. As was mentioned in the introduction, this also has a motivation since because of the Lyth bound [8],

Lyth bound

$$\frac{\Delta\phi}{M_p} \gtrsim 0.25\left(\frac{r}{0.01}\right)^{1/2} , \tag{2.45}$$

a detection of tensor modes at the level of $r \sim 0.01$ is the smoking gun signature of large field inflation.

2.5 Problems of Inflation

Despite being an elegant solution to the various fine-tuning problems of the early universe cosmology the implementation of the inflationary paradigm in effective field theory leads to a few possible problems. A very nice account of the treatment of inflation in effective field theory is given in [24]. In particular radiative corrections to the potential could lead to higher order corrections to the potential such that the smallness of ε, η is in fact spoiled for large $\Delta\phi$. For describing inflation in a consistent effective field theory from a bottom-up perspective the Wilsonian cutoff scale Λ of our effective description should satisfy

$$H \lesssim \Lambda \lesssim M_p \,. \tag{2.46}$$

The effective field theory Lagrangian will then contain higher dimension operators suppressed by appropriate powers of the cutoff and dressed with dimensionless Wilson coefficients

$$\mathscr{L} \supset \sum_i c_i \frac{\mathscr{O}_i[\phi, g]}{\Lambda^{\delta_i - 4}} \,. \tag{2.47}$$

In particular, the mass of a scalar inflaton is renormalised quadratically in the cutoff

$$\Delta m^2 \sim \Lambda^2 \,, \tag{2.48}$$

and hence radiative corrections destroy the parametric smallness of the slow roll parameter η, which changes upon keeping the Hubble scale fixed as

$$\delta\eta \simeq \frac{\Lambda^2}{H^2} \gtrsim 1 \,. \tag{2.49}$$

Even though supersymmetry naturally protects scalar masses due to their running being reduced to their fermionic partners logarithmic one, it cannot solve this problem since supersymmetry is necessarily broken at $\Lambda \sim H$ and so the scalar masses then naturally get driven to the Hubble scale, implying that η still gets corrections of order one. This is the so-called *eta problem*. The two possible solutions to it are either fine-tuning or protecting the potential by a symmetry. The degree of severeness of the eta problem differs dramatically between small and large field inflation models. In small field inflation models one can show that control over η requires at least control over all Planck suppressed operators up to dimension 6 [24]. This means that one needs to know at least the leading order corrections coming from the UV completion of the specific inflationary model under consideration. In large field inflation the situation can be much more dramatic. For $\Delta\phi > M_p$ an infinite number of possible radiative corrections can become relevant. This is because the integrating out of Planck scale fields might in principle induce every possible operator suppressed by the Planck scale in the effective action [24]. Even though an inflaton with only a very small mass term is technically natural because of the approximate shift symmetry[3], the UV completion of a given model in quantum gravity will not necessarily respect this. In particular there might be degrees of freedom that are needed for quantum gravitational consistency of the theory at high energies which spoil a particular model in *any possible* UV completion. In fact such strong constraints are suggested by the *Weak Gravity Conjecture* and the *Swampland Conjecture* which we will both introduce in chapter 4.

A well-motivated approach to implement large field inflation in a possibly UV complete way is to equip the scalar inflaton with an *exact* but discrete shift symmetry. Such an "angular" degree of freedom is usually called an axion, irrespective of its relation to the QCD axion. The only possible potential compatible with the symmetry is periodic and can be decomposed into

[3]The corrections to the potential will be proportional to the symmetry breaking parameter m and hence controllably small.

Fourier components. Models using such axions for inflation are termed *natural inflation* [31]. These are actually strongly disfavored by the CMB observations as of 2015 [30], see also figure (2.5). Nevertheless, it is worthwhile to study them in order to understand the consistency of large field inflation in a UV completion such as string theory.

2.6 Aspects of Natural Inflation

The most naive model of natural inflation [31] consists of a single axion with a periodic potential.

$$\mathscr{L} = -\tfrac{1}{2}(\partial\theta)^2 - \Lambda^4\left(1-\cos(\theta/f)\right) . \tag{2.50}$$

The dimensionful constant f governs the periodicity of θ or equivalently its charge under the discrete shift symmetry. It is usually called the axion decay constant. The constant Λ is usually related to the energy scale of the non-perturbative effects (instantons) that generate this potential. At small θ/f the theory effectively reduces to one of a massive scalar field with mass given by $m = \Lambda^2/f$. It is thus clear that in this regime we need $M_p \ll \Delta\theta \ll f$ and thus a super-Planckian decay constant. The basic inflationary parameters of this model are computed to be

$$\begin{aligned} \varepsilon &= \frac{1}{2}\left(\frac{M_p}{f}\right)^2\left(\frac{\sin(\theta/f)}{1-\cos(\theta/f)}\right)^2 , \\ \mathscr{N}_e &= 2\left(\frac{f}{M_p}\right)^2 \ln\left(\frac{\cos(\theta_{\text{end}}/2f)}{\cos(\theta_{\text{start}}/2f)}\right) . \end{aligned} \qquad \eta = \left(\frac{M_p}{f}\right)^2\left|\frac{\cos(\theta/f)}{1-\cos(\theta/f)}\right| , \tag{2.51}$$

Also from this it is clear that the simultaneous smallness of both slow roll parameters over a broad range requires $f \gtrsim M_p$. As we will discuss in chapter 3 obtaining such super-Planckian decay constants is a non-trivial task in string theory [19] and as of today no fully controlled constructions are known which produce these. This motivates us to try to evade super-Planckian decay constants already at the level of effective field theory and this is the setting for the next sections. We will discuss several such mechanisms which have been proposed in the literature to construct large field inflation models inspired by natural inflation which do not involve super-Planckian decay constants.

2.6.1 Effective Field Theory for N Axions

Since single axions are always bound to have sub-Planckian field ranges in string theory it is natural to consider more carefully the interplay of many axions. We will follow closely the notational conventions of [32]. The most general Lagrangian for N axions coupling to $P \geq N$ instantons, generating a non-perturbative potential, is given by

$$\mathscr{L} = \frac{1}{2}K_{ij}\partial\theta^i\partial\theta^j - \sum_{i=1}^{P}\Lambda_i^4\left[1-\cos\left(\mathscr{Q}^i{}_j\theta^j\right)\right] , \tag{2.52}$$

where K_{ij} is the metric on field space and $\mathscr{Q}$ is a $P \times N$ charge matrix describing the coupling of linear combinations of the θ to the P instantons. An UV scale Λ is introduced for each instanton for dimensional reasons, which is undetermined in the effective field theory. The loci of the first

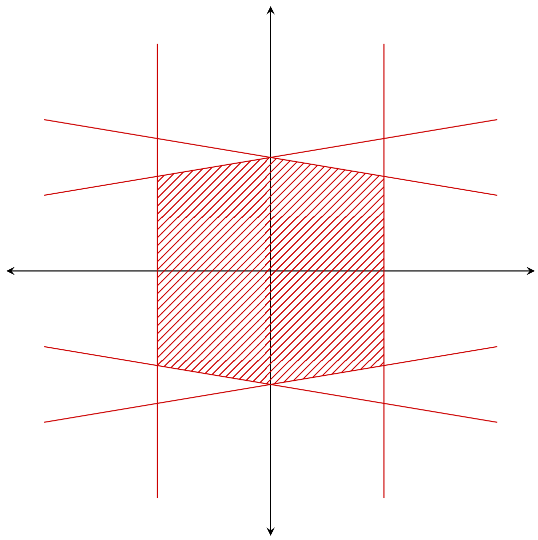

Figure 2.7: The axion fundamental domain for $N = 2$, $P = 3$.

maxima of each of the cosine terms in the potential define P hyperplane pairs in $\mathbb{R}^n$. In general position and for $P \geq N$, these in turn define a convex polytope whose interior is the *fundamental domain* for the θ fields (2.7). All other points in field space can be related to points in it by symmetry operations

$$\mathscr{Q}^i{}_j \theta^j \mapsto \mathscr{Q}^i{}_j \theta^j + 2\pi \,. \tag{2.53}$$

A conservative estimate for the available field range for large field inflation is the diameter $\mathscr{D}$ of the fundamental domain measured in terms of the quadratic form K. The field range can in principle be larger as is illustrated in figure (2.8). For the case $N = P$ (and assuming full rank of the charge matrix, i.e. no exactly flat directions) we can always change to the basis $\Phi = \mathscr{Q}\theta$ so the Lagrangian takes the form

$$\mathscr{L} = \frac{1}{2} (\partial\Phi)^T \, \Xi \, \partial\Phi - \sum_{i=1}^{N} \Lambda_i^4 \left[1 - \cos\left(\Phi^i\right)\right] \,, \qquad \Xi = (\mathscr{Q}^{-1})^T K \mathscr{Q}^{-1} \,. \tag{2.54}$$

Note that the fundamental domain of the Φ fields is now simply a hypercube. For the case where we have more terms in the potential than axions the potential cannot be completely diagonalised. The best one can do is to diagonalise N terms, leaving $P - N$ cross-couplings. Explicitly, order the rows of $\mathscr{Q}$ such that the first N rows form a full rank $N \times N$ submatrix Q, while the remaining $P - N$ rows form a $(P - N) \times N$ cross-coupling submatrix Q_R

$$\mathscr{Q} = \begin{pmatrix} Q \\ Q_R \end{pmatrix} \,. \tag{2.55}$$

Defining $\Phi = Q\theta$ and $\Xi = (Q^{-1})^T K Q^{-1}$, one has

$$\mathscr{L} = \frac{1}{2} (\partial\Phi)^T \, \Xi \, \partial\Phi - \sum_{i=1}^{P-N} \Lambda_i^4 \left[1 - \cos\left(\left(Q_R Q^{-1}\Phi\right)^i\right)\right] \,. \tag{2.56}$$

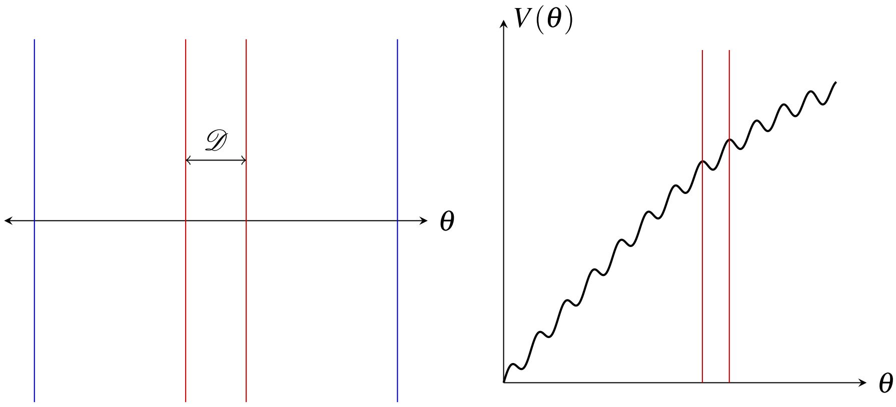

Figure 2.8: Two instanton contributions to the potential of an axion. The field range can in principle be vastly larger than the diameter of the fundamental domain. While the red instanton determines the diameter of the fundamental domain, it is sub-dominant to the blue one ($\Lambda_{\text{red}} \ll \Lambda_{\text{blue}}$). The blue instanton dominates the potential and for a large hierarchy also the effective field range. Exponential hierarchies can easily arise in string theory.

In this more general case the fundamental domain will be some general convex polytope, even in the Φ basis. This can be seen in figure (2.7) – one can always bring any one of the almost horizontal line pairs to a right angle with respect to the vertical line pair to create a square by a $GL(2,\mathbb{R})$ shear transformation. But this square will then have two corners cut off by the remaining pair of lines, still giving a hexagonal fundamental domain.

2.6.2 Alignment

One mechanism to achieve super-Planckian field ranges was proposed by J. Kim, H. P. Nilles and M. Peloso (thus termed *KNP alignment*) in their paper [33]. In the minimal model the idea is to have $N = 2$ axions and a non-diagonal charge matrix. This can then be tuned such that a certain linear combination of the original axions has enhanced field range. The general principle is pictured in figure (2.9). As discussed above we can always go to the Φ basis in which the enhancement of the field range is not manifest because of the square form of the fundamental domain. The enhancement of the field range is hidden in the $GL(2,\mathbb{R})$ basis change and the eigenvalues of Ξ will in general not be equal to those of K. Since the eigenvalues of Ξ and not K determine the field range of the canonically normalized fields, the geometric enhancement translates to the enhancement of the largest eigenvector of Ξ. To see this explicitly, consider the $N = 2$, $P = 2$ potential

$$V = \Lambda_1^4 \left[1 - \cos\left(q_{11}\theta^1 + q_{12}\theta^2\right)\right] + \Lambda_2^4 \left[1 - \cos\left(q_{21}\theta^1 + q_{22}\theta^2\right)\right] . \tag{2.57}$$

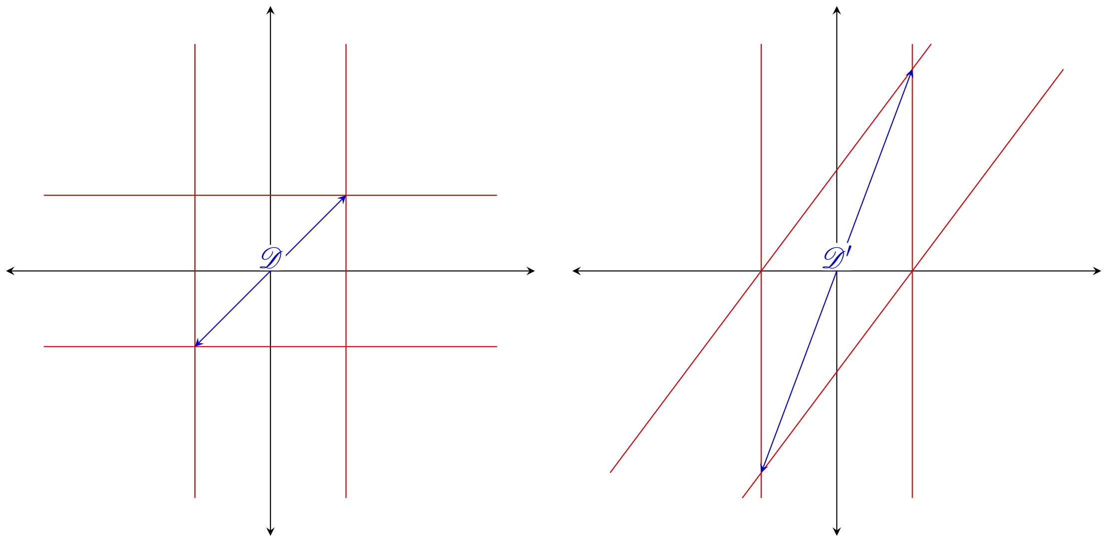

Figure 2.9: KNP alignment: The hyperplanes defining the periodic identifications are not at right angles to each other due to a non-diagonal instanton charge matrix. The fundamental domain is a distorted hypercube and has large diameter $\mathscr{D}' > \mathscr{D} = 2\sqrt{N}$ in certain directions.

Perfect alignment (a flat direction of V) occurs when the two hyperplanes pairs defined by the identifications

$$q_{11}\theta^1 + q_{12}\theta^2 = \pm\pi\,, \tag{2.58}$$
$$q_{21}\theta^1 + q_{22}\theta^2 = \pm\pi\,, \tag{2.59}$$

align in the geometric sense, i.e. when the charge matrix

$$\mathscr{Q} = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}\,, \tag{2.60}$$

degenerates in rank, or equivalently has vanishing determinant. Thus the perfect alignment condition can be written as

$$\det\mathscr{Q} = 0 \quad \Leftrightarrow \quad \frac{q_{11}}{q_{12}} = \frac{q_{21}}{q_{22}}\,, \tag{2.61}$$

for non-zero q_{ij}. In the θ basis this relates to $f \to \infty$ for one of the decay constants. Note that the flat direction in the potential discussed above arises in the θ basis. There will be also a flat direction in terms of the canonically normalized fields which might differ from the one in the θ basis, depending on the details of the Kähler metric K_{ij}. We will come back to this later. While in the simple case of $N = 2$ we can easily tune the charge matrix to give an exactly flat direction, corresponding to infinite diameter of the fundamental domain. In the large N case the alignment occurs generically [32]. To see this one can use random matrix theory, as discussed in appendix D. In [32] it was shown that if an $N \times N$ unit charge matrix $\mathscr{Q} = \mathrm{id}$ is perturbed by at least $2N$ random entries of variance $\sigma^2_{\delta\mathscr{Q}} \gtrsim 2/N$, the matrix $\mathscr{Q}^T\mathscr{Q}$ approaches a universal

Wishart limit. Let us for now restrict to the simple case $K = f^2\mathrm{id}$ and $P = N$. In the Φ basis, we have the Lagrangian (2.54) with Ξ now explicitly given as $\Xi = f^2\left(\mathscr{Q}^{-1}\right)^T\mathscr{Q}^{-1}$. Let us estimate the size of the largest eigenvalue of Ξ. In the Wishart limit eigenvalue repulsion as discussed in the appendix enforces an upper bound on the smallest eigenvalue λ_1 of $\mathscr{Q}^T\mathscr{Q}$. This is given by $\lambda_1 \lesssim \sigma_{\mathscr{Q}}^2/N$. The variance is approximately $\sigma_{\mathscr{Q}} \approx 2/\sqrt{N}$, so we find $\lambda_1 \lesssim 4/N^2$. This translates into a lower bound on the largest eigenvalue $\eta_N = 1/\lambda_1$ of $(\mathscr{Q}^{-1})^T\mathscr{Q}^{-1}$, $\eta_N \gtrsim N^2/4$. Finally, the largest eigenvalue $\xi_N = f^2\eta_N$ of the kinetic matrix Ξ is bounded from below as

$$\xi_N \gtrsim f^2N^2\,. \tag{2.62}$$

This means that the diameter of the fundamental domain is enhanced by a factor of N [34].

2.6.3 N-flation

The original N-flation proposal [35] considered N axions with $P = N$ instanton terms and a diagonal charge and kinetic matrix. The Lagrangian then takes the form

$$\mathscr{L} = \sum_{i=1}^{N}\left[\frac{1}{2}\left(\partial\theta^i\right)^2 - \Lambda_i^4\left[1-\cos\left(\frac{\theta^i}{f_i}\right)\right]\right]\,, \tag{2.63}$$

where f_i are the axion decay constants. For simplicity we consider $f_i \equiv f$, $\Lambda_i \equiv \Lambda$ and small displacements from the minimum at $\theta^i = 0$. The potential can be approximated to quartic order to give

$$\mathscr{L}_{\mathrm{eff}} = \sum_{i=1}^{N}\left[\frac{1}{2}\left(\partial\theta^i\right)^2 - \frac{m^2}{2}(\theta^i)^2 + \frac{\lambda}{4!}(\theta^i)^4\right]\,. \tag{2.64}$$

The mass is given in terms of the UV scale Λ and the decay constant f as the ratio $m = \Lambda^2/f$, while the dimensionless quartic coupling is $\lambda = (\Lambda/f)^4$. If we consider equidistant displacements $\theta^i = \alpha M_p$ while staying in the regime $\alpha^2 \ll f^2/M_p^2$ for consistency of the quadratic approximation, the radial mode $\rho^2 \equiv \sum(\theta^i)^2$ can easily reach a super-Planckian displacement $\rho = \sqrt{N}\alpha M_p$ for large enough N. For large N, the effective Lagrangian for ρ reduces to the simple form

$$\mathscr{L}_\rho = \left[\frac{1}{2}(\partial\rho)^2 - \frac{m^2}{2}\rho^2 + \frac{\lambda}{4!N}\rho^4 + \cdots\right]\,, \tag{2.65}$$

since the angular modes are over-damped [35]. It is crucial that the quadratic approximation for ρ can be trusted over much larger scales than the one for each θ^i, since the quartic and higher order couplings scale with inverse powers of N. For large enough N one should be able to trust the quadratic approximation of the potential over super-Planckian distances and use the radial mode as a natural inflaton. Another virtue of this model is that assuming randomly distributed initial conditions for the θ^i, one expects that for large N the radial mode ρ generically takes a super-Planckian value since the volume of an N-sphere is concentrated near its surface. The diameter of the fundamental domain as discussed above as an estimate of the available inflationary field range evaluates to $\sqrt{N}f$, corresponding to a $\sqrt{N}$ enhancement relative to the single axion case, see figure (2.10).

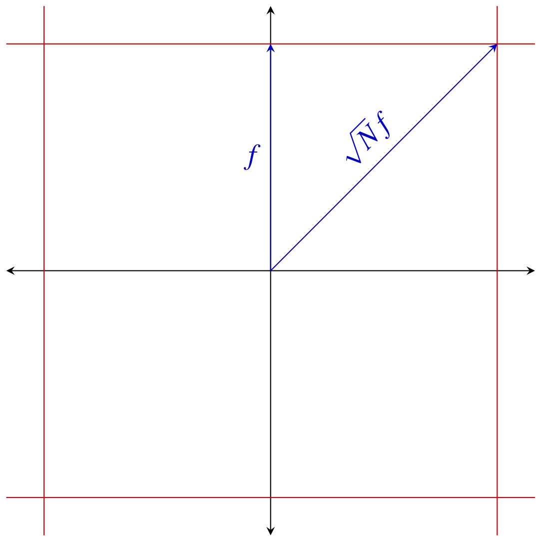

Figure 2.10: Diagonal N-flation: The radial mode has enhanced field range due to pointing along a diagonal in field space. The Kähler metric is rotationally invariant, so the axion decay constant is the same in any direction. Combining this gives a diameter of $\sqrt{N}f$ for the fundamental domain.

2.6.4 Kinetic Alignment

Kinetic alignment [36] takes the naive diagonal N-flation model to the next level by considering a generic kinetic matrix and again $P = N$ non-perturbative terms in the potential such that generically each axion is lifted. In the case of a charge matrix proportional to the identity, the Lagrangian reduces to

$$\mathscr{L} = \frac{1}{2}(\partial\theta)^T K\,\partial\theta - \sum_{i=1}^{N}\Lambda_i^4\left[1-\cos\left(\theta^i\right)\right] , \tag{2.66}$$

and one can go to a basis of canonically normalised fields by diagonalising the kinetic matrix as $K = R^T \mathrm{diag}\left(f_1^2, \cdots, f_N^2\right) R$, where f_i are the axion decay constants (conventionally ordered increasing in magnitude) and defining $\psi^i = f_i(R\theta)^i$ to give

$$\mathscr{L} = \frac{1}{2}(\partial\psi)^T\partial\psi - \sum_{i=1}^{N}\Lambda_i^4\left[1-\cos\left(\left(R^T\right)^i{}_j\frac{\psi^j}{f_j}\right)\right] . \tag{2.67}$$

If K were diagonal, R would be trivial and the fundamental domain in the ψ basis would be just an N-orthotope with largest side length $2\pi f_N$ and the diameter of the fundamental domain would be exactly this. The idea of kinetic alignment is that in general the non-trivial rotation will enhance the diameter. To see this, it is convenient to stay in the θ basis and consider the foliation of field space into ellipsoidal hypersurfaces (2.11)

$$\|\theta\|_K = \sqrt{\theta^T K\theta} = r\,. \tag{2.68}$$

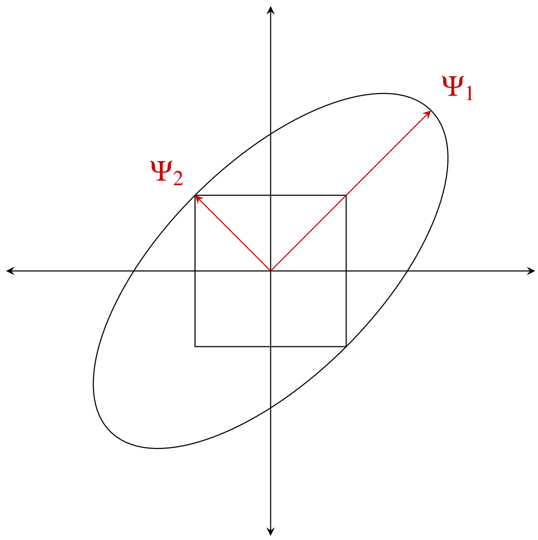

Figure 2.11: The ellipsoid of maximal invariant radius $r_{\max}$ still intersecting the hypercube of side length 2π for $N = 2$ in the favorable case where it is rotated such that the shortest principal axis is aligned with a diagonal.

It is clear that the diameter of the fundamental domain is then given by the radius of the largest ellipsoid still intersecting the N-cube of side length 2π. In the diagonal Kähler metric case the ellipsoid has its principal axes aligned with the coordinate axes, while in the most favourable case the rotation R rotates the ellipsoid such that the shortest principal axis (this corresponds to the largest eigenvalue of K) is aligned with a diagonal of the N-cube, as in figure (2.11). If Ψ_N is the corresponding eigenvector, rescaled to lie on the ellipsoid of maximal radius, i.e. on the boundary of the fundamental domain, then the maximal invariant diameter of the fundamental domain is given by

$$\Psi_N = (\pm\pi, \cdots, \pm\pi) \ , \tag{2.69}$$

$$\mathscr{D} = 2\sqrt{\Psi_N^T K \Psi_N} = 2 f_N \sqrt{\Psi_N^T \Psi_N} = 2\pi f_N \sqrt{N} \ . \tag{2.70}$$

In this way it is possible to get a $\sqrt{N}$ enhancement in the case of a non-trivial Kähler metric K. This alone does not help much, since we are interested in the generic case. The key insight made by [36] is that in the large N limit this "alignment accident" happens generically. Assuming a rotationally invariant ensemble of random choices for K_{ij} (see appendix D for a brief introduction to random matrix ensembles), the eigenvectors will be distributed uniformly on the unit sphere.[4] Since the number of diagonals of the N-cube grows exponentially with N, whereas the number of faces is obviously $2N$, the configuration where the ellipsoid has its shortest principal axis aligned with a diagonal becomes increasingly likely (details can be found in appendix D). So one actually expects the $\sqrt{N}$ enhancement (2.12) for a reasonable random choice of K.

[4] This phenomenon is known as eigenvector delocalisation.

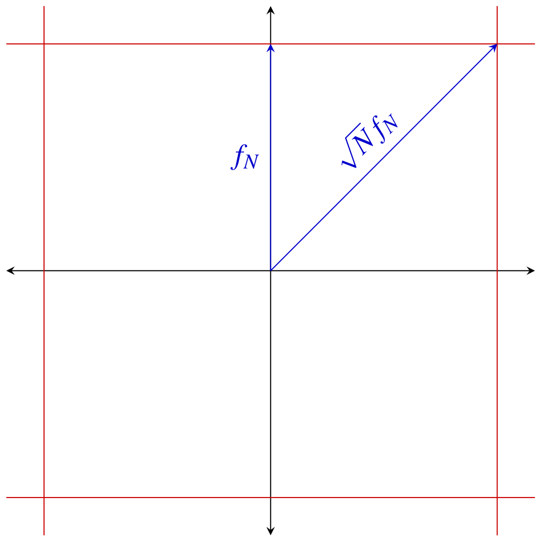

Figure 2.12: Kinetic Alignment: The eigenvector corresponding to the largest eigenvalue f_N points along the diagonal of the hypercube given by the periodic identifications. The result is a Pythagorean enhancement of the field range in that direction.

2.6.5 Combined Alignment

In [32] it was shown that the kinetic alignment can be combined with KNP style alignment and the typical enhancement of the fundamental domain diameter was calculated using random matrix theory. The authors found enhancements of $N^{3/2}$ in the case of $P = N$ and N^1 for the case of $P > N$.

① # Instantons = # Axions

Let us first look at the case where the number of instanton terms coincides with the number of axions. We have seen that for large N the eigenvalues of the matrix $\Xi = (\mathscr{Q}^{-1})^T K \mathscr{Q}^{-1}$ are enhanced by a factor of N via a KNP-like alignment mechanism if K is proportional to the identity matrix. This relied on the fact that the Kähler metric was proportional to the inverse of $\mathscr{Q}^T \mathscr{Q}$ and this matrix can be described in the large N limit by a Wishart matrix. Now in the case where we have a non-trivial Kähler metric K, we have to first see if Ξ still obeys eigenvector delocalisation and then estimate the eigenvalues of Ξ. In [32] it was found that, independent of K, if $\mathscr{Q}^T \mathscr{Q}$ is drawn from a rotationally invariant ensemble then the same holds for Ξ. So we get eigenvector delocalisation and thus kinetic alignment for Ξ. Furthermore, [32] made the guess that in the case where K is drawn from a Wishart ensemble, then the matrix Ξ will be approximately a rescaled inverse Wishart matrix

$$\Xi = \left(R\mathscr{Q}^{-1}\right)^T \mathrm{diag}(f_1^2, \ldots, f_N^2) R\mathscr{Q}^{-1} \sim \sigma_{f_i^2} \left(\mathscr{Q}^T \mathscr{Q}\right)^{-1} , \tag{2.71}$$

where $\sigma_{f_i^2}$ is the r.m.s. of the eigenvalues of K. For a Wishart matrix K, this is $\sigma_{f_i^2} \approx f_N^2/4$. This, combined with the estimate of the eigenvalues of $\mathscr{Q}^T \mathscr{Q}$, leads us to conclude that the total enhancement that can be achieved in this case is $\sqrt{N} \cdot N = N^{3/2}$ from a combination of kinetic alignment and KNP alignment. For K not a Wishart matrix, the results change. In fact, [32] find only a factor N enhancement for a particular choice of heavy tailed metric. After a careful evaluation one finds the following upper bounds on the fundamental domain diameter

$$\boxed{\mathscr{D} \lesssim \begin{cases} f_N N^{3/2} & \text{Wishart}\,, \\ f_N N & \text{Heavy-tailed}\,. \end{cases} \tag{2.72}}$$

2 # Instantons > # Axions

For the case where the instanton generated potential terms outnumber the axions, the $\sqrt{N}$ gain from kinetic alignment does not work any more. Intuitively, even after diagonalising N terms in the potential, the fundamental domain will be a hypercube with its diagonals sawed off, asymptoting to a sphere for $P/N \to \infty$. So generically there is no Pythagorean enhancement. The result is

$$\boxed{\mathscr{D} \lesssim \begin{cases} f_N N & \text{Wishart}\,, \\ f_N \sqrt{N} & \text{Heavy-tailed}\,. \end{cases} \tag{2.73}}$$

The intuitive argument above demands a more rigorous derivation [32]. We would like to compute the diameter of the fundamental domain in the Φ basis. As it was the case for $P = N$, the diameter will be twice the invariant distance to the surface of a maximal hyperboloid still intersecting the fundamental domain. It is useful to introduce an operator that takes a given vector and rescales it to lie on the boundary of the fundamental domain [32]

$$\varpi_{\mathscr{Q}}(w) \equiv \frac{\pi}{\max\limits_{i=1,\dots,P-N}\left(|(\mathscr{Q}Q^{-1}w)^i|\right)} w\,. \tag{2.74}$$

For a given vector w, the diameter of the fundamental domain in its direction is clearly given by

$$\mathscr{D}_w = 2\|\varpi_{\mathscr{Q}}(w)\|_{\Xi} = 2\sqrt{\varpi_{\mathscr{Q}}(w)^T\,\Xi\,\varpi_{\mathscr{Q}}(w)}\,. \tag{2.75}$$

The actual diameter is then obtained by taking the maximum of all $\mathscr{D}_w$ and each single $\mathscr{D}_w$ is a lower bound for the true diameter. The expression (2.75) simplifies greatly for w a linear combination of eigenvectors Ψ^i_{Ξ} of Ξ, corresponding to eigenvalues ξ_i^2. In particular, taking $v = \sum_i \xi_i \Psi^i_{\Xi}$ with associated unit vector $\hat{v}$ we have

$$\mathscr{D}_{\hat{v}} = 2\|\varpi_{\mathscr{Q}}(\hat{v})\|\sqrt{\frac{\sum_{i=1}^N \xi_i^4}{\sum_{i=1}^N \xi_i^2}}\,, \tag{2.76}$$

where the norm is now the standard Euclidean one, i.e. the dependence on the metric Ξ has been absorbed into the sums over powers of the eigenvalues. This estimate for the diameter is particularly useful because at the end we want to relate the diameter to the kinetic matrix eigenvalues. The claim of [32], which the authors support through numerical simulation, that the entries of the vector $Q_R Q^{-1}\hat{v}$ are distributed as $\mathcal{N}(0,\sqrt{2})$ for fixed Q, independent of σ_Q, N and P. This means that the maximum of these entries has typical (median) size [1]

$$\max_{i=1,\dots,P-N}\left(|(Q_R Q^{-1}\hat{v})^i|\right) \approx 2\mathrm{erf}^{-1}(2^{-\frac{1}{P-N}}) \approx \sqrt{4\log(P-N)}, \tag{2.77}$$

that is we have

$$\varpi_{\mathcal{Q}}(\hat{v}) \approx \hat{v}\,, \tag{2.78}$$

up to logarithmic corrections in the difference $P-N$. This means that the diameter of the fundamental domain is not enhanced by kinetic alignment in this case. The only enhancement can come from relating the eigenvalues ξ_i^2 of Ξ to the eigenvalues f_i^2 of K and thus we only get the factor N enhancement from KNP alignment in the Wishart case. For the heavy-tailed metrics, the authors of [32] find an enhancement of $\sqrt{N}$. This is summarised in (2.73).

2.6.6 Axion Monodromy

In contrast to typical models of natural inflation which use the instanton generated periodic potential (2.52), the idea of *axion monodromy* [38] is to ensure that there is another source of a potential for the axion which dominates the periodic one and is used for inflation. In order to do this in a controlled way one has to include additional degrees of freedom which reinstate the shift symmetry that would be broken by any non-periodic potential. At the level of 4d effective field theory, one possibility is to introduce a coupling to a four-form field strength [39]

$$S = \int \left(\frac{M_p^2}{2} \star R - \frac{1}{2} d\phi \wedge \star d\phi - \frac{1}{2} F_4 \wedge \star F_4 + \mu\phi F_4 - \star V_p(\phi) \right)\,. \tag{2.79}$$

Here V_p is a possible non-perturbative potential respecting the periodicity of ϕ. The coupling to the four form does not break the shift symmetry of ϕ because its variation can be cancelled by appropriate shifts of F_4 up to a total derivative. The field F_4 is non-dynamical in four dimensions and can be integrated out. In principle there can be 3-dimensional membranes (domain walls) coupling to F_4 and special care has to be taken in including appropriate boundary terms into (2.79). Upon integrating out F_4 this results in the appearance of a Lagrange multiplier field q which is locally constant but changes across the domain walls. The resulting effective action is

$$S = \int \left(\tfrac{M_p^2}{2} \star R - \tfrac{1}{2} d\phi \wedge \star d\phi - \star \left[\tfrac{1}{2}(q+\mu\phi)^2 + V_p(\phi) \right] \right)\,. \tag{2.80}$$

Thus a monomial potential is generated which might in principle dominate the periodic non-perturbative one. Note that the shift symmetry is still unbroken, since we may undo shifts of ϕ by appropriate shifts of q. It is only spontaneously broken by the local expectation value

[1] The cumulative distribution function (CDF) of the maximum of $P-N$ independent random variables is just the product of the individual CDFs. The result then follows because the CDF of the normal distribution is essentially the error function and the median is easily calculated by setting its $(P-N)^{\text{th}}$ power equal to $1/2$. For the asymptotic expansion of the inverse error function see e.g. [37]. This is valid for $P-N \gg 1$.

of q. We can view the potential in (2.80) as a multi-branched quadratic potential with branches indexed by an integer q and small wiggles on top. The aim of monodromy inflation is to suppress the wiggles and stay in a single branch of the potential over the inflationary period. Microscopic realisations in string theory typically involve wrapped, spacetime filling branes and the integer q or an analog thereof then discribes a brane charge. Different monomial potentials have been obtained from string theory. For an incomplete list of works on this topic see [38–46].

2.6.7 Consistency of Many Axions in Effective Field Theory

While the N-flation potential is protected from radiative corrections by the N shift symmetries of the individual axions in the same way as in single field natural inflation [31], there is a problem with having that many scalars in an effective field theory coupled to gravity. As pointed out in the original N-flation paper [35], the squared Planck mass is quadratically renormalised in the presence of N light scalars as

$$\delta M_p^2 \simeq \pm \frac{N}{16\pi^2} \Lambda_{\mathrm{UV}}^2 \,. \tag{2.81}$$

Since the inflationary slow roll parameters are proportional to M_p^2, one might be worried that any gain from having many fields is killed by this effect. The severeness of this renormalisation depends on the actual embedding into a UV-complete theory, as we will see in the next chapter.

3 The Difficulties of Large Field Inflation in String Theory

In this chapter we will discuss how the assorted models of natural inflation discussed in chapter 2 can be implemented in string theory. First we briefly review how four dimensional low energy effective field theories arise from the fundamental ten dimensional superstring theories. We will see that having axions in the low energy effective field theory is not a problem but producing the required super-Planckian axion decay constants proves to be an obstacle that seems to be impossible to overcome. We will not be concerned with constructing or reviewing actual inflation models but rather with the possibility of obtaining large axion fundamental domains as these are required for (non-monodromic) axion inflation. A very good reference for inflation in string theory is the review book by D. Baumann and L. McAllister [24]. For general string theory references, see [47, 48][49][50].

3.1 Low Energy Effective Supergravity

The basic strategy for deriving a 4d effective action is as follows. The effective field theory for the massless modes of the superstring is ten dimensional *supergravity* (SUGRA), which can be derived by imposing conformal symmetry of the string sigma model. This is then compactified on a product space $M_4 \times Y_6$ where Y_6 is compact, allowing for consistent VEVs of the fields such that the 4d Poincaré symmetry is not broken, and these are expanded around their VEVs into harmonics of Y_6. In this section we will desribe the 10d effective SUGRA actions of the different superstring theories.

There are five distinct supersymmetric string theories in ten dimensions. They are connected by a web of *string dualities* as depicted in (3.1) and should themselves be thought of as perturbative descriptions of different aspects of a single overarching theory. At strong coupling the IIA and $E_8 \times E_8$ theories can be seen to effectively grow an additional dimension and are conjectured to be the weak coupling limits of an eleven-dimensional *M-theory*, which reduces in the low energy limit to the unique 11d supergravity. The two type II string theories correspond to two $\mathcal{N} = 2$ effective supergravity theories, type IIA and type IIB, differing in the chirality of their supersymmetries. The type IIA SUGRA action can be deduced either by a stringy computation as indicated above or by dimensional reduction of the 11d SUGRA.

Due to its simplicity, we start precisely with this 11d SUGRA. For classical solutions and in particular cosmology, one is interested only in the bosonic field content. The bosonic sector consists only of the graviton and a 3-form

$$S = \frac{1}{2\kappa_{11}} \int d^{11}X \sqrt{-G} \left(\frac{1}{2} |F_4|^2 \right) - \frac{1}{6} \int A_3 \wedge F_4 \wedge F_4 \, . \tag{3.1}$$

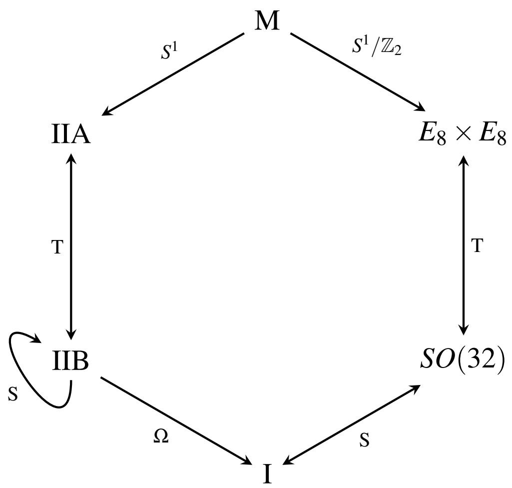

Figure 3.1: The different possible supersymmetric string theories in 10 dimensions are connected by a web of dualities. The IIA and heterotic $E_8 \times E_8$ theories can be thought of as the two possible compactifications of a unifying 11d theory whose low energy dynamics is described by the unique 11d SUGRA.

11D	10D
gravitino	2 gravitini + 2 dilatini
graviton	graviton, 1-form, dilaton
3-form	3-form + NS 2-form

Table 3.2: The field content of type IIA SUGRA as arising from compactification of eleven dimensional SUGRA.

The 3-form couples to extended objects, electrically to *M2-branes* and magnetically to *M5-branes*. These are stable supersymmetric BPS states with tensions

$$T_{\mathrm{M2}} = 2\pi(2\pi l_{11})^{-3} \qquad T_{\mathrm{M5}} = 2\pi(2\pi l_{11})^{-6}\,, \tag{3.2}$$

where the 11D Planck length is defined by $2\kappa_{11} = (2\pi l_{11})^9/2\pi$. One can then compactify this on an S^1 of radius R_{11} and this leads to the 10d type IIA supergravity. The correspondence between the 11d and 10d fields is schematically summarized in (3.1). The 11D 3-form couples to an extended object, the M2-brane, which when wrapped on the S^1 leads to a stringy object in ten dimensions. This is precisely the fundamental string of the type IIA string theory. The volume modulus of the S^1 is then identified with the string coupling

$$R_{11} = g_s \ell_s \qquad l_{11} = g_s^{1/3} \ell_s\,. \tag{3.3}$$

We will state the type IIA and IIB actions without derivation. The different fields arise as the lowest excitations of the string worldsheet. Depending on the boundary conditions imposed on the fields on the string worldsheet there are two different sectors, namely the Ramond (R) and Neveu-Schwarz (NS) sectors. For closed string theories such as the type II and heterotic ones one furthermore has to take the tensor product of left- and rightmoving excitations.

The NS-NS sector is the same in the IIA and IIB theories. The differences appear in the R-NS, NS-R and R-R sector. Since the R-NS/NS-R sectors contain only fermions we omit them. The R-R sector contributes several p-form gauge fields, namely the odd ones in IIA and the even ones in IIB. The p-forms couple to extended objects in the theory, namely electrically to $(p-1)$-branes and magnetically to $(D-p-3)$-branes. Only these branes which are charged under the form fields can be stable. The universal NS-NS action contains the graviton, Kalb-Ramond 2-form B_2 and the dilaton Φ, which determines the string coupling via $g_s = \exp(\Phi)$. It is given by

$$S_{\rm NS} = \frac{1}{2\kappa_{10}^2} \int d^{10}x \sqrt{-G} e^{-2\Phi} \left(R + 4(\partial\Phi)^2 - \frac{1}{2}|H_3|^2 \right) , \tag{3.4}$$

where $H_3 = dB_2$ and $2\kappa_{10}^2 = (2\pi)^7 \ell_s^8$.

Type IIA Supergravity
The IIA SUGRA has the action

$$S_{\rm IIA} = S_{\rm NS} + S_R^{\rm IIA} + S_{CS}^{\rm IIA} , \tag{3.5}$$

where the Ramond and Chern-Simons terms are given by

$$\begin{aligned} S_R^{\rm IIA} &= -\frac{1}{4\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left(|F_2|^2 + |\tilde{F}_4|^2 \right) , \\ S_{CS}^{\rm IIA} &= -\frac{1}{4\kappa_{10}^2} \int B_2 \wedge F_4 \wedge F_4 , \end{aligned} \tag{3.6}$$

and $F_{p+1} = dC_p$ are the p-form field strengths, $\tilde{F}_4 = F_4 + C_1 \wedge H_3$.

Type IIB Supergravity
The IIB SUGRA has the action

$$S_{\rm IIB} = S_{\rm NS} + S_R^{\rm IIB} + S_{CS}^{\rm IIB} , \tag{3.7}$$

with Ramond and Chern-Simons terms given by

$$\begin{aligned} S_R^{\rm IIB} &= -\frac{1}{4\kappa_{10}^2} \int d^{10}x \sqrt{-G} \left(|F_1|^2 + |\tilde{F}_3|^2 + |\tilde{F}_5|^2 \right) , \\ S_{CS}^{\rm IIB} &= -\frac{1}{4\kappa_{10}^2} \int C_4 \wedge H_3 \wedge F_3 , \end{aligned} \tag{3.8}$$

with $\tilde{F}_3 = F_3 - C_0 \wedge H_3$, $\tilde{F}_5 = F_5 - \frac{1}{2}C_2 \wedge H_3 + \frac{1}{2}B_2 \wedge F_3$ and the supplementary condition that $\star\tilde{F}_5 = \tilde{F}_5$. The latter self-duality condition is needed for SUSY and must be imposed in addition to the equations of motion following from the IIB action.

The 10d actions above appear in the so-called *string frame*, where the dilaton appears in front of the Einstein-Hilbert term. One can remove this by a Weyl rescaling to *Einstein frame*, $G_E = \exp(-\Phi/2)G$, and for example in the IIB case one arrives at

$$S_{\mathrm{IIB}} = \frac{1}{2\kappa_{10}^2}\int d^{10}x\sqrt{-G_E}\left[R_E - \frac{|\partial\tau|^2}{2\,\mathrm{Im}\,(\tau)^2} - \frac{|G_3|^2}{2\,\mathrm{Im}\,(\tau)} - \frac{|\tilde{F}_5|}{4}\right] - \frac{i}{8\kappa_{10}^2}\int\frac{C_4\wedge G_3\wedge \overline{G}_3}{\mathrm{Im}\,(\tau)}\,, \quad (3.9)$$

where the axio-dilaton is defined as $\tau = C_0 + i\exp(-\Phi)$ and the complexified 3-form is $G_3 = F_3 - \tau H_3$. This form of the action makes an $SL(2,\mathbb{R})$-symmetry

$$\tau \mapsto \frac{a\tau+b}{c\tau+d}\,, \qquad \begin{pmatrix} a & b \\ c & d\end{pmatrix} \in SL(2,\mathbb{R}) \quad (3.10)$$

manifest. In the full string theory this extends only to an $SL(2,\mathbb{Z})$-symmetry, the S-selfduality of the type IIB string.

Finally there are also the 10D $\mathcal{N}=1$ supergravity theories. These are the type I and heterotic supergravities. The bosonic sector of the heterotic supergravity is

Heterotic Supergravity

$$S = \frac{1}{2\kappa_{10}^2}\int d^{10}X\sqrt{-G}e^{-2\Phi}\left[R + 4(\partial\Phi)^2 - \frac{1}{2}|\tilde{H}_3|^2 - \frac{\ell_s^2}{120}\mathrm{tr}_a\left(|F_2|^2\right)\right]\,, \quad (3.11)$$

where $F = dA + A\wedge A$ is the Yang-Mills field strength of the heterotic string gauge field and here $\tilde{F}_3 = dC_2 + \ell_s^2\omega_3/4$ and $\omega_3 = \omega_{\mathrm{L}} - \omega_{\mathrm{YM}}$,

$$\begin{aligned} \omega_{\mathrm{L}} &= \mathrm{tr}\left(\omega\wedge d\omega + \tfrac{2}{3}\omega\wedge\omega\wedge\omega\right)\,, \\ \omega_{\mathrm{YM}} &= \mathrm{tr}\left(A\wedge dA - \tfrac{2i}{3}A\wedge A\wedge A\right)\,. \end{aligned} \quad (3.12)$$

The type I SUGRA is related to this by a field redefinition under which $\Phi\to-\Phi$. This exchanges strong and weak coupling and is the basis for the S-duality between the corresponding string theories.

3.2 Brane Worldvolume Actions

The branes to which the p-forms of the different SUGRA theories couple are dynamical objects in themselves. In the simplest case of a single brane the low energy brane excitations are described by the *Dirac-Born-Infeld action*

$$S_{\mathrm{DBI}} = -T_p\int_{\Sigma_{p+1}} d^{p+1}\xi\,\sqrt{-\det i^*\left(G+\mathscr{F}\right)}\,, \quad (3.13)$$

where i is the embedding of the brane into spacetime and $2\pi\alpha'\mathscr{F} = 2\pi\alpha' F + B$ is the gauge invariant field strength. The brane couples to various p-forms via the brane Chern-Simons action

$$S_{\mathrm{CS}} = \pm T_p \int_{\Sigma p+1} i^* \left(\sum_k C_k \right) \mathrm{ch}\left({\ell_s}^2 \mathscr{F} \right) \sqrt{\frac{\hat{A}(T\Sigma)}{\hat{A}(N\Sigma)}} \Bigg|_{p+1 \text{ part}} . \tag{3.14}$$

Here the sum is understood to be only over those p-forms that appear in the given theory, $\hat{A}$ denotes the A-roof genus of the tangent and normal bundles respectively and $\mathrm{ch}(\dots)$ is the Chern character.

3.3 Compactification

To get a realistic four dimensional theory we have to compactify six of the ten dimensions of the superstring. Let us denote the coordinates of the full ten-dimensional spacetime X_{10} by X^M, those of the four dimensional M_4 one by x^μ and finally those of the compactification space Y_6 as y^m. The most naive ansatz is a product manifold $X_{10} = M_4 \times Y_6$ equipped with the metric

$$G_{10}(X) = g_4(x) + g_6(y) . \tag{3.15}$$

A central question in determining the low energy effective field theory in four dimensions is its massless field content. To determine this we investigate the equation of motion for a ten-dimensional field and expand it in eigenfunctions of the corresponding differential operator. For example for a free 10d p-form C_p the equation of motion in Lorenz gauge, $d^\dagger C_p = 0$, is

$$0 = \Delta_{10} C_p = \left(dd^\dagger + d^\dagger d\right)\big|_{10\mathrm{d}} C_p = \Delta_4 C_p + \Delta_6 C_p . \tag{3.16}$$

Now if we expand $C_p = \sum_{q,i} A^i_{p-q} \wedge \omega^i_q$ in terms of eigenforms of the $6d$ Laplace operator of g_6, where ω_q has indices and functional dependence only along Y_6 and likewise A_{p-q} only along M_4, this will provide an effective mass term to the resulting four dimensional fields A_{p-q}

$$0 = \Delta_4 A^i_{p-q} + m_i^2 A^i_{p-q} , \quad \Delta_6 \omega^i_q = m_i^2 \omega^i_q . \tag{3.17}$$

Hence we see that massless form fields only arise from *harmonic forms* with $\Delta_6 \omega = 0$ on the compactification space and provided $p \geq q$ we get one such massless field for each harmonic form (these can be classified topologically, see appendix A). A similar reasoning applies to fermions. Since the Dirac operator also splits into a sum of a 4d and 6d part the classification of massless fermions is by zero modes of the 6d Dirac operator. In the above discussion it was critical that the metric split into strictly 4d and 6d parts with no cross-coupling (3.15). While this might seem like an innocent assumption the most general form compatible with 4d Poincare invariance is actually a *warped product* metric

$$G_{10}(X) = e^{2A(y)} g_4(x) + g_6(y) . \tag{3.18}$$

The *warp factor* $A(y)$ can be sourced by background gauge fields (*fluxes*) and localised sources in Y_6. For warped compactifications the effective mass term (3.16) has another contribution from the gradient of the warp factor. We will not go into any detail here but simply state that in regions where A is approximately constant the massless fields are obviously still given in terms of harmonic forms and refer the reader to [51, 52].

Finally the rest of the massless fields comes from expanding the $6d$ metric around its background VEV. The metric fluctuations can be decomposed into scalar, vector and tensor fluctuations. Of particular importance are the scalar deformation modes of the compact metric, because they can take on vacuum expectation values without breaking Lorentz symmetry. These are termed *metric moduli*.

If we start with one of the type II supergravities in ten dimensions, the requirement that the four dimensional effective field theory to be $\mathcal{N}=2$ supersymmetric demands the internal space Y_6 to be of *Calabi-Yau type* (see appendix A for a discussion of Calabi-Yau manifolds). The same holds for heterotic $\mathcal{N}=1$ compactifications. In order to obtain $\mathcal{N}=1$ supersymmetry from type II string theory one can perform *orientifolding* to a Calabi-Yau. This mods out a combination of string worldsheet parity and a geometric involution of the Calabi-Yau. In effect certain fields are projected out of the theory. For a review on 4d string compactifications with branes, fluxes and orientifolds, see [53].

3.4 De Sitter Vacua in String Theory

Before we get to discuss the implementation of natural inflation models in string theory let us discuss some generalities of inflation in string theory. As inflation is described by an approximate de Sitter evolution it is crucial to understand how de Sitter vacua may arise. De Sitter vacua are not only needed to describe inflation but also for describing the late time asymptotics of our universe. De Sitter space necessarily breaks supersymmetry. This can be seen for example because there is no de Sitter superalgebra, in contrast to the case of anti de Sitter space [54]. This already presents a technical problem as SUSY tends to simplify things significantly and in particular helps to control radiative corrections. To get a realistic model all moduli have to be stabilised meaning that there has to be a potential for the scalars that describe the extra-dimensional geometry and the value of the potential at this minimum has to be positive for de Sitter. Stabilising all moduli in a string compactification is a non-trivial task and usually one first looks for an AdS vacuum for which the stabilisation is under good control[1] and then *lifts* this solution *up* to de Sitter by adding addiational SUSY breaking ingredients such as anti-branes. These do however backreact on the solution and this leads generically to instabilities.[2] In the following we will not touch issues of moduli stabilisation and SUSY breaking as we are not interested in fully explicit inflation models here but rather the general issue of super-Planckian field displacements in quantum gravity.

[1] The two main schemes for moduli stabilisation in IIB are known as KKLT [55] and the Large Volume Scenario [56]. In IIA all geometric moduli can be fixed at the classical level [57].

[2] For recent work on interpreting dS as a resonance in the transition amplitude between supersymmetric flat vacua, see [58, 59]

3.5 Axions from String Theory

We will now see how axions arise in four dimensional string compactifications (see [60] for a review). We will focus on the type II theories. The ten dimensional effective SUGRA actions (3.5,3.7) each contain p-form gauge fields and in particular the B-field. This can be expanded into a purely four dimensional part and a product of four dimensional scalars times harmonic two-forms ω_i of the compactification space Y

$$B_2(X) = b_{\mu\nu}(x)dx^\mu \wedge dx^\nu + \sum_i b^i(x)\omega_i + \dots \,. \tag{3.19}$$

The *model independent axion* arises from $b_{\mu\nu}$. Upon dualisation this leads to a scalar field and it can be proven that this also has a shift symmetry, which is broken only by world-sheet instantons [24, 48]. Second, there are the *model dependent axions* b^i arising depending on the actual compactification geometry. Other form fields lead to axions in a completely analogous way.

The kinetic matrix of the axions determines their decay constants. In the string theory setting it is given in terms of geometric data of the compactification. Omitting for now the model independent axion and pre-factors we have for the B-field

$$\begin{aligned} S \supset \int_{M_4}\int_{Y_6} dB_2 \wedge \star_{10} dB_2 &= \int_{M_4}\int_{Y_6} (db^i \wedge \omega_i) \wedge (\star_4 db^j \wedge \star_6 \omega_j)\,, \\ = \int_{M_4} \left(\int_{Y_6} \omega_i \wedge \star_6 \omega_j \right) db^i \wedge \star_4 db^j &\equiv \int_{M_4} K_{ij} db^i \wedge \star_4 db^j \,. \end{aligned} \tag{3.20}$$

3.6 Single Field Natural Inflation

As we have seen, single field natural inflation demands a super-Planckian value of the axion decay constant f. The status of such large decay constants in string theory has been assessed in [19]. It was found that it is impossible to have $f > M_p$ in a controlled way. For example in the heterotic theory we have the same model dependent and independent axions as discussed above for the type IIB theory but also Wilson lines from integrating the heterotic gauge field over one-cycles in the compactification. The most simple setting analysed in [19] was the toroidal compactification on a T^6 with all radii set equal to R. In this setting the decay constants take the form (3.6). At first sight it looks like it is possible to get super-Planckian decay constants by simply going to sub-string length radii. But in fact, the heterotic string is T-dual to itself with the two possible gauge groups $E_8 \times E_8$ and $SO(32)$ being interchanged. In the T-dual description the sub-stringy radius gets mapped to a super-stringy one $\tilde{R} = {\ell_s}^2/R$. So in the T-dual description it is clear that the effective axion decay constant is in fact not super-Planckian. In the original description this can be traced to the winding modes becoming lighter than the momentum modes. The periodicity of the effective potential generated by the winding states is not $1/R$ but rather $R/{\ell_s}^2$. This illustrates the big picture. The idea is that when the decay constants seemingly become super-Planckian in a string construction certain degrees of freedom such as winding modes or wrapped branes become light and induce an effectively smaller periodicity in the axion effective field theory. This indicates that super-Planckian decay constants are a sign of loss of perturbative control.

Axion Species	f/M_p
Model Depended B-axion	$\left(\frac{\ell_s}{R}\right)^2$
Model Independed B-axion	1
Wilson Line A-Axion	$\left(\frac{\ell_s}{R}\right)^1$

Table 3.3: Axion decay constants for a toroidal heterotic string compactification.

3.7 Aligned Natural Inflation

The question of whether or not it is possible for two axions of both sub-Planckian decay constants to effectively produce a super-Planckian field range is more subtle in string theory. In chapter 4 we will find that general quantum gravity reasoning lends support to the belief that this is indeed *not* possible. We will here briefly mention the work of [61]. There it was attempted to construct an alignment scenario in type IIA string theory on a Calabi-Yau orientifold.[3] The IIA setting was chosen because here it is possible to stabilise all the geometric moduli (saxions) at the classical level. Only one linear combination of the axionic superpartners of the saxions is stabilised. The orthogonal complement is left massless at tree level. The linear combination being fixed can be adjusted by tuning flux numbers through various cycles. By integrating out the massive axion an alignment is induced.

The effective action for type IIA orientifolds has been derived in [62]. The moduli stabilisation by fluxes was achieved in [57]. The moduli stabilisation is briefly summarized in [61] and we use the same notational conventions as there. The relevant scalar fields are the dilaton s, the Kähler- and complex structure moduli t^i,v^λ and their associated axionic partners. In the large complex structure limit, the Kähler potential is

$$K = -\log(8\mathcal{V}) - \log(S+\bar{S}) - 2\log(\mathcal{V}')\,, \tag{3.21}$$

where $\mathcal{V}$ is the physical volume of the Calabi-Yau in string units and $\mathcal{V}'$ is the "complex structure volume", which is interpreted as the physical volume of the mirror Calabi-Yau

$$\mathcal{V} = \frac{1}{6}k_{ijk}t^i t^j t^k\,, \qquad \mathcal{V}' = \frac{1}{6}d_{\lambda\rho\sigma}v^\lambda v^\rho v^\sigma\,. \tag{3.22}$$

These can also be expressed in terms of the dual (mirrors of) 4-cycle volumes

$$\begin{aligned} \tau_i &= \partial_{t^i}\mathcal{V}\,, & u_\lambda &= \partial_{v^\lambda}\mathcal{V}'\,, \\ \mathcal{V} &= \mathcal{V}(t(\tau))\,, & \mathcal{V}' &= \mathcal{V}'(v(u))\,. \end{aligned} \tag{3.23}$$

As already anticipated above, by $\mathcal{N}=1$ SUSY, the moduli form chiral multiplets and are each paired with an axion coming from a form field of the appropriate rank. Explicitly

$$S = s + i\sigma\,, \qquad T_i = b_i + it_i\,, \qquad U_\lambda = u_\lambda + iv_\lambda\,. \tag{3.24}$$

[3] This is a $\mathcal{N}=1$ theory, see appendix B

In presence of fluxes $f_0, \tilde{f}_0, f_i, \tilde{f}^i, h_0, q^\lambda$, a superpotential is generated

$$W = f_0 k_{ijk} T^i T^j T^k + \frac{1}{2}\tilde{f}^i k_{ijk} T^j T^k - f_i T^i + \tilde{f}_0 - i h_0 S - i q^\lambda U_\lambda \ . \tag{3.25}$$

The SUSY conditions (B.5) for the various fields give

$$h_0 \sigma + q^\lambda \nu_\lambda = -\operatorname{Re}\left(W^T\right) \ , \tag{3.26}$$

$$\frac{q^\lambda}{K_\lambda} = 2 h_0 s = -\operatorname{Im}(W) \ , \tag{3.27}$$

$$\partial_{T^i} W^T = i(\partial_{T^i} K) \operatorname{Im}\left(W^T\right) \ , \tag{3.28}$$

where the T superscript denotes the part of the superpotential independent of S, U. The dilaton is fixed by (3.27), while (3.28) fixes the Kähler moduli. Equation (3.26) fixes a single linear combination of the axions. As a corollary of (3.27),

$$\frac{q^\gamma}{q^\delta} = \frac{K_\gamma}{K_\delta} \ , \tag{3.29}$$

fixes the complex structure moduli in terms of flux ratios.

The most simple model of alignment studied in [61] uses a moduli space spanned by two axions. Since there is always the axionic partner of the dilaton, σ, the simplest case is the one where there is one complex structure modulus

$$\mathcal{V}' = u^{3/2} \ . \tag{3.30}$$

The moduli stabilisation equations lead to

$$\frac{s}{u} = \frac{q}{3 h_0} \ , \tag{3.31}$$

where q, h_0 are the only 3-form fluxes present. In [61] it was then shown that the axion linear combination ψ orthogonal to the fixed one $h_0 \sigma + q \nu$ obtains an effective potential

$$V_{\text{eff}} = V_0 + A' e^{-s}\left(1 - \cos\frac{q\psi}{N}\right) + B' e^{-u}\left(1 - \cos\frac{h_0 \psi}{N}\right) \ , \tag{3.32}$$

with $N = \sqrt{h_0^2 K_{U\bar{U}} + q^2 K_{S\bar{S}}}$.[4] One can then see that the effective decay constant of the u instanton term can be enhanced with respect to f_σ by tuning for large flux q

$$f_\psi^{u_1} = \frac{N}{h_0} = \frac{q f_\sigma}{h_0}\sqrt{1 + \left(\frac{h_0 f_\nu}{q f_\sigma}\right)^2} \ . \tag{3.33}$$

Using the moduli stabilisation equation (3.31), one can show that

$$\frac{f_\nu}{f_\sigma} = \frac{q}{\sqrt{3} h_0} \ , \tag{3.34}$$

[4] The kinetic matrix determines the decay constants as $K_{U\bar{U}} \sim f_\nu^2 \sim 1/u^2$ and $K_{S\bar{S}} \sim f_\sigma^2 \sim 1/s^2$.

which leads in the critical case of large q to

$$f^u_\psi = 2f_v \ . \tag{3.35}$$

This means that even though there is a relative enhancement of decay constants this is only because in this limit one of the original decay constants gets smaller. The effective decay constant is not enhanced with respect to the other fundamental one and thus faces the same sub-Planckian constraint. More complicated examples were studied in [61] with the conclusion that in the cases where an actual super-Planckian enhancement of a decay constant was possible, there were other and in fact dominant terms in the effective potential with shorter periodicity which spoil the super-Planckian field range and lead to a sub-Planckian fundamental domain. For a different attempt at alignment in string theory see [63].

3.8 N-flation

Moving on to N-flation, there are also generic problems with having a very large number of fields in string theory. First of all there is the stringy incarnation of (2.81). In the string theory interpretation the scalar fields arise as geometric moduli and the internal manifold of the compactification adjusts itself to packing a large amount of geometry inside a small space. Here one can compute explicitly the corrections to the 10d Einstein-Hilbert action in terms of the string scale α'. The first order α' renormalisation of M_p can be shown to take the form [35]

$$\delta M_p^2 = M_p^2 \chi(Y) \left(\frac{\alpha'}{2\pi}\right)^3 \frac{\zeta(3)}{\mathcal{V}_Y} \ , \tag{3.36}$$

where $\chi(Y)$ and $\mathcal{V}_Y$ are the Euler characteristic and volume of the internal manifold respectively [35]. We see that this can be much better behaved than the naive field theoretic estimate. Since the Euler characteristic of a Calabi-Yau is the *alternating* sum of Betti numbers, there can be a cancellation between the complex structure and Kähler moduli. This means that the first order corrections to the Planck mass pose no direct obstruction for N-flation as long as one is able to find Calabi-Yau manifolds of large Hodge numbers but small Euler characteristic.

There is another objection to the N-flation paradigm. We have seen that the diameter of the axion fundamental domain and thus the field range could be enhanced with respect to the largest eigenvalue of the kinetic matrix, but what happens to the eigenvalues/decay constants themselves? For each single axion these are sub-Planckian and the importance of this was highlighted in the last section. If string theory censors super-Planckian axion field ranges as suggested by the Weak Gravity Conjecture, which we will discuss in chapter 4, we might guess that the decay constants could scale as $1/\sqrt{N}$, spoiling any gain in field range. But why should this be true? The squared axion decay constants are eigenvalues of the Kähler metric, which is the second derivative of a Kähler potential. In string theory this is, focusing on the Kähler moduli sector in the large volume limit for now, just the logarithm of the Calabi-Yau volume. This is given in terms of the triple intersection numbers of 2-cycles and their volumes, so it depends crucially on the internal geometry of the compactification space. The geometry adjusts to the large number of fields and thus cycles by growing in total volume. The derivatives of $\log(\mathcal{V})$ scale with powers of $1/\mathscr{V}$ and since the volume scales with N also with negative powers of N. The next section is devoted to analysing this in more detail. One might object that the volume might not grow with N if the volumes of the individual cycles scale with inverse powers of N. But in this case one is driven to the limit in moduli space where the cycles are smaller than the string scale and stringy corrections such as multi-wrapping of instantons on these cycles become important, hence we lose perturbative control. For attempts at constructing N-flation models in string theory, see [34, 64, 65].

3.8.1 Inverse Scaling of the Decay Constants

In the following we will make the N-scaling of the decay constants more precise.

Properties of the Kähler Metric

Let us record some properties of the metric on the Kähler moduli space derived from the Kähler potential

$$K = -\log(8\mathscr{V}) , \qquad \mathscr{V} = \frac{1}{3!} k_{ijk} t^i t^j t^k . \tag{3.37}$$

The volume can also be expressed in terms of the dual 4-cycle volumes

$$\tau_i = \frac{\partial \mathscr{V}}{\partial t^i} , \qquad \mathscr{V} = \mathscr{V}(t(\tau)) . \tag{3.38}$$

It is a homogeneous function of degree 3 and 3/2 in the t^i and τ_i respectively. Let us denote derivatives of K and $\mathscr{V}$ with respect to t^i by lower and τ_i by upper indices. The homogeneity properties are then summarized by

$$\mathscr{V}_i t^i = \tau_i t^i = 3\mathscr{V} , \qquad \mathscr{V}^i \tau_i = \tfrac{3}{2}\mathscr{V} . \tag{3.39}$$

Homogeneous functions of degree δ are "eigenfunctions of the Legendre transform operator" with eigenvalue $\delta - 1$, so

$$\mathscr{L}[\mathscr{V}](\tau) = 2\mathscr{V}(\tau) , \tag{3.40}$$

and it follows that

$$\mathscr{V}^i = \tfrac{1}{2} t^i . \tag{3.41}$$

Homogeneity imposes further constraints on derivatives:

$$\mathscr{V}_{ij} t^j = 2\mathscr{V}_i , \qquad \mathscr{V}^{ij}\tau_j = \tfrac{1}{2}\mathscr{V}^i = \tfrac{1}{4} t^i , \qquad \mathscr{V}^{ij}\mathscr{V}_{jk} = \frac{1}{2}\frac{\partial t^i}{\partial \tau_j}\frac{\partial \tau_j}{\partial t^k} = \frac{1}{2}\delta^i_k . \tag{3.42}$$

The Kähler metric for the t^i is given by

$$K_{ij} = -\frac{\mathscr{V}_{ij}}{\mathscr{V}} + \frac{\tau_i \tau_j}{\mathscr{V}^2} , \tag{3.43}$$

while the one for the τ_i is given by

$$K^{ij} = -\frac{\mathscr{V}^{ij}}{\mathscr{V}} + \frac{1}{4}\frac{t^i t^j}{\mathscr{V}^2} . \tag{3.44}$$

One can easily check that up to a factor of the volume they are inverse to each other

$$K^{ij}K_{jk} = \frac{\delta^i_k}{2\mathscr{V}^2} \,. \tag{3.45}$$

N-scaling

The Kähler moduli t^i are the expansion coefficients of the Kähler form of the Calabi Yau in terms of a basis of harmonic $(1,1)$-forms ω_i

$$J = t^i \omega_i \,. \tag{3.46}$$

Allowed deformations must preserve the positivity of volumes of holomorphic submanifolds in various co-dimensions

$$\int_C J > 0\,, \qquad \int_D J \wedge J > 0\,, \qquad \int_{\mathrm{CY}} J \wedge J \wedge J > 0\,, \tag{3.47}$$

where C denotes curves, D divisors and CY the whole Calabi-Yau. The set of all such deformations is invariant under positive rescaling and thus forms a cone, the *Kähler cone*. Integrating over the Poincaré dual cycles to the ω_i, one obtains $t^i > 0$ and hence in a natural parametrisation of the Kähler cone,

$$t^i \in (0,\infty). \tag{3.48}$$

Positivity of $\mathscr{V}$ then implies semi-positivity of the triple intersection numbers [64]

$$k_{ijk} \geq 0\,, \qquad \forall i,j,k\,. \tag{3.49}$$

Now naively $\mathscr{V}$ contains three summations over the index range $i = 1,\ldots,N$

$$\mathscr{V} = \frac{1}{3!} k_{ijk} t^i t^j t^k \,, \tag{3.50}$$

thus $\mathscr{O}(N^3)$ terms, while

$$\tau_i = \mathscr{V}_i = \frac{1}{2} k_{ijk} t^j t^k \tag{3.51}$$

contains only two summations and $\mathscr{O}(N^2)$ terms. Similarly, $\mathscr{V}_{ij}$ should contain $\mathscr{O}(N)$ terms. These must all be positive, so there can be no cancellations and if we take all t^i of the same order of magnitude $t^i \sim \hat{t}$

$$\mathscr{V} \simeq N^3 \hat{t}^3\,, \qquad \mathscr{V}_i = \tau_i \simeq N^2 \hat{t}^2\,, \qquad \mathscr{V}_{ij} \simeq N\hat{t}\,. \tag{3.52}$$

It follows that the Kähler metric for the 2-cycle volumes scales with N as

$$K_{ij} \sim N^{-2}\,. \tag{3.53}$$

Now the fact (3.45) comes into play. Because the two metrics are inverse to each other, up to a factor of $1/\mathscr{V}^2$ which scales like N^{-6}, we see that the metric for the 4-cycle volumes has to scale like

$$K^{ij} \sim N^2 \cdot N^{-6} = N^{-4}\,, \tag{3.54}$$

so the axion decay constants in the 2-cycle and 4-cycle sector scale like $1/N$ and $1/N^2$ respectively. In reality, the triple intersection form only contains order N non-zero entries. For example, in the diagonal case,

$$\mathscr{V} = \frac{1}{6}\sum_{i=1}^{N}(t^i)^3 = \frac{\sqrt{2}}{3}\sum_{i=1}^{N}(\tau_i)^{3/2} \simeq \frac{N}{6}\hat{t}^3 \simeq \frac{\sqrt{2}N}{3}\hat{\tau}^{3/2} , \tag{3.55}$$

we have $K_{ij} \sim 1/N$ and $K^{ij} \sim 1/N$,too. For the manifestly positive parametrisation (3.48), the volume always has to scale at least with one power of N, so $K^{ij}K_{jk}$ scales like $N^{-2\gamma}$ with $\gamma \geq 1$. Then $K_{ij} \sim N^{-\alpha}$ and $K^{ij} \sim N^{-\beta}$ with $2\gamma = \alpha + \beta$ and in addition $\alpha, \beta \geq 0$, since any derivative of the volume must scale with some smaller power than the volume itself. Inspecting the first term of K_{ij} we have

$$\frac{\mathscr{V}_{ij}}{\mathscr{V}} \sim \mathscr{V}_{ij}N^{-\gamma} \sim N^{-\alpha} , \tag{3.56}$$

and thus

$$\mathscr{V}_{ij} \sim N^{\gamma-\alpha} = N^{(\beta-\alpha)/2} . \tag{3.57}$$

From this we can see that $\beta \geq \alpha$. Since

$$\left(\frac{\mathscr{V}_{ij}}{\mathscr{V}}\right) t^i t^j = 3 = \mathscr{O}(1) , \tag{3.58}$$

and the sum contains at least (in the diagonal case) $\mathscr{O}(N)$ positive terms, K_{ij} has to scale at least like $1/N$. This again implies that K^{ij} scales at least like $1/N$, so we finally conclude that

$$K_{ij} \sim N^{-\alpha} \qquad K^{ij} \sim N^{-\beta} \qquad \text{with } \alpha, \beta \geq 1 . \tag{3.59}$$

This shows that the axion decay constants scale with negative powers of N and the $1/\sqrt{N}$ case is indeed the ideal one. This kills any N-hancement of the collective field range at least for naive N-flation. Taking the geometric mean between no scaling and maximal scaling (3.52)

$$\mathscr{V} \sim N^{3/2} , \qquad \mathscr{V}_i \sim N , \qquad \mathscr{V}_{ij} \sim N^{1/2} , \tag{3.60}$$

one has $K_{ij} \sim 1/N$ and $K^{ij} \sim 1/N^2$, reproducing the estimate of T. Grimm [64].

Positivity of Triple Intersection Numbers

In the above argument it was crucial that one could find a basis of 2-cycles with manifestly positive triple intersection numbers. In the original N-flation paper [35] and later in [65] it was argued that one could avoid the problems associated with a volume that scales with N by having a cancellation due to negative triple intersection numbers, i.e.

$$\mathscr{V} \sim \left(\sum_{ijk} k_{ijk}\right)\hat{t}^3 \simeq \hat{t}^3 . \tag{3.61}$$

The problem with this is that one always has to specify the physical field ranges of the Kähler moduli and taking $t^i \simeq \hat{t}$ might not be consistent. To see how this comes about let us first look at a real Calabi-Yau, the $P_{[1,1,1,6,9]}$ manifold as discussed for example in [66]. This has two Kähler moduli and the volume is given in terms of these as

$$\mathcal{V} = \frac{1}{3!}\left(3(t^1)^2 t^2 + 18 t^1 (t^2)^2 + 36 (t^2)^3\right) , \tag{3.62}$$

where t^i both take values in $\mathbb{R}^+$. The dual 4-cycle volumes are

$$\tau_1 = t^2(t^1 + 3t^2) , \qquad \tau_2 = \tfrac{1}{2}(t^1 + 6t^2)^2 . \tag{3.63}$$

The volume assumes a simple form only in terms of non-trivial linear combinations of the 4-cycle volumes τ_i

$$\mathcal{V} = \frac{\sqrt{2}}{18}\left(\tau_b^{3/2} - \tau_s^{3/2}\right) , \qquad \tau_b = \tau_2 , \quad \tau_s = \tau_2 - 6\tau_1 . \tag{3.64}$$

From their definitions, it is clear that the domain of the redefined τ is *not* $\mathbb{R}^+$ any more. The small cycle τ_s is constrained to be smaller than the big one. This ensures positivity of the overall Calabi-Yau volume. Geometrically the cycle τ_s is the exceptional divisor arising from the blowup of a singularity. In the limit $\tau_s \to \tau_b$, the blowup cycle devours the whole manifold. Now (3.64) suggests a generalisation. Take a manifold with many 4-cycles, half of which are blowups, contributing negatively to $\mathcal{V}$

$$\mathcal{V} = \sum_{i=1}^{N_+} (\tau_{i,+})^{3/2} - \sum_{i=1}^{N_-} (\tau_{i,-})^{3/2} . \tag{3.65}$$

It seems that if we take all τ of the same order of magnitude $\tau_{i,\pm} \simeq \hat{\tau}$ and at the same time $N_+ \simeq N_-$ we could achieve a volume that does not scale with N, $\mathcal{V} \sim \hat{\tau}^{3/2}$, and still have all cycles at large volume. To see that this is not the case it is instructive to translate the volume into the 2-cycle form and see what happens. It is clear that in terms of a basis of positively intersecting 2-cycles the only way a cancellation in the volume can happen is if certain cycles shrink to zero size and we will explicitly see that this is exactly the case. For ease of calculation we take $N_+ = 2$ and $N_- = 1$

$$\mathcal{V} = \tau_1^{3/2} + \tau_2^{3/2} - \tau_s^{3/2} . \tag{3.66}$$

This can be rewritten in terms of positively intersecting 2-cycles as

$$\mathcal{V} = (t^1)^3 + (t^2)^3 + (t^3)^3 + 3t^3(t^1+t^2)^2 + 3(t^3)^2(t^1+t^2) . \tag{3.67}$$

The dual 4-cycles are given by

$$\tau_1 = 3(t^1+t^3)^2 \qquad \tau_2 = 3(t^2+t^3)^2 \qquad \tau_3 = \tau_1 + \tau_2 - 3(t^3)^2 . \tag{3.68}$$

The blowup mode τ_s is again a non-trivial linear combination of the τ_i

$$\tau_s = \tau_1 + \tau_2 - \tau_3 = 3(t^3)^2 . \tag{3.69}$$

Now from the definitions we see that the domains of definition of the 4-cycles are $\mathbb{R}^+$ for τ_1 and τ_2 but τ_3 has to take values in $(0, \min\{\tau_1, \tau_2\})$. In particular this makes it impossible to go to the limit $\tau_i, \tau_b \simeq \hat{\tau}$ without collapsing 2-cycles.

3.9 Combined Alignment

The above findings seem to suggest that even if the metric on moduli space scales with inverse powers of N, the enhancement from the several alignment mechanisms as discussed in (2.6.5) might in combination still lead to an effective super-Planckian enhancement of the field range. At first sight, the explicit string theory on Calabi-Yau example of [32] supports the possible enhancement above the naive $\sqrt{N}$ level. In this case $N = 51$ is moderately large. The largest metric eigenvalue is $f_N \approx 0.013 M_p$, while the total field range along the lightest direction is computed to be $\mathscr{D} = 1.13 M_p$. While this seems to correspond to an $\mathscr{O}(N)$ enhancement (expected because $P > N$), the total field range is still $\mathscr{O}(M_p)$ and not parametrically super-Planckian.[5] This suggests that in this case the decay constants themselves scale as $1/N$. In fact, we have actually seen that the decay constants for the 4-cycle axions which are considered here typically scale exactly like $1/N$. The authors of [32] predict the diameter by their random matrix estimate correctly and suggest that one might simply look for larger N examples such that the enhancement is enough to get parametrically super-Planckian, but their estimate is only for the enhancement *relative* to the decay constants. We have no convincing argument why the cancellation should always happen, but as we will see in (5) there are general arguments against diameters of super-Planckian size.

[5] Actually it was argued in [11] that a more careful analysis of this model might lead to a sub-Planckian diameter.

4 Conjectures on the Moduli Space of Quantum Gravity

In lack of experimental guiding in finding the right theory of quantum gravity it is important to sometimes adopt an agnostic point of view and consider general theoretical constraints on *possible* theories of quantum gravity that can be deduced from low energy physics and consistency arguments. As we still miss a final understanding of all aspects of quantum gravity, most of these constraints naturally take a rather conjectural form. Given the tiny amount of bottom-up information that we have — the low energy effective field theory of gravity, Lorentz invariance, locality, black hole physics — it seems at first quite surprising that one can accomplish anything at all. Nevertheless, one can still gain some ground from analysing for example the quantum structure of black holes through their Hawking evaporation. Not without reason have black holes been termed the harmonic oscillator of the 21st century [67]. Such conjectures can then be checked explicitly in string theory, for which a tremendous amount of evidence supports its existence as an actual UV finite and constistent theory of quantum gravity. On the other hand string theory itself suggests several properties of quantum gravity for which one can then try to find more general evidence not relying on string theory. The aim of this chapter is to introduce and motivate several conjectures on quantum gravity that have been put forward.

4.1 The Swampland Conjecture

The tremendous amount of string vacuum constructions could lead one to think that in string theory anything goes and in particular that one could get any consistent low energy effective field theory from a stringy construction. This is in fact not so. The swampland program of Vafa [68] aims to discriminate between low energy effective field theories which can be UV completed to consistent quantum gravitational theories, so-called *(string-) landscape* theories, and those *swampland* theories which do not admit such a completion. Mainly motivated by string theory, Vafa conjectured a finiteness property of the scalar moduli space of quantum gravity. Concretely, if we fix a cutoff Λ and integrate out all fields with higher mass, this defines a restricted moduli space B_Λ. The volume of this should be generically finite in the limit as $\Lambda \to 0$ or at most diverge logarithmically in Λ

$$V_\Lambda = \int_{B_\Lambda} d\Phi \sqrt{g(\Phi)} \lesssim \log(1/\Lambda)\,, \quad \text{as } \Lambda \to 0\,. \tag{4.1}$$

A key example of this is the two-dimensional dilaton-axion moduli space in 10d type IIB string theory. The relevant metric on the moduli space is

$$g = \frac{dz \otimes d\bar{z}}{\mathrm{Im}\,(z)^2}\,, \tag{4.2}$$

and the integral of the associated volume form over $\mathbb{C}$ apparently diverges. The crucial point is that the integration range must be restricted to the fundamental domain of the $SL(2,\mathbb{Z})$ S-duality. This renders the volume of the moduli space finite. Other important examples include the moduli space of toroidal as well as $K3$ compactifications of type II string theory. The case of S^1 compactification actually leads to the logarithmic divergence (4.1). The mass of the Kaluza-Klein tower of states goes like $m \sim 1/R$. In order to consistently integrate out the Kaluza-Klein modes, we have to demand that $R \lesssim 1/\Lambda$. The integral of the volume form dR/R of the scalar moduli space from the T-dual radius to the UV cutoff then leads precisely to the logarithmic behaviour. Different examples of swampland theories that were suggested in [68] are theories with exotic gauge groups or theories with too many matter fields which cannot be constructed in string theory.

In a follow-up paper by Ooguri and Vafa [18] the question of finiteness of the scalar moduli space was adressed in a more precise manner. Among the conjectures proposed was the following, which we will call the *Swampland Conjecture* for simplicity since it will play a prominent role in the rest of this thesis.

Swampland Conjecture:
For any point p_0 in the continuous scalar moduli space of a consistent quantum gravity theory, there exist other points p at arbitrarily large distance. As the distance $d(p_0, p)$ diverges, an infinite tower of states exponentially light in the distance appears, meaning that the mass scale of the tower varies as

$$m \sim e^{-\alpha d(p_0,p)} \,. \tag{4.3}$$

The number of states in the tower which are below any finite mass scale diverges as $d \to \infty$.

The intuitive picture for this was already given in the introduction (1.3). Crucially, the evidence for this conjecture comes directly from string theory. Even if there is such a vast amount of stringy constructions they all seem to satisfy this property. Two particular examples of this were given above. In the S^1 reduction of the type II string, the infinite tower of light states in the one direction is precisely the Kaluza-Klein tower whose mass scales exponentially in the *proper* distance in moduli space. In the other direction it is the T-dual tower of winding strings. For the dilaton-axion moduli space of type IIB the pullback of the metric to the imaginary axis reproduces the metric of the S^1 reduction moduli space and we see that indeed for any point in the fundamental domain, the points at $i\infty$ and 0 are infinitely far away with logarithmic distance divergence. This is of course no coincidence as the type II dilaton can be interpreted as the radial modulus of the actual S^1 compactification of M-theory. From the type IIB perspective the degrees of freedom which are exponentially light in the distance are the F- and D-strings at $i\infty$ and 0, respectively. We see that points at infinte distance in moduli space can often be thought of as decompactification points and the swampland tower is realised as an actual Kaluza-Klein tower. Since string theory seems to be a consistent theory of quantum gravity it is conceivable that the fulfilment of this conjecture is indeed a necessary property of every such theory. A more radical point of view would be to conjecture that string theory is the only fully consistent realisation of quantum gravity and in this case the evidence from string theory alone is quite convincing. We will nevertheless give a bottom-up argument for why this conjecture should be true from general quantum gravitational arguments in chapter 6.

4.2 Absence of Global Symmetries

The statement of the following conjecture is sufficiently simple, so we will first state it and then explain the evidence for it.

> **No Global Symmetries:**
> A consistent theory of quantum gravity cannot have any continuous global symmetries.

The general quantum gravitational rationale behind this conjecture is that an exact continuous global symmetry implies an associated conserved charge by Noether's theorem. Such a conserved charge could be carried by black holes. It turns out that such a conserved quantum number is highly problematic if it is not associated with an accompanying gauge force. The exterior solution to such a black hole is independent on the charge and in effect it cannot be detected from outside because of the no-hair theorem. This means that an outside observer associates an infinite amount of entropy with this black hole because he cannot discriminate the infinite number of microstates, indexed by the value of the conserved charge, which live inside. This is in sharp violation of so-called *entropy bounds* such as the Bekenstein bound [69] or the covariant entropy bound [70] which bound the amount of entropy that a physical system can have. A different but related problem with global symmetries can be seen as follows. Consider the following Gedankenexperiment. Take a black hole carrying a conserved global symmetry charge Q_{glob}. This charge is undetectable from the outside and cannot be radiated away in Hawking particles because there is no force / chemical potential towards discharge. The net amount of particles with charge Q_{glob} and $-Q_{\text{glob}}$ will be equal and the black hole cannot decay completely. Thus the end product of Hawking evaporation must be a black hole *remnant* with mass $m \approx M_p$ and radius $r \approx l_p$ that is stable. Since the charge is not bounded from above, this leads to an infinite number of stable particles at each given mass. This is clearly pathological in many ways. It has been argued that an infinite number virtual remnant species appearing as intermediate states in scattering processes lead to the divergence of any scattering amplitude and thus to a meaningless physical theory.[1] Also an infinite number of species drive Newtons constant to zero [71]. While the arguments above make it extremely unplausible that a continuous global symmetry could be consistent with quantum gravity, string theory lends further evidence to this hypothesis. From the world sheet perspective it seems impossible to introduce global symmetries [72]. In brief, a continuous symmetry on the worldsheet leads to a conserved current which is actually the vertex operator for the associated gauge boson. For example the global super Poincare invariance of the worldsheet action leads to target space supergravity, the global symmetries of the heterotic string world sheet fermions lead to a target space Yang-Mills theory, just as the global Chan-Paton symmetry of the open string. A different aspect of this is that global symmetries in the low energy

[1] This is not entirely convincing since it depends cruicially on the UV completion. In fact string theory shows explicitly that an infinite tower of particles *can* make sense and lead to finite and exceptionally well-behaved scattering amplitudes.

effective theories obtained from string vacua always arise from gauge symmetries in the UV completion. An example for this phenomenon are the closed string axions. These can arise for example from form fields with indices along cycles of the internal compactification manifold. In this case the gauge symmetry of the form field forbids the appearance of non-derivative terms in the action of the axion. This results in a global continuous shift symmetry [73]. But even this apparent global symmetry is violated non-perturbatively by instanton effects from branes wrapping the corresponding cycle.

4.3 The Weak Gravity Conjecture

The *Weak Gravity Conjecture* [9] is a natural sharpening of the conjectured absence of global symmetries. Gobal symmetries are the $g \to 0$ limit of local ones. Since we have seen that global symmetries are pathological in quantum gravity there should be some feature of effective field theories coming from a consistent quantum gravity theory that fights against taking this limit. In particular any effective description should break down as one takes $g \to 0$. The breakdown of an effective field theory is usually accompanied by the appearance of new light states. Thus, if our effective description is valid up to a cutoff of Λ, we expect the limiting behaviour

$$\lim_{g \to 0} \Lambda(g) = 0 \,, \tag{4.4}$$

and new light states should enter the stage at around this scale. One could suspect a bound on particle masses charged under a local symmetry in terms of their gauge coupling as follows. The charge Q_{loc} of a black hole can be measured from outside by measuring the effect of the long-range gauge force on test particles. But if the gauge coupling is tiny, the force will be weak and one needs charged test particles of very small mass to detect anything [74]. If no such particle is in the spectrum the problems from above will reappear in the limit $g \to 0$. Thus quantum gravity should forbid this limit. Let us see how this works and first fix conventions. We will consider charged black hole solutions to the action

$$S = \frac{1}{2\kappa^2} \int \star R - \frac{1}{2g^2} \int F \wedge \star F \,, \tag{4.5}$$

where electric charges are defined by

$$Q = \frac{1}{g^2} \int_{S^2} \star F \,. \tag{4.6}$$

In order to avoid pathological black hole remnants, we demand that every charged black hole in the theory can eventually decay. It is clear that if extremal black holes are kinematically allowed to decay, then every subextremal black hole will also be able to decay. Thus it suffices to consider an extremal black hole of mass and charge $M = \sqrt{2} Q M_p$. Its mass to charge ratio is of order the Plack mass, $\Gamma = M/Q = \sqrt{2} M_p$. Suppose the minimal mass to charge ratio in the spectrum is $\gamma_{\text{min.}}$ and the black hole decays into a collection of particles with $\gamma_i = m_i/q_i$.

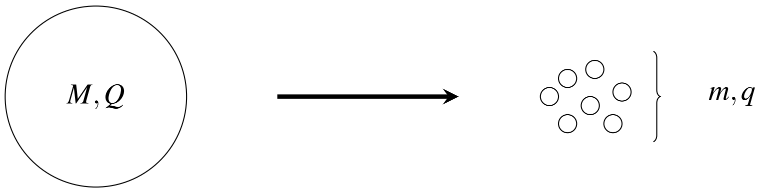

Figure 4.1: The statement of the Weak Gravity Conjecture follows from the requirement that all black holes (M,Q) should be able to decay to particles (m,q).

Kinematics demand that $\sum m_i \leq M$ and charge conservation demands that $\sum q_i = q$. It follows that

$$\Gamma \geq \frac{1}{Q}\sum_i m_i = \frac{1}{Q}\sum_i \gamma_i q_i > \frac{1}{Q}\gamma_{\text{min.}}\sum_i q_i = \gamma_{\text{min.}}\,. \tag{4.7}$$

From this one can see that the following Weak Gravity Conjecture[2] must be necessarily fulfilled.

Electric Weak Gravity Conjecture :
For a U(1) gauge theory coupled to quantum gravity there exists a charged particle with mass m and charge q satisfying

$$m \leq \sqrt{2}qgM_p\,. \tag{4.8}$$

This is the weakest form of the WGC. It states that some state in the theory is super-extremal and is agnostic about the precise nature of that state.[3] The actual argument from black hole decay (4.7) hints that the above inequality should in fact hold for the state of smallest mass to charge ratio. There exists an even stronger form which demands that the state should actually be the lightest charged particle in the spectrum [9]. We will often omit the factor of q in (4.8) since it can be absorbed into the definition of the gauge coupling. Before discussing generalisations of this conjecture we will investigate the behaviour of the WGC under electric-magnetic duality. First, we could consider magnetically charged black holes and the corresponding decay to magnetic monopoles. This would result in the same claim as (4.8) but with the electric gauge coupling g replaced by the magnetic $1/g$ one. In the far field limit the magnetic monopole's field is given by

$$\mathbf{B} = \frac{p}{g^2r^2}\mathbf{e}_r\,. \tag{4.9}$$

Suppose now that we deal with a $U(1)$ theory that has a cutoff Λ. The mass of a magnetic monopole should be of order of its field energy

$$m_{\text{mon}} \approx \mathscr{E}_{\text{mon}} = \int_{r\leq 1/\Lambda} dV\,\varepsilon(x) = \int_{r\leq 1/\Lambda} dV\,\mathbf{B}^2 \propto \frac{p^2}{g^2}\int_{1/\Lambda}^{\infty}\frac{1}{r^2}dr = \frac{p^2}{g^2}\Lambda\,. \tag{4.10}$$

[2] It was called Weak Gravity Conjecture in [9] because it implies that there are states in the spectrum for which the gravitational attraction is trumped by the electric repulsion.

[3] The state implied by the Weak Gravity Conjecture can actually be extremal and this is the case for example with the BPS branes in string theory. These states then marginally fulfil the WGC bound. In [75] it was conjectured that this is indeed the case *iff* the theory is supersymmetric.

Together with the decay argument this leads to the

Magnetic Weak Gravity Conjecture :
A U(1) gauge theory coupled to quantum gravity is at most valid up to a cutoff of

$$\Lambda \lesssim g M_p \, . \tag{4.11}$$

A somewhat different argument can be given as follows. Consider the unit charged monopole. We do not expect it to be a black hole but rather to behave as an ordinary particle. Thus its radius $r_\Lambda \sim 1/\Lambda$ should be bigger than its associated Schwarzschild radius $r_S = 2GM$ in order to not collapse to a black hole. This gives

$$r_\Lambda = \frac{1}{\Lambda} \gtrsim r_S \sim \frac{\Lambda}{g^2 M_p^2} \, . \tag{4.12}$$

Hence we recover the magnetic Weak Gravity Conjecture (4.11). It should be noted that the existence of a cutoff usually means loss of perturbative unitarity at that scale. Loss of unitarity means that there is a non-zero probability for particles to scatter out of the physical Hilbert space and hence into new states that one did not take into account. In this sense the existence of a low cutoff signals the presence of the new states which are needed for the black holes to decay. The magnetic Weak Gravity Conjecture was challenged in [76], where it was explicitly argued that certain non-perturbative and extended, monopole-like objects can still be below the black hole threshold even if the fundamental monopole of corresponding charge would in fact be a black hole.

Next, we will consider the generalisation of the Weak Gravity Conjecture to the product gauge group $\prod_{a=1}^{N} U(1)$. This has been worked out in [74]. Suppose the theory contains particles with charge vectors $\mathbf{q}_i = (q_{i_a})_{a=1\ldots N}$. Define the dimensionless charge-to-mass ratios $\mathbf{z}_i = \sqrt{2}\mathbf{q_i} M_p/m_i$. By $SO(N)$ invariance of the Einstein-Maxwell system, where the $SO(N)$ rotates the gauge fields into each other, one can see that the generalisation of the Weak Gravity Conjecture should be $SO(N)$ invariant (an exception being theories with non-trivial scalar dependent kinetic matrix). It is instructive to repeat the black hole decay argument. Consider decay of an extremal black hole with charge vector $\mathbf{Q}$, mass M and charge-to-mass ratio $\mathbf{Z} = \sqrt{2}\mathbf{Q} M_p/M$ into a final state of n_i particles with charge-to-mass vector $\mathbf{z}_i$ as defined above. Charge conservation implies $\mathbf{Q} = \sum n_i q_i$, while energy conservation implies $M > \sum n_i m_i$. Definining also the mass fraction per particle species $\sigma_i = n_i m_i/M$ one finds that

$$1 > \sum_i \sigma_i \, , \qquad \mathbf{Z} = \sum_i \sigma_i \mathbf{z_i} \, . \tag{4.13}$$

This is the requirement that

$$\mathbf{Z} \in \mathrm{Conv}(\pm\mathbf{z}_1, \ldots, \pm\mathbf{z}_N) \, . \tag{4.14}$$

For extremal black holes $|\mathbf{Z}| = 1$ so the convex hull of the charge-to-mass vectors should contain the unit ball. This is the *convex hull condition* of the Weak Gravity Conjecture for multiple $U(1)$

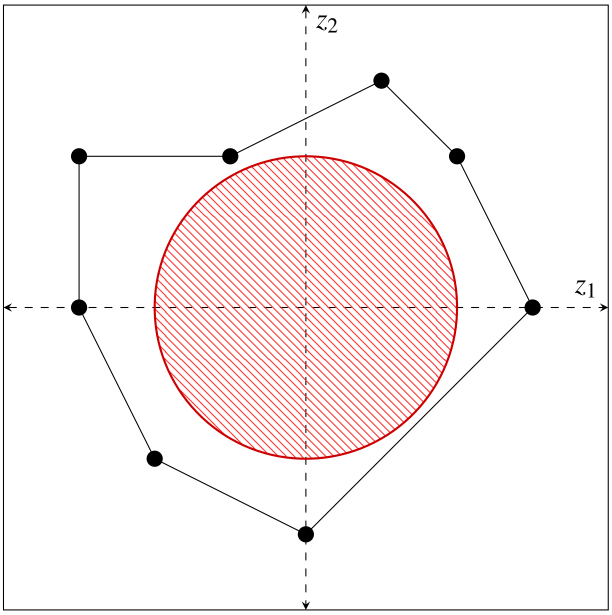

Figure 4.2: The convex hull condition for multiple $U(1)$s. The filled open disk contains sub-extremal black holes, while the boundary circle consists of extremal black hole solutions. The black dots spanning the convex hull are the super-extremal Weak Gravity Conjecture states.

gauge fields. It is altered in theories where the kinetic matrix of the gauge bosons is scalar dependent and the $SO(N)$ symmetry is broken. The unit ball is then replaced by the appropriate co-dimension 1 shape corresponding to extremal black hole solutions of the theory.

The Convex Hull Condition for Multiple $U(1)$s
For a $U(1)^n$ gauge theory, the convex hull of the Planck normalised charge-to-mass vectors $\mathbf{z}_i = \sqrt{2}\mathbf{q_i}M_p/m_i$ of the charged particles in the theory should contain the extremal black hole solutions of the theory.

Another quite different way to generalise the Weak Gravity Conjecture is to consider p-form gauge theories in d dimensions as ubiquitous in string theory [9, 77]. For sake of generality we also allow for an additional dilatonic scalar field

$$S = \frac{1}{2\kappa^2}\int\left(\star R - \frac{1}{2}d\phi\wedge\star d\phi\right) - \frac{1}{2g^2}\int e^{-\alpha\phi}F_{p+1}\wedge\star F_{p+1}\ . \tag{4.15}$$

In the conventions of [77], the charges in the theory are defined by

$$Q = \frac{1}{g^2}\int_{S^{d-p-1}} e^{-\alpha\phi}\star F_{p+1}\ , \tag{4.16}$$

and the generalised Weak Gravity Conjecture states that [77]

p-Form Weak Gravity Conjecture
For each Abelian p-form gauge field in the theory, there must exist a p-dimensional object with tension T_p and charge q such that

$$\left[\frac{\alpha^2}{2}+\frac{p(d-p-2)}{d-2}\right]T_p^2 \lesssim \frac{q^2g^2}{\kappa^2}\,. \tag{4.17}$$

For $0<p<d-2$ the argument is precisely the same as for $p=1$. The black holes simply get replaced by black $(p-1)$-branes and these should be able to decay in order to prevent $(p-1)$-remnants. The formula (4.17) is actually degenerate for $p=0,d-2,d-1,d$. The case for a Weak Gravity Conjecture for $p=0$ is not so solid since there are no objects charged under a 0-form. For $p=d-2$ we are dealing with codimension 2 objects and these induce a deficit angle in their transverse geometry (these are cosmic strings, D7 branes, etc.) so there actually is a maximum tension for these objects such that the surrounding space is not destroyed. For $p=d-1$ the gauge field is non-dynamical and finally for $p=d$ we are dealing with spacetime filling objects. Nevertheless, in [11, 78] it was argued that in string theory similar constraints should arise by relating the questionable cases to the particle and gauge field case via string dualities.

4.3.1 The (Sub-) Lattice Weak Gravity Conjecture

Now we would like to look at a proposed extension of the Weak Gravity Conjecture which implies that there must be not only a single charged particle of appropriate mass but a whole *tower* of particles fulfilling the Weak Gravity Conjecture. This is the *Lattice Weak Gravity Conjecture* (LatWGC) [77, 79].

Lattice Weak Gravity Conjecture:
For a $U(1)$ gauge theory with charge lattice Γ, there exists a sub-lattice $\Gamma_{\text{WGC}}\subseteq\Gamma$ such that every site is occupied by a super-extremal particle.

Why should such a strong statement be true? A first reason is that the $g\to 0$ limit does not have a very drastic effect in the original WGC setting (4.8,4.11). Even if the effective field theory breaks down for $g\to 0$ in a conservative interpretation this only means that we have to include the new massless state implied by the Weak Gravity Conjecture in our effective description. The LatWGC actually implies something much stronger that is the breakdown of *any* effective description since an infinite tower of states becomes massless as we take $g\to 0$. This then prohibits the appearance of global symmetries as a limit of gauge symmetries at the level of consistent EFTs arising from quantum gravity. Ultimately, the question must of course be settled in a proper UV complete description of the LatWGC tower. For example in string theory in a simple S^1 reduction the gauge coupling will be given as the radius of the circular dimension. If this is sent to zero we reduce the number of dimensions by one. This is not consistent since string theory fixes the number of spacetime dimensions. Another reason is that in [77] it was noted that the Weak Gravity Conjecture as in (4.13) is not stable under dimensional reduction. If one reduces a $U(1)$ theory consistent with the Weak Gravity Conjecture to a lower dimensional theory, the mixing with the Kaluza-Klein photon can spoil the Weak Gravity Conjecture in the lower dimensional theory. This is because the KK modes are the Weak Gravity Conjecture states for the KK photon and these sit precisely at the Weak Gravity Conjecture *threshold* so there is not much wiggle room. In [79] evidence was presented that the LatWGC is actually satisfied in string theory and consistency under dimensional reduction was shown. The reader is referred to the above references for further details.

4.4 The Completeness Conjecture

The Lattice Weak Gravity Conjecture can be seen as a sharpening of the Weak Gravity Conjecture, but it can moreover also be seen as a sharpening of a different and unrelated conjecture. This is the *Completeness Conjecture* [20, 21]. The Dirac quantisation condition is a constraint on the charges of magnetic monopoles in a theory that also contains electrically charged objects. It states that for pair of electric q and magnetic p charges in the theory

$$qp = 2\pi n\,, \quad n \in \mathbb{Z}\,. \tag{4.18}$$

In particular the fundamental magnetic charge quantum must be an integer multiple of 2π times the fundamental electric charge quantum. It is not implied that the fundamental magnetic charge is the one for which $n = 1$. This is the statement of the

Completeness Conjecture: Every site in the lattice of possible charges allowed by Dirac quantisation is occupied.

The Lattice Weak Gravity Conjecture states that this lattice has a sublattice occupied by superextremal, (approximately) stable states. In [21] it was argued that it is not possible to render charged fields non-dynamical in a quantum gravity theory by sending their mass to infinity — before being able to do so, the state would collapse to a black hole. Even if no field of a given charge is included in the theory to begin with, the corresponding Wilson line/t'Hooft operator will obtain a renormalised mass by interacting with the gravitational field and the black hole state of beforementioned charge must be in the spectrum.

4.5 Consistency of Natural Inflation

Some of the conjectures that we discussed in this chapter would have important implications for large field inflation if true. This is in particular clear for the Swampland Conjecture but also true for the Weak Gravity Conjecture as will become clear.

4.5.1 Single Field Natural Inflation

In this case the moduli space is spanned by a single axion θ with a discrete, gauged shift symmetry $\theta \to \theta + 2\pi f$ and thus is a topological S^1. The swampland behaviour (4.3) is incompatible with the shift symmetry so the only consistent realisation of such an axion in quantum gravity is with a sub-Planckian periodicity. This ensures that there is no need for the swampland behaviour to set in since we can never traverse super-Planckian distances in moduli space. Consequently, the Swampland Conjecture requires all axion decay constants in quantum gravity to satisfy $f \leq M_p$.

The zero form version of the Weak Gravity Conjecture (4.17) provides a similar constraint. We can interpret the axion as a 0-form (scalar) gauge field with a 1-form field strength $d\theta$. A 0-form gauge field naturally couples to dimension 0 objects, that is objects localised in space and time. In the Euclidean path integral these are just instantons. The action of an instanton can be interpreted as its tension and the coupling of the axion to the instantons is given by the inverse decay constant. The 0-form version of the Weak Gravity Conjecture is then the statement that

$$Sf \leq M_p \,. \tag{4.19}$$

The instanton expansion is governed by powers of e^{-S} and demanding this to be controlled ($S > 1$) we again find that

$$f \leq M_p \,. \tag{4.20}$$

The upshot is that axionic fields can only have sub-Planckian decay constants in quantum gravity and this is also what is observed in string theory [19]. This rules out single field natural inflation models because for these we would need precisely the super-Planckian decay constants. We are thus led to look for ways to enhance the effective decay constants such as in aligned natural inflation or N-flation.

4.5.2 Aligned Natural Inflation, N-flation and Kinetic Alignment

As we have seen in section (2.6) the effective field theory of N axions allows for the possibility to displace beyond the individual axion decay constants (which become an ambiguous concept). As a measure of the available field range we took the diameter of the fundamental domain. We have reviewed the possible enhancements of the field range and one might hope to evade the Weak Gravity Conjecture in this way. Nevertheless, we have seen at least in the case of particles and gauge fields that the Weak Gravity Conjecture actually also gets stronger if we increase the number of gauge symmetries due to the convex hull condition (4.13). In [11] a similar convex hull condition was derived by T-dualising the corresponding convex hull condition in the gauge field and charged particle setup. Using this, in [80] it was shown that in the regime of perturbative control the convex hull condition for N axions leads to the result that the fundamental domain diameter is bounded by

$$\mathscr{D} \leq 2\pi \,. \tag{4.21}$$

This means that in all these cases the only way to evade the Weak Gravity Conjecture in order to get super-Planckian displacements for inflation is to assume that the instantons cutting out the fundamental domain are sub-dominant to stronger and less constraining instanton contributions.

This is equivalent to the question whether the strong or the weak version of the Weak Gravity Conjecture is realised. The situtation was illustrated in (2.8) and the generalisation to the higher dimensional moduli space of N axions is obvious. String theory lends support to the belief that it is indeed the strong version of the Weak Gravity Conjecture which is true [9] and thus the fundamental domain diameter should be a good proxy for the available inflationary field range.

We conclude that the Weak Gravity Conjecture and Swampland Conjecture most likely rule out inflationary scenarios involving only N axions in absence of monodromy.[4]

4.5.3 Monodromy

Since the axion monodromy potential does not respect the shift symmetry $\theta \to \theta + 2\pi f$, the above constraints on f do not directly constrain models of axion monodromy. Nevertheless in [78] it was shown that the magnetic version of the Weak Gravity Conjecture actually does have something to say about the domain walls describing the tunneling between different branches of the potential. The general philosophy of the Swampland Conjecture suggests that even in the case of monodromy for $\Delta\phi \gtrsim M_p$ the infinite tower of massive states which become light should lead to a breakdown of the effective description and could possibly spoil the inflationary potential. We conclude by remarking that while there might be some tension, $\Delta\phi \gtrsim M_p$ is not sharply excluded for monodromy models and these provide the most promising UV realisation of large field inflation.

[4]This is at least likely for inflation models coming from string theory.

5 Sharpening Conjectures

The aim of this chapter is to sharpen and refine two of the conjectures introduced in chapter 4, namely the Swampland Conjecture and the Weak Gravity Conjecture. This and the next chapter are based on our recent paper [17]. While we follow the general logic of [17], we aim to expand on certain points that were only briefly discussed therein.

5.1 The Local Weak Gravity Conjecture

The first conjecture that we will discuss is the Weak Gravity Conjecture in its most simple incarnation[1]

$$m \leq g M_p \ . \tag{5.1}$$

We notice that although this is no necissity, we expect that in proper quantum gravity all couplings become dynamical fields. This is definitely the case for string theory, where they are controlled by moduli fields which often parametrise a compactification geometry. This in turn implies that the Weak Gravity Conjecture should hold *locally* in the moduli space of vacua of the theory

$$m(\phi) \leq g(\phi) M_p \ . \tag{5.2}$$

In effect, the Weak Gravity Conjecture should be viewed as a constraint on the moduli space of vacua $\mathscr{M}$ of a consistent quantum gravity theory. In the next chapter we will be interested in studying displacements of moduli from particular points in $\mathscr{M}$. It is clear that if we displace in $\mathscr{M}$ while keeping the moduli VEVs homogeneous the Weak Gravity Conjecture continues to hold for each value of the displacement. It is however not so clear how we should interpret the Weak Gravity Conjecture if we displace the moduli in four dimensional space. The most sensible interpretation seems to be that the Weak Gravity Conjecture should hold also locally in space, away from a flat space vacuum configuration

$$m(x) \leq g(x) M_p \ . \tag{5.3}$$

We should note that from the black hole decay argument in favour of the Weak Gravity Conjecture the naive requirement seems to be only that it should be asymptotically satisfied

$$m_\infty \leq g_\infty M_p \ , \tag{5.4}$$

since we are worried about the asymptotic decay of black holes. We will see however that there are other reasons to believe that (5.3) should be true. To illustrate the above, consider the following simple example. For an asymptotically flat extremal black hole we have that the

[1] Here and in the following chapter we will be not very careful about $\mathscr{O}(1)$ factors.

solution interpolates between two vacua of the theory — four dimensional Minkowski space at infinity and $AdS_2 \times S^2$ on the horizon.[2] It is evident that the Weak Gravity Conjecture should hold in these regions of the spacetime. The crucial question is what happens for $r_h < r < \infty$. This question is hard to answer in full generality. If the solution preserves some supersymmetry and the WGC state is a BPS state then the states mass will sit at the WGC threshold everywhere in space, so indeed we can take this as evidence for (5.3).

There should also be a local version of the magnetic Weak Gravity Conjecture (4.11). This should bound some local energy scale of the theory by the gauge coupling. To get some intuition we should revisit the derivation. The magnetic WGC follows from demanding that the unit charge magnetic monopole in the theory should be a particle and not a black hole. If we associate a cutoff radius r_Λ to the monopole, then we should have

$$\frac{1}{r_\Lambda} < g M_p \ . \tag{5.5}$$

In the case where the gauge coupling was constant over space we would interpret the left hand side as a constant Wilsonian cutoff to the theory. In the case where g is allowed to vary over space this would be a very strong constraint. We can also reinterpret this in terms of the local energy density in the gauge field $\rho \simeq 1/g^2 r^4$ and write

$$g M_p > \frac{\rho(r_\Lambda)^{\frac{1}{2}}}{M_p} \ , \tag{5.6}$$

where the right hand side can be interpreted as the Hubble scale of the theory $H M_p \simeq \sqrt{V}$. We can thus propose the mild generalisation of the magnetic Weak Gravity Conjecture that the local energy scale, set by the Hubble scale, should be below the mass of the WGC states. If we want to consistently integrate out a state which interacts only through the gravitational force, its loops would induce operators suppressed by powers of $m M_p$, so we should impose a cutoff of $\Lambda^2 < m M_p$. This means that the consistent decoupling of the Weak Gravity Conjecture states is achieved if[3]

$$H M_p \simeq \sqrt{\rho} < m_{\mathrm{WGC}} M_p < g M_p^2 \ . \tag{5.7}$$

Note that this interpretation is in fact stronger than the electric and magnetic WGC alone since it requires the cutoff scale to be below the mass scale of the WGC states. To summarise, for spatially varying gauge couplings the local generalisations of the electric and magnetic WGC that we propose are

[2] In [81] pathologies were shown to arise when the WGC is not fulfilled from an analysis of the near horizon geometry of extremal black holes.

[3] It should be noted that the WGC states also couple through gauge interactions and these are indeed stronger due to the weak gravity requirement, although the gravitational coupling is universal and should be the relevant one for a quantum *gravity* constraint. Alternatively, one could think of imposing $m_{\mathrm{WGC}} > \rho^{1/4}$.

The Local Weak Gravity Conjectures (Weak Curvature)

$$m_{\mathrm{LWGC}}(x) \leq g(x)M_p \,, \tag{5.8}$$
$$\sqrt{\rho(x)} < g(x)M_p^2 \,. \tag{5.9}$$

We are assuming that some form of the lattice WGC holds such that m_{LWGC} is actually the mass scale of an infinite tower of states. A useful example to keep in mind is the gauge theory arising from a simple circle Kaluza-Klein compactification. In this case the WGC states are precisely the KK modes of the massless scalar which parametrises the diameter fo the circular extra dimension and these sit exactly at the WGC threshold. The mass scale of the KK tower is $m_{\mathrm{LWGC}} \sim 1/R$, which coincides with the gauge coupling. This is also true over all of space if we displace $R(x)$ locally in four dimensional space and in order for the effective 4d description not to break down we should not excite the KK modes, leading to the requirement (5.9).

We could also have postulated a more constraining form of the magnetic WGC, namely that the mass of the WGC states should stay above the *maximum* local energy scale of the solution. This would in fact be appropriate if we were to describe the solution in terms of a Wilsonian EFT with a constant cutoff. We adopt the point of view that this is indeed too strong since the weaker requirement (5.9) still enables a local observer to set up an effective description in its vicinity. Furthermore, the operators which appear in the EFT upon integrating out the WGC states are subleading when evaluated on the actual solution. This means that from the viewpoint of locality (5.9) is just strong enough. It is also unlikely that a much weaker form than (5.9) should suffice as the Kaluza-Klein example illustrates.

We should note that the proposed magnetic local Weak Gravity Conjecture (5.9) does not transform covariantly under diffeomorphisms and thus the energy density on the left hand side should be replaced by an appropriate scalar quantity when we study strongly curved backgrounds. The energy momentum tensor (of which ρ is the tt-component in flat space) is related to the Ricci tensor through the Einstein field equations $M_p^2 R \sim T$ and so natural candidates are scalar invariants built out of the curvature tensors like the the Ricci scalar $M_p\sqrt{R}$, the square of the Ricci tensor $M_p(R_{ab}R^{ab})^{1/4}$ or the Kretschmann scalar $M_p(R_{abcd}R^{abcd})^{1/4}$. The result are the following generalised local Weak Gravity Conjectures

The Local Weak Gravity Conjectures (Strong Curvature)

$$m_{\mathrm{LWGC}}(x) \leq g(x)M_p \,, \tag{5.10}$$
$$\frac{1}{R_c(x)} < g(x)M_p \,. \tag{5.11}$$

where R_c is a typical curvature radius of the solution, which can be quantified by one of the above mentioned curvature scalars.

To illustrate the above conjectures it is useful to study a concrete example of a theory with a spatially varying gauge coupling. For simplicity we will restrict to a single gauge field with gauge coupling controlled by a single modulus and we work in the weak gravity approximation of the Newtonian limit of GR. We will furthermore restrict to spherically symmetric solutions. The action is

$$S = \frac{1}{2}\int \left(\star R - 2d\phi\wedge\star d\phi - 2e^{2\alpha\phi}F\wedge\star F\right) . \tag{5.12}$$

The Newtonian approximation of GR amounts to working in a regime where

$$ds^2 = -[1+2\Phi(r)]dt^2 + [1-2\Phi(r)](dr^2 + r^2 d\Omega^2) , \qquad |2\Phi| \ll 1 , \tag{5.13}$$

and Φ is the Newtonian potential. It is defined as the general solution to the Laplace equation

$$\Delta\Phi = \frac{1}{4}\rho , \tag{5.14}$$

where ρ is the energy density of the matter fields and we impose the asymptotic boundary condition $\Phi(r\to\infty) = 0$. We set the reduced Planck mass equal to one $8\pi G = 1$. The solution in the spherically symmetric case is given by

$$\begin{aligned} -2\Phi(r) &= \frac{1}{8\pi}\int dV \frac{\rho(r')}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{8\pi}\int dr'\wedge d\Omega \frac{\rho(r')}{\sqrt{r^2 + r'^2 + 2rr'\cos(\theta)}} \\ &= \frac{1}{2r}\int_0^r dr' r'^2 \rho(r') + \frac{1}{2}\int_r^\infty dr' r' \rho(r') . \end{aligned} \tag{5.15}$$

At this point let us specialise to a magnetic monopole solution, which we will term the *Dilaton monopole*, where the gauge field assumes the form

$$F = p\sin(\theta)d\theta\wedge d\phi , \qquad \Rightarrow \qquad F\wedge\star F = \star\frac{p^2}{r^4} . \tag{5.16}$$

In this background the equation of motion for ϕ,

$$d\star d\phi = \alpha e^{2\alpha\phi}\star\frac{p^2}{r^4} \qquad \Leftrightarrow \qquad \partial_r\left(r^2\partial_r\phi\right) = \alpha e^{2\alpha\phi}\frac{p^2}{r^2} , \tag{5.17}$$

admits the solution [82]

$$\phi = -\frac{1}{\alpha}\ln\left[g_\infty\left(1+\frac{r_F}{r}\right)\right] , \qquad r_F = \frac{\alpha p}{g_\infty} . \tag{5.18}$$

Here r_F is the radius at which the scalar starts to asymptote to a free field,

$$\phi \simeq -\frac{1}{\alpha}\left(\ln(g_\infty) + \frac{r_F}{r}\right) , \qquad r \gg r_F . \tag{5.19}$$

Since we are working in natural units α is a dimensionless constant. As we approach the monopole source the solution necessarily breaks down since gravity becomes strong. The Newtonian potential (5.15) at a radius r contains contributions from above and below r. Since we do not trust the solution down to arbitrarily low r let us pick out the contribution above r and define the radius at which this equals to one to be r_N

$$\tilde{\Phi}(r_N) = \frac{1}{2}\int_{r_N}^{\infty} dr' r' \rho(r') \stackrel{!}{=} 1 \ . \tag{5.20}$$

For the Dilaton monopole, the energy density from the scalar field gradient is precisely the energy density stored in the gauge field

$$\rho_\phi = 2\phi'^2 = \rho_F = 2\frac{p^2}{r^4}e^{2\alpha\phi} = \frac{2}{\alpha^2}\frac{r_F^2}{r^2(r+r_F)^2} \ . \tag{5.21}$$

Crucially, for $r \ll r_F$, the scalar behaves logarithmically with r,

$$\phi \simeq -\frac{1}{\alpha}\ln\left(\frac{r}{r_F}\right) \ , \tag{5.22}$$

and thus the energy density drops only like $1/r^2$ in this regime as opposed to $1/r^4$ in the free field regime. Accounting for both contributions of the energy density we find

$$\tilde{\Phi}(r_N) = \frac{2r_F^2}{\alpha^2}\int_{r_N}^{\infty}\frac{dr'}{r'(r'+r_F)^2} = \frac{2}{\alpha^2}\left[-\frac{r_F}{r_N+r_F} + \ln\left(\frac{r_N+r_F}{r_N}\right)\right] \stackrel{!}{=} 1 \ . \tag{5.23}$$

Introducing the variable $x = r_F/r_N$ we have

$$\frac{\alpha^2}{2} = \ln(1+x) - \frac{x}{1+x} \ , \tag{5.24}$$

which can be solved approximately in the two regimes $x \gg 1$ and $x \ll 1$.

$$x \simeq \begin{cases} e^{\frac{\alpha^2}{2}} \ , & x \gg 1 \\ \frac{\alpha}{\sqrt{2}} \ , & x \ll 1 \end{cases} \ . \tag{5.25}$$

From this it is evident that the solution for $x \gg 1$ corresponds to $\alpha \gg \sqrt{2}$, whereas the one for $x \ll 1$ corresponds to $\alpha \ll \sqrt{2}$.

$$\frac{r_F}{r_N} \simeq \begin{cases} e^{\frac{\alpha^2}{2}} \ , & \alpha \gg \sqrt{2} \\ \frac{\alpha}{\sqrt{2}} \ , & \alpha \ll \sqrt{2} \end{cases} \ . \tag{5.26}$$

We see that α controls the separation between the free field radius r_F and the radius r_N at which the Newtonian approximation breaks down. The ratio between the gauge coupling and the energy density is

$$\frac{g(r)}{\sqrt{\rho(r)}} = \frac{1}{2}\alpha^2 q\left(1+\frac{r}{r_F}\right)^2 \geq \frac{1}{2}\alpha^2 q\left(1+\frac{r_N}{r_F}\right)^2 \ . \tag{5.27}$$

Let us consider the unit charge monopole $p = 1$. The limit $\alpha \to 0$ corresponds decoupling the scalar from the gauge field and in this case

$$\frac{g(r)}{\sqrt{\rho(r)}} \gtrsim \frac{1}{2}\left(\alpha+\sqrt{2}\right)^2 \simeq 1 \ . \tag{5.28}$$

We see that the magnetic WGC is always satisfied and the same conclusion holds for large α. In the small α case we flow into the usual magnetic WGC constraint (4.12) because the gauge coupling is approximately constant over a larger and larger range. For large α the local WGC deviates significantly.

$$\rho\left(r_N\right)^{\frac{1}{2}} \simeq \frac{2g\left(r_N\right)}{\alpha^2} \simeq g_\infty \frac{2e^{\frac{\alpha^2}{2}}}{\alpha^2} \, . \tag{5.29}$$

Because the gauge coupling at infinity is exponentially smaller than the energy density in the strong gravity regime, imposing the magnetic WGC in a global way by considering the gauge coupling at infinity as the relevant quantity renders this solution inconsistent. Nevertheless every magnetic point source in a corresponding UV completion should flow to the above in the IR. We view this as an indication that the local version of magnetic WGC (5.9) should be the correct one.

5.2 The Refined Swampland Conjecture

After motivating a natural local version of the electric and magnetic Weak Gravity Conjecture, we proceed with the Swampland Conjecture, which was introduced in section (4.1). The original form of the Swampland Conjecture (4.3) is not very precise about the amount of displacement in field space needed to induce the exponential mass drop of the SC tower of states. It seems unlikely that the local geometry of the moduli space $\mathscr{M}$ is in general severely constrained by quantum gravity physics — we do not expect the exponential drop to necessarily set in for infinitesimal displacements. Nevertheless, we can imagine a stronger statement than just one about displacements asymptoting to infinity, since quantum gravity comes equipped with a built in scale, namely the Planck mass M_p. The idea behind [17, 83] is thus that the exponential drop cannot be delayed much beyond the Planck scale. We will first look at the motivating example of [83] and then discuss our general proposal for a *refined Swampland Conjecture*.

The *exponential* drop in masses for super-Planckian distances in field space of the refined Swampland Conjecture can be related to a *logarithmic* behaviour of the proper distance in field space as follows. Consider a one-dimensional sub-manifold of moduli space parametrised by a single scalar field φ, equipped with the pullback metric $g_{\varphi\varphi}$. If φ or some power of it controls the mass of a tower of fields, as is the case for the KK scalars in compactifications, then the mass of the tower is exponential in the proper distance, calculated as

$$\Delta\phi = \int_{\phi_1}^{\phi_2} \sqrt{g_{\varphi\varphi}} d\varphi \, , \tag{5.30}$$

if it grows only logarithmically with φ for $\Delta\Phi \geq M_p$,

$$\Delta\phi \simeq \frac{1}{\alpha} \log\left(\frac{\varphi_2}{\varphi_1}\right) . \tag{5.31}$$

Equivalently, for $\Delta\phi \geq M_p$ the field space metric should asymptote to $g_{\varphi\varphi} \simeq 1/\varphi^2$. This is certainly evident for toroidal compactifications as was noted in (4.1). In [83] evidence was presented that this behaviour of the field space metric indeed emerges for $\Delta\phi \geq M_p$ in the setting of monodromy axions in Type IIA on Calabi-Yaus. As we have reviewed in section (3.7), in these models a single linear combination of the axions obtains a potential from fluxes while the other represent flat directions. If the fixed linear combination is given as $\varphi = \sum_i h_i v^i$ in terms of fluxes h_i, then the field space proper distance is

$$\Delta\phi = \int_{\varphi_i}^{\varphi_f} (h_i g^{ij} h_j)^{-1/2} d\varphi \,. \tag{5.32}$$

Even though this apparently depends on the values of the fluxes h_i, it was found in [83] that due to a scaling symmetry the moduli are fixed in such a way in terms of the fluxes such that the overall expression (5.32) is in fact flux-independent and the universal behaviour (5.31) sets in with a flux-independent constant α.

In the next and final chapter we will present evidence for a logarithmic behaviour of $\Delta\phi$ in *physical space* for super-Planckian distances [17]. It will be argued that this in combination with the local Weak Gravity Conjecture leads to evidence for the Swampland Conjecture.

Let us finally state the precise form of the refined Swampland Conjecture as proposed in [17].

The Refined Swampland Conjecture

$$m_{\mathrm{SC}}\left(\phi_0 + \Delta\phi\right) = m_{\mathrm{SC}}\left(\phi_0\right)\Gamma\left(\phi_0, \Delta\phi\right) e^{-\alpha\Delta\phi/M_p} \,. \tag{5.33}$$

The mass drop flows to an exponential one quickly after passing $\Delta\phi = M_p$.

More precisely, the a priori arbitrary function $\Gamma(\phi_0, \Delta\phi)$ encodes the local structure of $\mathscr{M}$, while α is the strength of the exponential drop off. The above example suggests that even if $\Gamma \neq 1$, it could in fact be that generally still $\Gamma(\phi_0, \Delta\phi)\exp(-\alpha\Delta\phi/M_p) < 1$ and the decrease in mass continues monotonically with an approximate minimal rate of $\exp(\alpha\Delta\phi/M_p)$, for field displacements $\Delta\phi > \mathscr{O}(1)M_p$ [17]. We saw that this is supported by string theory, where simple examples feature $\Gamma = 1$ and more complicated ones as [83] quickly flow to the SC behaviour for $\Delta\phi \geq M_p$. As was anticipated above, we will try to shed some light on this from a general quantum gravitational perspective in the next chapter, based on our recent paper [17]. A central aim will be to make the $\mathscr{O}(1)$ and also the nature of α more precise.

6 Connecting Conjectures

After proposing refined versions of the (lattice-) Weak Gravity Conjecture and Swampland Conjecture in the last chapter we will now try to establish a connection between them [17]. It is easy to see how such a connection should come about. Both conjectures imply the existence of an infinite tower of states. Since the inclusion of such a tower is quite a big step, one is inclined to think that quantum gravity might prefer the most economical way to satisfy both conjectures — the two towers could be one and the same. In fact, we will present evidence that in the case where the scalar field being displaced controls a gauge coupling the mass scale of the Weak Gravity Conjecture tower of states actually drops exponentially in the displacement. While the Swampland Conjecture was originally motivated in a top-down fashion from string theory examples this provides genuine bottom-up evidence for it. Before getting into technicalities let us outline the way this works. The mass scale of the Weak Gravity Conjecture tower is given by gM_p. In highly supersymmetric setups such as toroidal string compactifications or Calabi-Yau compactifications of type II strings, the gauge coupling is often an exponential function of a (canonically normalised) modulus

$$g(\phi) = e^{-\alpha\phi} \, . \tag{6.1}$$

If this modulus is displaced, the mass scale of the LatWGC tower drops exponentially

$$m_{\text{LatWGC}} \sim e^{-\alpha\Delta\phi} M_p \, , \tag{6.2}$$

implying that it in fact plays the double role mentioned above. It should be kept in mind that the Swampland Conjecture is actually a much more general statement in that it does not rely on the scalar being displaced controlling a gauge coupling. The above situation seems to be very special in that it depends so crucially on the exponential structure of the gauge coupling function. In fact, we will see that this structure is *generic* for super-Planckian displacements, leading to evidence for the refined Swampland Conjecture. The setup under consideration is again that of spatially varying moduli as in the last section. The Weak Gravity Conjecture was originally motivated by studying monopoles and charged black holes in a gauge theory coupled to gravity. As we have seen in the last section, when the gauge coupling is dependent on a scalar field these localised sources induce a flow of the scalar from its value at infinity to a generally different value at the monopole center/black hole horizon. If the scalar fields are moduli, the value at infinity is a free parameter, whereas if there is a potential it is fixed at a minimum of this. For extremal black holes the attractor mechanism (see for example [84–87]) can fix the value of the scalar on the horizon. Since this flow of the scalar field can range over super-Planckian distances in field space it is natural to study the relation between the Swampland Conjecture and the local Weak Gravity Conjecture in this context. Both for the attractor black holes, where the proper distance to the horizon diverges, and for the Swampland Conjecture we reach universal behaviour after travelling a long distance — in the first case in physical space and in the second case in field space. If we want to identify the lattice Weak Gravity Conjecture tower of states with the Swampland Conjecture one, we expect in analogy to the refined Swampland Conjecture (5.33), that the gauge coupling as a function of a modulus ϕ should take the form

$$g(\phi_0 + \Delta\phi) = g(\phi_0)\Gamma(\phi_0, \Delta\phi)e^{-\alpha\Delta\phi/M_p} , \tag{6.3}$$

which then sets the mass scale for the Weak Gravity Conjecture and hence Swampland Conjecture tower. The aim of the following sections is then to provide evidence for the above formula. We first study the case of weakly curved backgrounds and then proceed to the analysis of large spatial displacements in general relativity. The general philosophy is that large scalar field gradients induce strong gravitational backreaction and to avoid the collapse of the whole configuration to a black hole scalars can grow at most logarithmically at super-Planckian distances. This is then related to an exponential structure of the gauge coupling function. Thus we restrict to spherically symmetric solutions which minimize the gravitational backreaction. This has the neat side-effect that these are of course much easier to analyse.

6.1 Weakly Curved Backgrounds

Now that we have seen evidence for a local Weak Gravity Conjecture in section (5.1), we would like to apply it to general static spherically symmetric solutions of the generalised Dilaton-Maxwell system

$$S = \frac{1}{2}\int \left[\star R - 2d\phi \wedge \star d\phi - \frac{1}{g(\phi)^2} F \wedge \star F\right] , \tag{6.4}$$

where $g(\phi)$ is an arbitrary gauge kinetic function. We recover the Dilaton monopole (5.12) in the special case of an exponential gauge coupling (6.1).[1] In this section we restrict to the weak curvature limit of Newtonian gravity (5.13). To evaluate the Newtonian potential in this background we need to calculate the energy density associated to the action (6.4)

$$\rho(\phi, F) = 2(\partial_r\phi)^2 + \frac{1}{g(\phi)^2}\left(\mathbf{B}^2 + \mathbf{E}^2\right) \equiv \rho_\phi + \rho_F . \tag{6.5}$$

This is of course impossible to evaluate directly because of the unknown function $g(\phi)$ as well as the gauge field background. Nevertheless, since the only source for the scalar is the gauge kinetic term we will assume that the induced gradient energy is of order of the energy stored in the gauge fields

$$\rho_\phi \simeq \rho_F \qquad \Rightarrow \qquad \rho(\phi, F) \simeq 4(\partial_r\phi)^2 . \tag{6.6}$$

While we have seen that this is exactly satisfied for the Dilaton monopole (5.21), it is less clear that this should be true for the general case especially because for $g \to$ const, the scalar and gauge field decouple.[2] While we will show later that for large $\Delta\phi$ the energy densities in fact

[1] We adopt a different convention for the normalisation of the gauge kinetic term in this section compared to the Dilaton monopole. The gauge couplings are simply related by a factor of $\sqrt{2}$.

[2] For the Dilaton monopole solution we still have that in the decoupling limit $\alpha \to 0$ the energy densities match, since we have shown them to be equal for arbitrary α. This is because the space of solutions to (5.12) is singular as $\alpha \to 0$. The limit is only regular if $g_\infty = 1$, which means that we lose an integration constant. The proper two-parameter solution for ϕ at $\alpha = 0$ does not necessarily satisfy $\rho_\phi = \rho_F$.

track each other, let us for now simply assume (6.6). While it might seem that this assumption in fact is so strong that we are simply led back to the Dilaton monopole, this is not the case because we allow for arbitrary background gauge fields and thus arbitrary $\phi(r)$. We parametrise our ignorance about the solution for ϕ in terms of two functions γ and $\tilde{\gamma}$ which describe the deviation of the energy density from the Weak Gravity Conjecture bound (5.9) and the deviation of the energy density from the $1/r^2$ scaling of the logarithmic profile (5.22)

$$g(r) \equiv \gamma(r)\sqrt{\rho(r)} \equiv \frac{\gamma(r)\tilde{\gamma}(r)}{r} \, . \tag{6.7}$$

In the case of spatially constant γ and $\tilde{\gamma}$, we are in fact lead back to the Dilaton monopole example. First of all, the energy density can be integrated to give the logarithmic profile of ϕ

$$\sqrt{\rho} \simeq 2\phi' = \frac{\tilde{\gamma}}{r} \, , \qquad \Rightarrow \qquad \phi = \frac{1}{\alpha}\ln(r) \, , \qquad \text{for } \tilde{\gamma} = \text{const} \, . \tag{6.8}$$

Second, for two points $r_{\text{UV}} < r_{\text{IR}}$ we have that

$$\frac{g(r_{\text{IR}})}{g(r_{\text{UV}})} = \frac{g(\phi(r_{\text{UV}}) + \Delta\phi)}{g(\phi(r_{\text{UV}}))} = \frac{r_{\text{UV}}}{r_{\text{IR}}} = e^{-\alpha\Delta\phi} \, , \qquad \text{for } \tilde{\gamma}, \gamma = \text{const} \, . \tag{6.9}$$

Insofar as our aim is to relate the refined Swampland Conjecture and the local Weak Gravity Conjecture, it is of course of great importance to constrain the functional form of γ and $\tilde{\gamma}$. These should flow to approximately constant behaviour once $\Delta\phi$ passes the Planck scale. To constrain these we will use the requirement of weak gravity, i.e. staying in the Newtonian approximation and also the magnetic local Weak Gravity Conjecture. A general result for scalar fields in the Newtonian approximation is that free scalars are bound to sub-Planckian variations, while for parametrically super-Planckian variations they can grow at best logarithmically [88]. Here we briefly present the results. For the free scalar field outside any source (without loss of generality consider $\phi_\infty = 0$),

$$\phi = \phi_F \frac{r_F}{r} \, , \tag{6.10}$$

the contribution from the Newtonian Potential outside the source is

$$\Phi = \int_{r_F}^{\infty} \tilde{r}\phi'^2 d\tilde{r} = \frac{1}{2}\phi_F^2 \ll 1 \, , \qquad \Rightarrow \qquad \Delta\phi = \phi_F \ll 1 \, . \tag{6.11}$$

Furthermore, in [88] it was shown that for fixed and small Newtonian potential the field profile which maximises $\Delta\phi$ is the logarithmic one

$$\phi = \phi_F + \frac{1}{\alpha}\ln(r/r_F) \, . \tag{6.12}$$

Because of (6.11), following [88], we can focus on the region below r_F and pick out the contribution to the Newtonian potential above a given point r_{UV} and below r_F[3]

$$\Delta\Phi \equiv \frac{1}{2}\int_{r_{\text{UV}}}^{r_F} r'\rho(\phi, F; r')dr' \simeq \int_{r_{\text{UV}}}^{r_F} r'\rho_\phi(r')dr' \, . \tag{6.13}$$

[3] Here we assumed that is well described by a logarithm below the radius r_F at which it transitions to a free field. This restriction will be lifted in the following.

For the logarithm, this evaluates to

$$\Delta\Phi = \frac{2\Delta\phi}{\alpha} < 1 \, , \qquad \Rightarrow \qquad \alpha > 2\Delta\phi \, . \tag{6.14}$$

The above restriction on the Dilaton-Maxwell parameter α is not necessary in the presence of strong curvature, as discussed in the next section. Even though the analysis of [88] indicates that for asymptotically large displacements $\Delta\phi \to \infty$, the requirement of weak gravitational backreaction leads to the logarithmic profile of the scalar field, we are interested in the transition regime and would like to show that the profile changes to logarithmic quickly after passing $\Delta\phi = M_p$, in the spirit of the Swampland Conjecture. This amounts to determining $\tilde{\gamma}$ at finite $\Delta\phi$.

6.1.1 Quantitative Behaviour of $\tilde{\gamma}$ at finite $\Delta\phi$

To study $\tilde{\gamma}$ at finite $\Delta\phi$, we look at a simple family of field profiles and show that as $\Delta\phi \gtrsim M_p$, this family indeed flows to the logarithmic case. Since non-monotonic profiles only lead to larger $\Delta\Phi$ at fixed $\Delta\phi$, we restrict to monotonic ones. The most natural family of field profiles to study is the power law one[4]

$$\phi(r) = \phi_{\mathrm{IR}} + \frac{\beta}{\alpha}\left(1 - \left(\frac{r}{r_{\mathrm{IR}}}\right)^{\frac{1}{\beta}}\right) \, . \tag{6.15}$$

We can restrict here to positive $\beta > 0$, since the analysis of the negative powers proceeds completely analogous and set ϕ_{IR} to zero without loss of generality. Since

$$\phi(r) = \frac{\beta}{\alpha}\left(1 - e^{\frac{1}{\beta}\ln\left(\frac{r}{r_{\mathrm{IR}}}\right)}\right) \, , \tag{6.16}$$

one can easily see that

$$\begin{aligned} \phi(r_{\mathrm{UV}}) &= \frac{1}{\alpha}\ln\left(\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}}\right) \times \frac{1}{\delta}\left(e^{\delta} - 1\right) \\ &= \frac{1}{\alpha}\ln\left(\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}}\right) \times \left(1 + \frac{\delta}{2} + \mathscr{O}\left(\delta^2\right)\right) \end{aligned} \, , \qquad \delta \equiv \frac{\ln(r_{\mathrm{UV}}/r_{\mathrm{IR}})}{\beta} \, , \tag{6.17}$$

thus for fixed r_{UV} this converges indeed to the logarithmic profile for $\beta \to \infty$. The field displacement at a given radial coordinate r_{UV} and the Newtonian potential evaluate to[5]

$$\begin{aligned} \Delta\phi &= \frac{\beta}{\alpha}\left(1 - \left(\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}}\right)^{\frac{1}{\beta}}\right) \, , \\ \Delta\Phi &= \frac{\Delta\phi}{\alpha}\left(1 + \left(\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}}\right)^{\frac{1}{\beta}}\right) \, . \end{aligned} \tag{6.18}$$

[4] For large field variations we are interested in the limit $\beta \to \infty$. The constant β dependent renormalisation is only needed for regularity of this limit. Also note that in the following r_{IR} is an arbitrary fixed radial coordinate, in general not related to the free field radius r_F.

[5] From now on we define $\Delta\Phi$ as the integral (6.13) with the integration understood to be from r_{UV} to r_{IR}. This puts a lower bound on the total $\Delta\Phi$.

We do not want to directly take the limit $\beta \to \infty$ but rather would like to consider $\Delta\phi \to \infty$. The two limits are in fact related because

$$1 > \Delta\Phi \geq \frac{\Delta\phi}{\alpha}\left(1-\left(\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}}\right)^{\frac{1}{\beta}}\right) = \frac{\Delta\phi^2}{\beta}\,, \qquad \Rightarrow \qquad \beta > \Delta\phi^2\,, \tag{6.19}$$

so that going to very large field variation in the Newtonian limit necessitates going to even larger β and hence to the logarithmic regime. We clearly also need to have large α

$$1 > \Delta\Phi > \frac{\Delta\phi}{\alpha}\,, \qquad \Rightarrow \qquad \alpha > \Delta\phi\,. \tag{6.20}$$

Now that we have seen that the limit $\Delta\phi \to \infty$ implies logarithmic behaviour, we would like to reassess the crucial assumption of equality of the energy densities of the gauge field and scalar. Taking the r-derivative of the equation of motion for ϕ, using the chain rule to relate r and ϕ derivatives and inserting the given power law profile (6.15) results in

$$\frac{\partial}{\partial r}\left[\frac{1}{g(\phi)^2}\mathbf{B}^2 - 2\frac{\beta+1}{\beta-1}\left(\partial_r\phi\right)^2\right] = 0\,. \tag{6.21}$$

Here we assumed a purely magnetic background and the purely electric case is related by electromagnetic duality. The dyonic case is not as simple but we will later see in an example that a dyonic background only leads to super-Planckian displacements when either the electric or magnetic charge strongly dominates, leading back to the pure electric and magnetic cases. The above can also be written as

$$\frac{\partial}{\partial r}\left(\rho_F - \frac{\beta+1}{\beta-1}\rho_\phi\right) = 0\,, \tag{6.22}$$

which makes it evident that in the case of super-Planckian displacements $\beta \gg 1$ the gauge field energy density tracks the scalar one.

Having this settled, we proceed in analysing the asymptotics of (6.15) for large $\Delta\phi$. While we have seen that in the limit $\Delta\phi \to \infty$ necessarily $\alpha, \beta \to \infty$, there are several ways to take this limit, the important constraint being $\Delta\Phi \ll 1$. Introducting the parameter

$$\varepsilon \equiv \frac{\alpha^2}{2\varepsilon}\,, \tag{6.23}$$

we assume that indeed as in (6.17)

$$\delta \ll 1\,, \tag{6.24}$$

and derive a consistency condition on ε from this. As by the above assumption, we are in the logarithmic regime and hence

$$\Delta\phi \simeq \frac{\beta}{\alpha}\delta\,, \qquad \Delta\Phi \simeq \frac{2\Delta\phi}{\alpha} \simeq \frac{\delta}{\varepsilon}\,. \tag{6.25}$$

This means that staying in the Newtonian approximation is equivalent to

$$\delta \ll \varepsilon\,. \tag{6.26}$$

We can enforce our assumption by sending $\varepsilon \to 0$. Taking the limits $\alpha, \beta \to \infty$ while $\varepsilon \to 0$ we thus approach the logarithmic profile for ϕ. Assuming for now that γ is constant, this leads then to (6.9). At finite ε we find

$$\frac{g(\phi(r_{\mathrm{UV}})+\Delta\phi)}{g(\phi(r_{\mathrm{UV}}))} = \left(1-\frac{2\varepsilon}{\alpha}\Delta\phi\right)^{\frac{\alpha^2}{2\varepsilon}-1} \equiv \Gamma(\phi(r_{\mathrm{UV}}),\Delta\phi)e^{-\alpha\Delta\Phi}\;, \tag{6.27}$$

which evidently converges to the exponential for $\varepsilon \to 0$ because of

$$e^x = \lim_{\varepsilon\to 0}(1+\varepsilon x)^{\frac{1}{\varepsilon}}\;. \tag{6.28}$$

The first thing to note is that for super-Planckian variations, $1 < \Delta\phi < \beta = \alpha^2/2\varepsilon$, and thus the exponent is always positive. Since the term in the parentheses is always smaller than one, the result is

$$\Gamma(\phi(r_{\mathrm{UV}}),\Delta\phi)e^{-\alpha\Delta\Phi} < 1\;, \qquad \text{for } \Delta\phi > 1\;. \tag{6.29}$$

Furthermore, we find numerically that the parameter ε controls the range of $\Delta\phi$ for which $\Gamma \simeq 1$ and that this is approximately the case for $\Delta\phi \ll 1/\sqrt{2\varepsilon}$. For $\Delta\phi > 1.2$, we find that $\Gamma < 1$ and hence

$$g(\phi(r_{\mathrm{UV}})+\Delta\phi) \leq g(\phi(r_{\mathrm{UV}}))e^{-\alpha\Delta\phi}\;, \tag{6.30}$$

even for finite ε. As we have seen above, in the limit $\varepsilon \to 0$ the inequality is saturated.

6.1.2 Quantitative Behaviour of γ at finite $\Delta\phi$

In the last section we assumed a constant $\gamma(r)$ in deriving the exponential drop of the gauge coupling (6.3). Assuming now a constant $\tilde{\gamma}(r)$ we have

$$\frac{g\left(\phi\left(r_{\mathrm{UV}}\right)+\Delta\phi\right)}{g\left(\phi\left(r_{\mathrm{UV}}\right)\right)} = \frac{\gamma(\phi\left(r_{\mathrm{UV}}\right)+\Delta\phi)}{\gamma(\phi\left(r_{\mathrm{UV}}\right))}e^{-\alpha\Delta\phi} \equiv \Gamma(\phi(r_{\mathrm{UV}}),\Delta\phi)e^{-\alpha\Delta\phi}\;. \tag{6.31}$$

Thus we need to constrain the change in γ between r_{UV} and r_{IR}. Half of the job is done by the magnetic local WGC (5.9), which provides us with the lower bound

$$\gamma(r) \geq 1\;. \tag{6.32}$$

The idea is now that we only need to constrain γ at r_{IR}, corresponding to $\phi(r_{\mathrm{UV}})+\Delta\phi$, where the field ϕ has already travelled a long distance in field space an thus should have reached its universal long distance behaviour. Another simplification is that we only need to determine the maximal value of γ. We can first look at the case where $\gamma(r)$ is a monotonic function. We have seen above that free scalars cannot support $\Delta\phi > 1$ so we are only interested in those radii in which the scalar is non-trivially sourced and hence the maximum relevant value of γ occurs at the free field radius r_F. To bound $\gamma(r_F)$ let us first estimate r_F itself. Let us assume that upon reaching r_F we are already at $\Delta\phi > 1$, i.e. $\phi \simeq \frac{1}{\alpha}\ln(r)$. We can furthermore assume that as ϕ is approximately free, we are also at sufficiently large r such that

$$\mathbf{B} \simeq \frac{p}{r^2}\;, \qquad p = \int_{r\leq r_F}\rho_m\;, \tag{6.33}$$

where ρ_m is the magnetic charge density and we assume absence of electric charge. The equation of motion (5.17) evaluated on this background and at $r = r_F$ reads

$$r_F^2 = -\alpha p^2 \frac{\partial_\phi \ln(g)}{2g^2}\bigg|_{r\to r_F} . \tag{6.34}$$

The energy density of the total configuration is bounded from below by the electromagnetic one

$$\rho \geq \rho_F = \frac{p^2}{g^2 r^4} . \tag{6.35}$$

From this we get that

$$\gamma(r_F) = \frac{g}{\sqrt{\rho}}\bigg|_{r\to r_F} \lesssim \frac{g}{\sqrt{\rho_F}}\bigg|_{r\to r_F} \simeq -\alpha p \partial_\phi \ln(g)\big|_{r\to r_F} . \tag{6.36}$$

Since the Γ-factor is defined as a *ratio* of the γ-factors and the charge p should drop out of the ratio of gauge couplings, we have also[6]

$$\Gamma(\phi(r_{\mathrm{UV}}), \Delta\phi) \lesssim -\alpha \partial_\phi \ln(g)\big|_{r\to r_F} . \tag{6.37}$$

We thus find that the maximum value of Γ is dependent on the functional dependence of the gauge coupling on the scalar. Assuming that Γ is already approximately constant at $r = r_F$, the gauge coupling is approximately of the exponential form $g \sim \exp(-\alpha\phi)$, so we can estimate

$$\Gamma(\phi(r_{\mathrm{UV}}), \Delta\phi) \lesssim \alpha^2 . \tag{6.38}$$

If it is not the case that Γ is approximately constant at $r = r_F$, we have to account for this variation of Γ, which we assumed to be monotonically increasing. By $g = \gamma\sqrt{\rho}$, the radial increase of γ counteracts the decrease of the energy density, leading to smaller rate of change of g with respect to ϕ. This leads to a stronger bound on the actual maximum value of Γ via (6.37). To get some intuition, suppose that γ varies as some power law

$$\gamma(r) \sim r^{\frac{\alpha-\delta}{\alpha}} . \tag{6.39}$$

Then we have that

$$g \sim e^{-\delta\phi} , \tag{6.40}$$

and the bound on Γ modifies to

$$\Gamma(\phi(r_{\mathrm{UV}}), \Delta\phi) \lesssim \alpha\delta . \tag{6.41}$$

While this is technically a weaker bound in comparison to (6.38) for $\delta > \alpha$, this is compensated by the gauge coupling having dropped exponentially faster by a factor of

$$e^{-(\delta-\alpha)\Delta\phi} , \tag{6.42}$$

[6] This is true only in the case where r_{UV} is outside of any sources for F, so that the integrated charge densities at r_F and r_{UV} coincide. If this is not the case, the bound can be weaker.

which is very small since we assumed to be already in the super-Planckian regime at r_F.

Note that while the above analysis assumed monotonicity of the γ-factor from r_{IR} to the free field radius r_F, we can drop this assumption if we restrict to the special case $r_{\mathrm{IR}} = r_F$. In using the magnetic Weak Gravity Conjecture to constrain the γ-factor contribution from the UV point r_{UV}, we have been extremely conservative. In fact, while from (6.38) it might seem that generically Γ is tunable by adjusting the prefactor α of the logarithmic dependence of ϕ on r, in the case of the Dilaton monopole of section 5.1 we have indeed that $\gamma \sim \alpha^2$ but since this is an overall factor it drops out in the ratio $\gamma(r_{\mathrm{UV}})/\gamma(r_{\mathrm{IR}})$. We find that the ratio is not tunable to be bigger than four, since γ is indeed monotonic in this case and by equation (5.27)

$$\frac{\gamma(r_{\mathrm{IR}})}{\gamma(r_{\mathrm{UV}})} \leq \frac{\gamma(r_F)}{\gamma(r_N)} = \frac{4}{\left(1+\frac{r_N}{r_F}\right)^2} \leq 4\,. \tag{6.43}$$

The above general arguments suggest that as $\Delta\phi \gg 1$, where we are in the logarithmic regime of the scalar and the only contribution to Γ is indeed the ratio of γ since $\tilde{\gamma}$ asymptotes to a constant, Γ is indeed subdominant to the exponential drop in $\Delta\phi$. This, together with the analysis of the $\tilde{\gamma}$-factor in section 6.1.1 presents evidence for the exponential dependence of the gauge coupling function on super-Planckian scalar field displacements (6.3). As we have emphasised in the introduction to this chapter, this then together with the electric local Weak Gravity Conjecture leads to the first purely bottom-up evidence for the Swampland Conjecture, completely independent of string theory.

6.1.3 Some Comments

A few comments are in order, justifying several assumptions made in this chapter.

First of all, to get to the region where $\Delta\phi \gg 1$, we had to go to the limit of large α. In simple string compactifications the parameter α is usually a constant of order one, $\alpha = \mathscr{O}(1)$. This does not imply that super-Planckian displacements are impossible in this case. In fact, we will see in the next section that the requirement of large α is an artifact of staying in a weakly curved background (6.20). If we lift this restriction we can easily have super-Planckian $\Delta\phi$ for any α.

We have emphasised that it is crucial to stay in the regime $\Phi \ll 1$ if we neglect gravitational backreaction. To be careful we should also restrict to super-Planckian radii,

$$1 \overset{!}{\gg} \frac{1}{r_N} \simeq \frac{g_\infty}{\alpha p} e^{-\frac{\alpha^2}{2}}\,. \tag{6.44}$$

A similar constraint applies to the gauge physics. We should demand that the gauge coupling stays perturbative when evaluated on the solution. For example in the Dilaton monopole this is monotonically increasing towards the monopole centre and we should demand

$$g(r_N) = g_\infty\left(1+\frac{r_F}{r_N}\right) \simeq g_\infty\left(1+e^{\frac{\alpha^2}{2}}\right) \overset{!}{<} 1\,. \tag{6.45}$$

In equations (6.44,6.45) we assumed the large α required for $\Delta\phi > 1$. Clearly, both conditions require exponentially weak gauge coupling at infinity

$$g_\infty < e^{-\frac{\alpha^2}{2}}\,. \tag{6.46}$$

It is also intuitively clear that in the general setup this should be a sufficient condition, not only for the special case of the Dilaton monopole, since we look at a limit where g increases approximately as

$$g(r_N) \simeq e^{\alpha\Delta\phi} g(r_F) \simeq e^{\alpha\Delta\phi} g_\infty < e^{\frac{\alpha^2}{2}} g_\infty \,, \tag{6.47}$$

where we have used the linear upper bound on $\Delta\phi$ in terms of α (6.20).

We conclude this section by remarking that while we have restricted to purely magnetic charge densities, the results generalise to purely electric ones by electromagnetic duality. While the magnetic sources tend to drive the gauge coupling to large values when approached from infinity, the duality transformation includes an inversion of the gauge coupling $g \to 1/g$, thus electric sources drive it to small values. The result is for example in the Dilaton monopole that the scalar is driven in the opposite direction in field space. Since we are interested in large $\Delta\phi$, a competition between electric and magnetic sources is not desired. In fact, we will see in the next section that an explicit dyonic configuration only leads to super-Planckian $\Delta\phi$ if one type of charge effectively dominates.

6.2 Strongly Curved Backgrounds

After discussing the problem of $\Delta\phi > 1$ in weakly curved backgrounds in great detail, we will now remove this restriction and consider arbitrarily curved backgrounds. The price we have to pay is that we have to use a version of the magnetic local Weak Gravity Conjecture for strong curvature (5.11). We would like to study again the Dilaton-Maxwell system (6.4) with an arbitrary gauge coupling function. In the weakly curved case we eliminated the unknown functional form of $g(\phi)$ by the physical assumption $\rho_F \simeq \rho_\phi$. This assumption will be replaced in the strongly curved case by considering a curvature invariant for application in (5.11) that is not sensitive to the gauge kinetic term. The most simple curvature invariant, R, is appropriate for this since it is only sensitive to the trace of T_{ab} and the electromagnetic energy momentum tensor is traceless. A more complete treatment would have to use an invariant which is also sensitive to the traceless contributions in T_{ab} such as the square of the Ricci tensor or the Kretschmann scalar. Nevertheless, R is a lower bound to these more complete measures, so we still expect

$$\sqrt{R(r)} < g(r) M_p \,. \tag{6.48}$$

Let us first look at an example that will be analogous to the Dilaton monopole discussed in section 5.1.

6.2.1 The Dilaton Black Hole

We will first look at an explicit strongly curved dyonic black hole solution, which can be implemented into a concrete string theory setup. It is a particular limit of the STU black hole, which can be obtained for instance by a toroidal compactification of Type IIA string theory to four dimensions. The WGC states in this setup are given by wrapped branes and are BPS. The details of this construction are discussed in appendix E. The theory under consideration is

$$S = \frac{1}{2\kappa^2} \int \left(\star R - 2d\phi \wedge \star d\phi - e^{2\phi} F_1 \wedge \star F_1 - e^{-2\phi} F_2 \wedge \star F_2 \right) \,, \tag{6.49}$$

which is the bosonic action of $\mathscr{N}=2$ SUGRA coupled to a single vector multiplet, which can also be obtained by a truncation of pure $\mathscr{N}=4$ SUGRA. The solutions have been described in the context of $\mathscr{N}=2$ supergravity in [54] and $\mathscr{N}=4$ supergravity in [89]. Here we will be mainly interested in extremal solutions due to their simplicity. The general non-extremal solution can be found in [89]. For the extremal solutions, we use the metric ansatz

$$ds^2 = -e^{2U(r)}dt^2 + e^{-2U(r)}\left(dr^2 + r^2 d\Omega_2^2\right) \ . \tag{6.50}$$

In the notation of appendix C the general spherically symmetric solution of the Maxwell equations in such a background is given by

$$\begin{aligned} F_1 &= \frac{Q_1 e^{-2\phi}}{b^2} e^{\hat{t}} \wedge e^{\hat{r}} + \frac{P_1}{b^2} e^{\hat{\theta}} \wedge e^{\hat{\phi}} \ , \\ F_1 &= \frac{Q_2 e^{+2\phi}}{b^2} e^{\hat{t}} \wedge e^{\hat{r}} + \frac{P_2}{b^2} e^{\hat{\theta}} \wedge e^{\hat{\phi}} \ . \end{aligned} \tag{6.51}$$

Observers at asymptotic infinity will canonically normalise their gauge fields and measure the dilated charges as

$$\begin{aligned} \begin{pmatrix} \tilde{Q}_1 \\ \tilde{P}_1 \end{pmatrix} &:= \begin{pmatrix} Q_1 \\ P_1 \end{pmatrix}\Bigg|_\infty = \frac{1}{4\pi} \int_{S^2_\infty} e^{+\phi_\infty} \begin{pmatrix} \star F_1 \\ F_1 \end{pmatrix} = \begin{pmatrix} Q_1 e^{-\phi_\infty} \\ P_1 e^{+\phi_\infty} \end{pmatrix} \ , \\ \begin{pmatrix} \tilde{Q}_2 \\ \tilde{P}_2 \end{pmatrix} &:= \begin{pmatrix} Q_2 \\ P_2 \end{pmatrix}\Bigg|_\infty = \frac{1}{4\pi} \int_{S^2_\infty} e^{-\phi_\infty} \begin{pmatrix} \star F_2 \\ F_2 \end{pmatrix} = \begin{pmatrix} Q_2 e^{+\phi_\infty} \\ P_2 e^{-\phi_\infty} \end{pmatrix} \ . \end{aligned} \tag{6.52}$$

We will group those charges that dilate with $\exp(+\phi_\infty)$ and those which dilate with $\exp(-\phi_\infty)$ together as

$$\begin{aligned} \mathscr{Q}_+ &= \sqrt{P_1^2 + Q_2^2} \ , \\ \mathscr{Q}_- &= \sqrt{Q_1^2 + P_2^2} \ , \end{aligned} \tag{6.53}$$

as only these combinations will appear in the solution. From the action (6.49) we can read off the trace reversed energy momentum tensor in terms of the orthonormal co-frame defined in appendix C. Here we work with the trace reversed Einstein equations because the energy momentum tensor of the gauge fields is invariant under trace reversal, while the one for the Dilaton looks much simpler in the trace reversed form. In terms of the "effective potential" for ϕ in presence of the non-trivial gauge field background as defined in [85],

$$V_{\text{eff}}(\phi) = \frac{1}{2}\left(e^{2\phi}\mathscr{Q}_+^2 + e^{-2\phi}\mathscr{Q}_-^2\right) \ , \tag{6.54}$$

we have

$$\begin{aligned} \tilde{T}^{\hat{t}}_{\hat{t}} &= -\tilde{T}^{i}_{j} = -\frac{V_{\text{eff}}}{r^4} e^{4U} \ , \\ \tilde{T}^{\hat{r}}_{\hat{r}} &= -\frac{V_{\text{eff}}}{r^4} e^{4U} + 2\phi'^2 e^{2U} \ , \end{aligned} \tag{6.55}$$

where $\tilde{T}^a_b = \tilde{T}^a_b - \delta^a_b T/2$ is the trace reversed energy momentum tensor.

We take the Ricci tensor from (C.14) and find that the Einstein equations and Dilaton equation of motion result in the following set of equations

$$\begin{aligned} \phi'^2 + U'^2 = \nabla^2 U &= \frac{V_{\text{eff}}}{r^4} e^{2U} \,, \\ \nabla^2 \phi &= \frac{\partial_\phi V_{\text{eff}}}{2r^4} e^{2U} \,. \end{aligned} \tag{6.56}$$

Note that these equations enjoy a symmetry under combined dilations of the charges and shifts of the Dilaton

$$\mathscr{Q}_+ \to e^{\delta} \mathscr{Q}_+ \,, \qquad \mathscr{Q}_- \to e^{-\delta} \mathscr{Q}_- \,, \qquad \phi \to \phi - \delta \,, \tag{6.57}$$

which will ensure that the solution for the metric only contains the invariant charge combinations $\tilde{\mathscr{Q}}_+ = e^{\phi_\infty} \mathscr{Q}_+$ and $\tilde{\mathscr{Q}}_- = e^{-\phi_\infty} \mathscr{Q}_-$. This system is solved by the ansatz

$$U = -\frac{1}{2} \ln \left(H_1 H_2 \right) \,, \qquad \phi = -\frac{1}{2} \ln \left(\frac{H_1}{H_2} \right) \,, \tag{6.58}$$

where the H_i are harmonic functions of r

$$H_1 = e^{-\phi_\infty} + \frac{\mathscr{Q}_+}{r} \,, \qquad H_2 = e^{+\phi_\infty} + \frac{\mathscr{Q}_-}{r} \,. \tag{6.59}$$

We see that the value of the Dilaton at the horizon is fixed by the ratio of the magnetic charges of the gauge fields and independent of the asymptotic value of the Dilaton. This is the attractor mechanism that was briefly discussed at the beginning of this chapter.

Note that the solution (6.58) displays the same logarithmic behaviour of the scalar field as the Dilaton monopole

$$\phi = \phi_\infty + \frac{1}{2} \ln \left(\frac{r + \tilde{\mathscr{Q}}_-}{r + \tilde{\mathscr{Q}}_+} \right) \,. \tag{6.60}$$

There is also a one parameter deformation of this black hole, which depends on a parameter α analogous to the one in the Dilaton monopole, if we restrict to purely magnetic charges (P_1, P_2). This is a solution to the deformed action

$$S = \frac{1}{2\kappa^2} \int \left(\star R - 2 d\phi \wedge \star d\phi - e^{2\alpha\phi} F_1 \wedge \star F_1 - e^{-2\phi/\alpha} F_2 \wedge \star F_2 \right) \,. \tag{6.61}$$

The solution is explicitly given by[7]

$$\begin{aligned} U &= -\frac{\alpha}{1+\alpha^2} \ln \left(H_1^{\frac{1}{\alpha}} H_2^{\alpha} \right) - \tfrac{1}{2} \ln (\spadesuit) \,, \\ \phi &= -\frac{\alpha}{1+\alpha^2} \ln \left(\alpha \frac{H_1}{H_2} \right) \,. \end{aligned} \tag{6.62}$$

[7] This is basically the solution of [85] but the author has found a disagreement regarding the integration constants.

with two harmonic functions given by

$$H_1 = \frac{P_1}{r} + \frac{\sqrt{2}}{\sqrt{1+\alpha^2} e^{\alpha\phi_\infty}} ,$$
$$H_2 = \frac{P_2}{r} + \frac{\alpha\sqrt{2}}{\sqrt{1+\alpha^2} e^{-\phi_\infty/\alpha}} . \tag{6.63}$$

The α-dependent constant $\spadesuit$ is given by

$$\spadesuit = \tfrac{1}{2}\left[\left(\frac{1}{\alpha^2}\right)^{\frac{\alpha^2}{1+\alpha^2}} + \left(\frac{1}{\alpha^2}\right)^{\frac{-1}{1+\alpha^2}}\right] . \tag{6.64}$$

The constant but α-dependent shifts on U and ϕ are irrelevant for the equations of motion but convenient because they ensure that the boundary conditions $\phi \to \phi_\infty$ and $\exp(2U) \to 1$ are met at infinity. The same is of course true for the constant terms in the harmonic functions. The solution has a horizon at $r = 0$, where we cut it off. We can usefully rewrite the solution for the scalar ϕ in terms of two distance scales $\ell_{1,2}$ as

$$\phi = \phi_\infty - \frac{1}{\alpha + 1/\alpha} \ln\left(\frac{r+\ell_1}{r+\ell_2}\right) , \qquad \ell_1 = \frac{\sqrt{1+\alpha^2}\tilde{P}_1}{\sqrt{2}} , \qquad \ell_2 = \frac{\sqrt{1+1/\alpha^2}\tilde{P}_2}{\sqrt{2}} , \tag{6.65}$$

where we defined the charges measured at infinity as $\tilde{P}_1 = \exp(\alpha\phi_\infty)P_1$ and $\exp(-\phi_\infty/\alpha)P_2$.

In the case of $P_2 = 0$, we can see that this reduces to the Dilaton monopole for large α by rewriting it as

$$\phi = \phi_\infty - \frac{\alpha}{1+\alpha^2} \ln\left(1 + \frac{P_1 e^{\alpha\phi_\infty}\sqrt{1+\alpha^2}}{\sqrt{2}r}\right) \xrightarrow{\alpha\gg 1} \phi_\infty - \frac{1}{\alpha}\ln\left(1 + \frac{P_1 \alpha e^{\alpha\phi_\infty}}{\sqrt{2}r}\right) . \tag{6.66}$$

This coincides with the Dilaton monopole up to a factor of $\sqrt{2}$ from the gauge field normalisation. We had to go to the limit of large α since it is the parameter controlling the separation of r_N and r_F, so for a large region where the Newtonian approximation is valid we are necessarily in the large α limit[8].

The solution has an uninteresting limit of $\ell_1 = \ell_2$, in which both contributions to the running of the Dilaton cancel out exactly and it stays constant[9]. Apart from that, in order to analyse the solution let us assume without loss of generality that $\ell_1 > \ell_2$, so that ϕ is monotonically increasing towards infinity and the gauge coupling corresponding to F_1 in turn is decreasing, just as in the Dilaton monopole. The case of $\ell_2 > \ell_1$ is precisely the same except that the scalar runs in the opposite direction. The ℓ_i provide a natural way to separate the solution into three

[8] In a UV completion such as string theory α is not an arbitrary parameter. In fact we will see in appendix E that the possibility of obtaining large α in string theory might be constrained.

[9] While we write this solution as a purely magnetic one, there is always one gauge coupling that will be strong at any point r. We can go to a different electromagnetic duality frame where F_2 is electric and this leads to a dyonic solution with both gauge couplings weak at infinity. The cancellation in the running is then the one anticipated for dyonic sources with comparable electric and magnetic sources.

different regions. We can use $\ln(1+\varepsilon) \approx \varepsilon$ for small ε to investigate the behaviour of ϕ in these regions. In the first one, $r < \ell_i$ we have that the scalar approximately behaves linearly in r. In the second one, $\ell_2 < r < \ell_1$, ϕ grows logarithmically, and finally in the third region $r > \ell_i$, it asymptotes to its value at infinity like a free field

$$\phi \simeq \begin{cases} \ln\left(\frac{P_2}{P_1}\right) + \frac{1}{\alpha+1/\alpha}\left(\frac{1}{\ell_2} - \frac{1}{\ell_1}\right) r & r \ll \ell_2 \ll \ell_1 \\ \phi_\infty + \frac{1}{\alpha+1/\alpha} \ln\left(\frac{\ell_1}{r}\right) & \ell_2 \ll r \ll \ell_1 \\ \phi_\infty - \frac{1}{\alpha+1/\alpha}(\ell_1 - \ell_2)\frac{1}{r} & \ell_2 \ll \ell_1 \ll r \end{cases} . \tag{6.67}$$

We can compute the maximal field variation in these regions exactly. The result is

$$\Delta\phi = \begin{cases} \frac{1}{\alpha+1/\alpha} \ln\left(\frac{2\ell_1}{\ell_1+\ell_2}\right) \leq \frac{1}{2}\ln(2) & r < \ell_1 \text{ or } r > \ell_i \\ \frac{1}{\alpha+1/\alpha} \ln\left(\frac{(\ell_1+\ell_2)^2}{4\ell_1\ell_2}\right) < \frac{1}{\alpha+1/\alpha} \ln\left(\frac{\ell_1}{\ell_2}\right) & \ell_2 < r < \ell_1 \end{cases} . \tag{6.68}$$

As a result we see that in the regions $0 < r < \ell_2$ and $r_1 < r < \infty$, the field displacement is always sub-Planckian. The inequality in (6.68) is saturated for the case $\alpha = 1$ and $P_2 = 0$, corresponding to the Dilaton monopole. We see again that the field displacement is maximised if electric and magnetic sources do not compete. As was already stated in the discussion of the weakly curved case, super-Planckian $\Delta\phi$ are possible in the region $\ell_2 < r < \ell_1$ even for very small α, since we can tune the argument of the logarithm in (6.68) by going to large charge ratios. The important point is that the result is evidence for the logarithmic growth of scalars, corresponding to $\tilde{\gamma} \simeq$ const, at super-Planckian distances since in the solution the logarithmic growth sets in before we pass $\Delta\phi = 1$.

6.2.2 The General Case

We would now like to discuss the general case (6.4) with an arbitrary gauge coupling function. Here we restrict again to spherically symmetric solutions. We will introduce a γ-factor analogous to the weakly curved case (6.7)

$$g(r) = \gamma_R(r)\sqrt{R(r)} , \qquad \gamma_R(r) > 1 , \tag{6.69}$$

but there will be no $\tilde{\gamma}$-factor, since we will determine the profile of ϕ directly by solving the trace of the Einstein equations. The aim will be again to show that $\gamma \simeq$ const for $\Delta\phi > 1$. Note again, as was mentioned in the introduction to this section, the Ricci scalar R is not sensitive to the gauge field contribution to the energy momentum tensor and γ_R would be analogous to defining γ with respect to the scalar field gradient energy density only in the weakly curved case, whereas what we actually did was using the assumption $\rho_F \simeq \rho_\phi$. By spherical symmetry, we can use the following most general static ansatz for the metric

$$g = -e^{2U(r)}dt^2 + e^{-2U(r)}\left(dr^2 + f(r)r^2 d\Omega^2\right) . \tag{6.70}$$

We want to extract some information independent of the gauge coupling function $g(\phi)$. The relevant part of the Einstein field equations is the trace of it, which is ignorant of the electromagnetic sector. Using formulae from appendix C, one can readily check that the Ricci scalar takes the form

$$R = 2e^{2U}\left[\nabla^2 U - U'^2 - \frac{\nabla^2 f}{f} + \frac{1}{4}\left(\frac{f'}{f}\right)^2 + \frac{1}{r^2}\left(\frac{1}{f} - 1\right) + \left(U' - \frac{1}{r}\right)\frac{f'}{f}\right], \tag{6.71}$$

where

$$\nabla^2 = \partial_r^2 + \frac{2}{r}\partial_r \tag{6.72}$$

is the $3d$ flat space Laplace operator applied to isotropic functions and the trace of the Einstein equations reads

$$\phi'^2 = -U'^2 + \nabla^2 U - \frac{\nabla^2 f}{f} + \frac{1}{4}\left(\frac{f'}{f}\right) + \frac{1}{r^2}\left(\frac{1}{f} - 1\right) + \left(U' - \frac{1}{r}\right)\frac{f'}{f}\,. \tag{6.73}$$

To simplify this we parametrise the two unknown functions ϕ, U by introducing two arbitrary functions H_1 and H_2 such that

$$\begin{aligned} U &= -\frac{1}{\alpha + 1/\alpha}\ln\left(H_1^{\frac{1}{\alpha}} H_2^{\alpha}\right) + \frac{1}{2}\ln(f)\,, \\ \phi &= -\frac{1}{\alpha + 1/\alpha}\ln\left(\frac{H_1}{H_2}\right)\,. \end{aligned} \tag{6.74}$$

Using this, the trace of the Einstein equations simplifies tremendously and we get

$$\frac{2}{\alpha}\frac{\nabla^2 H_1}{H_1} + 2\alpha\frac{\nabla^2 H_2}{H_2} + (\alpha + 1/\alpha)\,\frac{\nabla^2(rf) - 2/r}{rf} = 0\,. \tag{6.75}$$

Since this is a single second order differential equation for three functions, we obviously will not get a unique solution, but we will see that there are some interesting special cases. The form of (6.75) suggests that we might look for solutions where

$$\frac{\nabla^2 H_i}{H_i} = \lambda_i\,, \tag{6.76}$$

meaning that the H_i are eigenfunctions of the flat space Laplacian. In this case the remaining equation for f is

$$\frac{\nabla^2(rf) - \frac{2}{r}}{rf} = \lambda_3\,, \tag{6.77}$$

with λ_3 given by

$$\frac{1}{\alpha}\lambda_1 + \alpha\lambda_2 + \frac{1}{2}\left(\alpha + \frac{1}{\alpha}\right)\lambda_3 = 0\,. \tag{6.78}$$

If $\lambda_3 = 0$, the general solution for f is given by[10]

$$1 + \frac{A}{r} + \frac{B}{r^2}\,. \tag{6.79}$$

[10] The non-extremal generalisation of the extremal Dilaton black hole (6.58) fits into this case.

For non-zero λ_3, shifting rf by a harmonic function, $H_3 \equiv rf + 2/\lambda_3 r$, we find that H_3 must be an eigenfunction of the Laplacian as well

$$\frac{\nabla^2 H_3}{H_3} = \lambda_3 \,, \qquad \Rightarrow \qquad f = \frac{H_3}{r} - \frac{2}{\lambda_3 r^2} \,. \tag{6.80}$$

Note that eigenfunctions of the 3d isotropic Laplacian are the spherical Bessel functions of order 0, hence

$$H_i = c_j j_0\left(\sqrt{-\lambda_i}r\right) + c_y y_0\left(\sqrt{-\lambda_i}r\right) \,, \qquad j_0(r) = \frac{\sin(r)}{r} \,, \qquad y_0(r) = -\frac{\cos(r)}{r} \,. \tag{6.81}$$

The eigenvalues $-\lambda_i$ are precisely the momentum squared of the solution. For real momentum, or negatve λ_i, we get oscillating solutions while for imaginary momentum, or positive λ_i, we get solutions which decay exponentially. We also have the zero modes of the Laplacian, the harmonic functions, as solutions

$$H_i = a_i\left(1 + \frac{\ell_i}{r}\right) \,. \tag{6.82}$$

The most simple solution is the one where $\lambda_i = 0$, hence H_1 and H_2 are harmonic, and we take f to be a constant, $f = 1$. We see that the solution for the scalar field takes the same form as in the Dilaton black hole, except that we have no formula for the ℓ_i in terms of the black hole charges, so the exact same analysis for the logarithmic behaviour of ϕ applies.

Let us check the general implications of the local Weak Gravity Conjecture (5.10,5.11) in this context. Here we take without loss of generality $\ell_1 > \ell_2$ and restrict to the most interesting logarithmic region $\ell_2 \ll r \ll \ell_1$. There we have

$$\frac{r_{\mathrm{UV}}}{r_{\mathrm{IR}}} \leq e^{-(\alpha + \frac{1}{\alpha})\Delta\phi} \,. \tag{6.83}$$

We want to relate this to the fall-off of the gauge coupling via the r-scaling of the Ricci scalar. By the trace of the Einstein equations we have on-shell

$$R = 2e^{2U}\phi'^2 = 2H_1^{\frac{-2}{1+\alpha^2}} H_2^{\frac{-2}{1+1/\alpha^2}} \phi'^2 \,. \tag{6.84}$$

In the logarithmic regime we have that

$$\phi' \sim \frac{1}{r} \,, \qquad H_1 \sim r^{\frac{2}{1+\alpha^2}} \,, \qquad H_2 \simeq \text{const} \,, \tag{6.85}$$

from which we extract the r scaling of the Ricci scalar

$$\sqrt{R} \sim r^{\frac{-\alpha^2}{1+\alpha^2}} \,. \tag{6.86}$$

This means that the magnetic Weak Gravity Conjecture constraints translates to

$$g\left(\phi(r_{\mathrm{UV}}) + \Delta\phi\right) \leq g\left(\phi(r_{\mathrm{UV}})\right) \frac{\gamma_R\left(\phi(r_{\mathrm{UV}}) + \Delta\phi\right)}{\gamma_R\left(\phi(r_{\mathrm{UV}})\right)} e^{-\alpha\Delta\phi} \,. \tag{6.87}$$

As in the weakly curved case, we need to bound γ_R at r_{IR} in order to arrive at the Swampland Conjecture statement (6.9). The argument proceeds just as in the weakly curved background since at the free field radius we are already in a weakly curved region. If we assume that before reaching r_F we already have $\Delta\phi > 1$, we will also have approximate equality of the energy densities of the gauge field and scalar, so we can relate the bound on $\gamma(r_F)$ to a bound on γ_R.

For improved clarity, we have restricted to the case where we are deeply in the logarithmic regime. In fact, we find that (6.87) is true also for general $r_{\mathrm{UV}} < r_{\mathrm{IR}}$. Even though the ratio of $R(r_{\mathrm{IR}})/R(r_{\mathrm{UV}})$ can be increased outside the deep logarithmic region, we still find that $\gamma_R(r_{\mathrm{UV}})$ also increases and cancels the increase from the ratio of R.

After discussing this case, which basically led back to the Dilaton black hole, we would like to look for more general solutions in terms of eigenfunctions of the Laplacian. The key equation here is (6.78). It tells us that not all λ_i can be positive or negative simultaneously.

Consider first the special case where we still fix $f = 1$. In this case the relation between the eigenvalues of H_1 and H_2 is $\lambda_2 = -\lambda_1/\alpha^2$, so we get one exponential and one oscillatory function in the ratio H_2/H_1 that determines ϕ

$$\phi \sim \frac{1}{\alpha + 1/\alpha} \ln\left(\frac{e^{-\alpha\sqrt{\lambda_2}r}}{\sin(\sqrt{\lambda_2}r)}\right) = \frac{-1}{\alpha+1/\alpha}\left(\sqrt{\lambda_2}r + \ln\sin\left(\sqrt{\lambda_2}r\right)\right) . \tag{6.88}$$

Here we may assume that $\lambda_2 > 0$. Even though the logarithm is cancelled by the exponential, leading to a linear dependence on r, this behaviour can only hold for a finite range of r since we have poles at the zeros of the sine. The variation from the linear term in (6.88) is actually bounded by $\Delta\phi_1 \leq \pi/(1+1/\alpha^2) \leq \pi$. Super-Planckian variations are still possible near the poles of the sine due to the second term. Close to the poles, the sine is well approximated by a linear function and we get back to the logarithmic behaviour for super-Planckian $\Delta\phi$. We have seen before that this logarithmic growth then leads to the exponential behaviour of g (6.9).

The most general case we will consider here is the one where all three λ_i are allowed to be non-zero. Since one of the three functions must be necessarily of sine or cosine type, we will again have to cut off the solution at the induced poles. One interesting case is where $\lambda_1, \lambda_2 < 0$, so that $\lambda_3 > 0$. Now the metric has poles at $r = 0, \pi$, while the scalar is a purely linear function of r

$$\phi \sim \frac{\alpha}{1+\alpha^2}\left(\sqrt{|\lambda_1|} \pm \sqrt{|\lambda_2|}\right) r . \tag{6.89}$$

The total variation is bounded by

$$\Delta\phi \leq \frac{\alpha}{1+\alpha^2}\left(\sqrt{|\lambda_1|} + \sqrt{|\lambda_2|}\right)\frac{\pi}{\sqrt{|\lambda_3|}} = \sqrt{\frac{\alpha}{1+\alpha^2}}\frac{\sqrt{|\lambda_1|}+\sqrt{|\lambda_2|}}{\sqrt{\frac{1}{\alpha}|\lambda_1| + \alpha|\lambda_2|}}\frac{\pi}{\sqrt{2}} \leq \frac{\pi}{\sqrt{2}} . \tag{6.90}$$

The other interesting cases are those in which ϕ itself has poles. Here we observe the same logarithmic behaviour close to the poles as for $f = 1$. One might worry that by tuning α it could be possible to induce a parametric separation between the λ_i but using (6.78) we find that this can be never larger than α^2. One thing that changes between the different possibilities for the signs of the λ_i is the behaviour of the Ricci scalar. We find that the exponent in (6.9) changes depending on this choice and the pole that is being approached, but there is a maximal difference of a factor of two.

While the above is quite nontrivial evidence for the Swampland Conjecture behaviour (6.9) in the strongly curved case, there are of course other solutions to (6.75) than the simple eigenfunctions of the Laplacian. We leave a more general analysis for future work.

7 Conclusion

In this thesis we have investigated the possibility of large scalar field displacements in quantum gravity. As a main motivation for studying these, we introduced the theory of cosmological inflation. We have focused on a large class of well motivated so-called large field inflation models, where the inflaton is a scalar axion with a discrete gauged shift symmetry. Furthermore, we have seen that the implementation of these into a UV complete quantum gravity theory such as string theory seems to be problematic. The maximum field range of these axions usable for inflation is given in terms of the axion decay constant f. It was shown that the required $f > M_p$ is difficult, if not impossible, to achieve in string theory.

While this could be simply a property of string theory, we have proceeded to highlight that this is in fact a general property of consistent quantum gravity theories, in the guise of the Weak Gravity Conjecture. The Weak Gravity Conjecture only constrains models of inflation where the inflaton is an axion. This led us then to the Swampland Conjecture of Ooguri and Vafa, which constrains super-Planckian scalar field displacements in general, since it implies the exponential decrease in mass of a tower of quantum gravity related states with the scalar field displacement at super-Planckian distances. The Swampland Conjecture is motivated by examples from string theory. Thus, if this conjecture is true, effective field theories descending from string theory must break down for scalar field displacements parametrically larger than M_p and this poses a direct obstruction for any large field inflation model, not only the axionic ones. Several other conjectures were discussed which are connected to the Swampland and Weak Gravity conjectures. Although the Weak Gravity Conjecture has been motivated by bottom-up arguments and has been successfully used to constrain a large class of axion inflation models, the Swampland Conjecture is much more general and thus it is very desirable to gather some evidence for it not only from a top-down string theory perspective but also from general quantum gravitational arguments.

This is in fact what we tried to establish in the rest of the thesis, based on our recent paper [17]. We have shown that the Swampland Conjecture might be implied by the lattice version of the Weak Gravity Conjecture if indeed the scalar that is being displaced controls a gauge coupling. This is true if the gauge coupling is an exponential function of the scalar and we have indeed shown evidence that it must flow towards this behaviour for super-Planckian displacements. In order to do this, we have investigated super-Planckian spatial scalar field displacements. These are constrained in a gravitational theory because very large gradients can induce such a large energy density that the system collapses to a black hole. We have seen that avoiding such collapse implies that scalars can at most grow logarithmically with the spatial displacement for $\Delta\phi > M_p$. Together with a power law dependence of the gauge coupling function, we have seen that this leads in effect to the required exponential dependece of g on $\Delta\phi$ at super-Planckian field displacements. While the case for this seems to be quite solid in the Newtonian approximation, we have seen that the full general relativistic description allows much more freedom. This immediately suggests further research into the general strongly curved case, which was out of scope of this thesis. We would also like to highlight the importance of constraining the parameter α that appears in the Swampland Conjecture from a string theory perspective as was briefly anticipated in appendix E. Another interesting research direction would be to establish bottom-up evidence for the Swampland Conjecture in the case where it cannot be related to the Weak Gravity Conjecture, but this is a far more ambitious goal since in this case one would need novel information about states related to quantum gravity which do not also necessarily interact via a gauge force. In general it is very desirable to research further the constraints on experimentally accessible quantum gravity physics such as large field inflation.

We conclude by stating that by now it is clear that quantum gravity, while out of experimental reach, is strongly constrained by theoretical consistency conditions. String theory seems to satisfy all these conditions and so far appears to be a consistent quantum gravity theory. It has passed highly non-trivial tests both from the particle physics perspective (e.g. super-Planckian scattering) but also from the general relativistic one, providing consistency with the expected properties of quantum gravity in strongly curved regions such as black holes. Precisely because of the absence of experimental guidance, it is crucial to explore further the tight web of theoretical constraints that discern quantum gravity from an arbitrary effective quantum field theory.

A Complex Manifolds

From a very abstract perspective, geometric spaces are just topological spaces with a special notion of functions on that space that is compatible with the topology — functions on overlapping subsets should glue if they coincide on the overlap and should coincide globally if their restrictions coincide on an open cover (this is the concept of *ringed spaces*). From this perspective the difference between complex and real manifolds is very simple — we replace functions from our space into the real numbers by functions into the complex numbers. Yet the theory of complex manifolds is much more constrained due to the stronger requirement of complex differentiability. In this appendix we will review some special classes of complex manifolds useful for string compactifications. This is textbook material and a very nice and brief account is given in [24].

> **Complex manifolds**
> A *complex n-fold* is a real $2n$ dimensional manifold $\mathscr{M}$ for which the coordinate maps are functions $\varphi : U \to \mathbb{C}^n$ and the transition functions are holomorphic.

It is often useful to think of complex manifolds as real manifolds with an additional distinguished $(1,1)$ tensor field. Even if one can locally always complexify the coordinates of a real even-dimensional manifold, $z = x + iy$, there is no natural and globally well-defined split of the complexified tangent bundle into holomorphic and antiholomorphic parts in general.

> **Complex Structures**
> An *almost complex structure* on a real manifold $\mathscr{M}$ is an automorphism of the tangent bundle $\mathscr{J} : T\mathscr{M} \to T\mathscr{M}$ which squares to minus the identity $\mathscr{J}^2 = -\mathrm{id}$. If one can cover $\mathscr{M}$ by local holomorphic coordinates z^μ (with conjugates $\bar{z}^{\bar{\mu}}$) such that
>
> $$\mathscr{J}_\nu^{\ \mu} = i\delta_\nu^\mu \ , \quad \mathscr{J}_{\bar{\nu}}^{\ \bar{\mu}} = -i\delta_{\bar{\nu}}^{\bar{\mu}} \ , \tag{A.1}$$
>
> then we call $\mathscr{J}$ a *complex structure* and $\mathscr{M}$ is a complex manifold as defined above. Such an almost complex structure has eigenvalues $\pm i$ and the positive and negative eigenvectors are *holomorphic* ($\partial/\partial z^\mu \equiv \partial_\mu$) and *antiholomorphic* ($\partial/\partial \bar{z}^{\bar{\mu}} \equiv \bar{\partial}_{\bar{\mu}}$) tangent vectors.

The tangent bundle thus splits into the direct sum of holomorphic and antiholomorphic parts

$$T_{\mathbb{C}}\mathscr{M} = T^{1,0}\mathscr{M} \oplus T^{0,1}\mathscr{M} \ . \tag{A.2}$$

The complexified cotangent bundle decomposes in the same way and is spanned by $dz^\mu, d\bar{z}^{\bar{\mu}}$. We have that

$$\Omega^k_{\mathbb{C}} = \bigoplus_{r+s=k} \Omega^{r,s} \ , \tag{A.3}$$

where $\Omega^{r,s}$ is spanned by the wedge products of r holomorphic and s antiholomorphic differentials. Forms of definite degree are termed (r,s)-forms. The exterior derivative refines to the *Dolbeault differentials*.

Dolbeault Operators
The natural complex version of the exterior derivative are the two operators

$$\begin{aligned} \partial &= dz^\mu \wedge \partial_\mu \ : \Omega^{r,s} \to \Omega^{r+1,s} \ , \\ \bar{\partial} &= d\bar{z}^{\bar{\mu}} \wedge \bar{\partial}_{\bar{\mu}} \ : \Omega^{r,s} \to \Omega^{r,s+1} \ . \end{aligned} \tag{A.4}$$

They add to the usual exterior derivative $d = \partial + \bar{\partial}$, square to zero $\partial^2 = \bar{\partial}^2 = 0$ and anti-commute $\{\partial, \bar{\partial}\} = 0$.

These can be in turn used to refine the usual notion of de Rham cohomology $H^i(\mathscr{M})$ to

Dolbeault Cohomology
Dolbeault Cohomology is the cohomology associated to the antiholomorphic Dolbeault differential. It inherits a grading from the corresponding grading of differential forms and we denote the cohomology of $\bar{\partial}$ closed (r,s)-forms modulo exact ones by $H^{r,s}(\mathscr{M})$. The $\mathbb{C}$-dimensions of the Dolbeault cohomologies are known as *Hodge numbers* and denoted by $h^{r,s}$ in analogy to the Betti numbers h^i. It holds that $h^i = \sum_{r+s=i} h^{r,s}$. One can define the *Laplace operator* associated to $\bar{\partial}$ as

$$\Delta_{\bar{\partial}} = \bar{\partial}\bar{\partial}^\dagger + \bar{\partial}^\dagger\bar{\partial} \ , \tag{A.5}$$

where $\bar{\partial}^\dagger$ is the adjoint of $\bar{\partial}$ with respect to the natural scalar product of forms. The space of harmonic (r,s)-forms is denoted by $\mathscr{H}^{r,s}(\mathscr{M})$ and it is a fundamental theorem that $\mathscr{H}^{r,s}(\mathscr{M}) \cong H^{r,s}(\mathscr{M})$.

After these fundamental definitions and theorems we are ready to define a few special classes of complex manifolds.

Kähler Manifolds
A Riemannian metric on a complex manifold is called *Hermitian* if $g(\mathscr{J}X, \mathscr{J}Y) = g(X,Y)$ for all pairs of vector fields. In coordinates this implies that the metric has only mixed indices $g_{\mu\bar{\nu}} = \overline{g_{\bar{\mu}\nu}}$. One can associate to such a metric its *Kähler form*

$$J = i g_{\mu\bar{\nu}} dz^\mu \wedge d\bar{z}^{\bar{\nu}} \ . \tag{A.6}$$

If a manifold admits a Hermitian metric with closed Kähler form, it is called a *Kähler manifold* and the metric is called Kähler as well. Locally such a metric is determined by a single real *Kähler potential*

$$J = i\partial\bar{\partial}K \; . \tag{A.7}$$

The holonomy group of such an n-fold is $U(n)$.

Kähler manifolds are quite restricted. While the restrictions on the metric impose restrictions on the curvature tensors, it is still possible for a Kähler manifold to not admit a Ricci flat Kähler metric. This motivates the following definition.

Calabi-Yau Manifolds
A (compact) *Calabi-Yau* n-fold is a compact n-dimensional Kähler manifold which admints a Ricci-flat metric. Equivalently, it has vanishing first Chern class, or holonomy $SU(n)$, or a global and non-vanishing $(n,0)$-form Ω.

Calabi-Yau manifolds have an extremely restricted cohomology. For one of the most interesting cases of $n = 3$ one has that $h^{0,0}, h^{3,0}, h^{0,3}, h^{3,3} = 1$ and all other Hodge numbers except for $h^{1,1}$ and $h^{1,2} = h^{2,1}$ vanish.

Moduli of Calabi-Yau manifolds
The geometric deformations of a Calabi-Yau metric are restricted by imposing Ricci-flatness. It follows that there are two classes of deformations. First of all deformations of type $\delta g_{a\bar{b}}$ can be shown to correspond to harmonic $(1,1)$-forms and are called *Kähler moduli* because they change the Kähler form. There are $h^{1,1}$ such deformations. Those deformations of type δg_{ab} and $\delta g_{\bar{a}\bar{b}}$ can be shown to be in correspondence with harmonic $(2,1)$-forms and are called *complex structure moduli* because the deformed metric is hermitian only with respect to a deformed complex structure. There are $2h^{2,1}$ such moduli.

B 4d $\mathcal{N} = 1$ Supergravity

Here we will briefly review the bosonic action of ungauged 4d $\mathcal{N} = 1$ supergravity. By $\mathcal{N} = 1$ SUSY the multiplets appearing in the action can be chiral (z, χ) and vector (A_μ, λ) multiplets plus a single gravity multiplet (e^a_μ, ψ_μ). Supersymmetry restricts the scalar sigma model target space to be of Kähler type, which means that the scalar kinetic metric is determined by a real Kähler potential $K(z,\bar{z})$. The scalar potential is given in terms of a holomorphic *superpotential* $W(z)$. The kinetic matrix for the vector multiplet sector is determined in terms of a holomorphic matrix $f_{AB}(z)$. The resulting bosonic action is [54]

$$S = \int \left[\frac{1}{2\kappa^2} \star R - g_{\alpha\bar{\beta}} dz^\alpha \wedge \star \bar{z}^{\bar{\beta}} - V - \mathrm{Re}\,(f_{AB})\, F^A \wedge \star F^B - \mathrm{Im}\,(f_{AB})\, F^A \wedge F^B \right] . \tag{B.1}$$

Here the Kähler metric $g_{\alpha\bar{\beta}}$ and scalar potential V are determined by

$$\begin{aligned} g_{\alpha\bar{\beta}} &= \partial_\alpha \partial_{\bar{\beta}} K(z,\bar{z}) \,, \\ V &= e^{\kappa^2 K} \left(g^{\alpha\bar{\beta}} D_\alpha W \overline{D}_{\bar{\beta}} \overline{W} - 3\kappa^2 |W|^2 \right) \,, \\ D_\alpha W &= \partial_\alpha W + \kappa^2 (\partial_\alpha K) W \,. \end{aligned} \tag{B.2}$$

Supersymmetric classical solutions to (B.1) are by definition invariant under supersymmetry transformations and since the supersymmetry transformations of the bosons vanish in a background without fermions the important condition is the vanishing of the fermion transformations

$$\delta_\varepsilon \mathrm{Fermi} = 0 \,. \tag{B.3}$$

We are interested in constant backgrounds for the scalars only and thus the necessary condition for unbroken sypersymmetry is

$$\delta\chi^\alpha = \frac{1}{\sqrt{2}} P_L \left(\partial\!\!\!/ z^\alpha - e^{\kappa^2 K/2} g^{\alpha\bar{\beta}} \overline{D}_{\bar{\beta}} \overline{W} \right) \varepsilon \,, \tag{B.4}$$

so since $\partial z = 0$ when evaluated on the background we should have

$$\boxed{D_\alpha W = 0 \,,} \tag{B.5}$$

for all chiral multiplets in the theory.

C Frames, Connections and Curvatures of Spherically Symmetric Spacetimes

Here we will briefly present results on the torsion free connections and curvatures for spherically symmetric spacetimes. By spherical symmetry we can foliate our spacetimes into hyperspheres. These n-spheres will be parametrised by n angles defining hyperspherical coordinates $\theta^1, \ldots, \theta^n$ of $\mathbb{R}^{n+1}$. These are defined recursively by [54]

$$x^{n+1}_{(n)} = \cos\theta^n \,, \qquad x^a_{(n)} = \sin\theta^n x^a_{(n-1)} \,, \qquad a = 1, \ldots, n \,, \qquad 0 \leq \theta^n \leq \pi \,. \tag{C.1}$$

The one-sphere is parametrised by $x^2 + ix^1 = \exp(i\theta^1)$, with the polar angle θ^1 running from 0 to 2π. The metric for S^n is then given by

$$d\Omega_n^2 = \sum_{i=1}^{n} \bar{e}^i \otimes \bar{e}^i \,, \tag{C.2}$$

with the co-frame $\bar{e}^i$ defined by

$$\bar{e}^i = \sin\theta^{i+1} \cdots \sin\theta^n d\theta^i \,. \tag{C.3}$$

Alternatively, these are given recursively by

$$\bar{e}^n_{(n)} = d\theta^n \,, \qquad \bar{e}^a_{(n)} = \sin\theta^n \bar{e}^a_{(n-1)} \,, \qquad a = 1, \ldots, n-1 \,. \tag{C.4}$$

In absence of torsion, the first Cartan structure equation can be solved to give the connection one-forms

$$\bar{\omega}^{ab}_{(n)} = \bar{\omega}^{ab}_{(n-1)} \,, \qquad \bar{\omega}^{an}_{(n)} = \cos\theta^n \bar{e}^a_{(n-1)} \,, \qquad a = 1, \ldots, n-1 \,. \tag{C.5}$$

The second structure equation gives the curvature of S^n, which is maximally symmetric, and thus

$$\bar{\rho}^{ab} = \bar{e}^a \wedge \bar{e}^b \,. \tag{C.6}$$

We will write the spherically symmetric metrics in $D = 2 + n$ dimensions in terms of an orthonormal co-frame $e^{\hat{t}}, e^{\hat{r}}, e^i_{(n)}$ as

$$g = -e^{\hat{t}} \otimes e^{\hat{t}} + e^{\hat{r}} \otimes e^{\hat{r}} + \sum_{i=1}^{n} e^i_{(n)} \otimes e^i_{(n)} \,, \tag{C.7}$$

where we allow for an arbitrary overall r-dependent factor of the S^n metric. Note that this can always be removed by changing to standard spherical coordinates $r \to f(r)$. Let us first consider the most general static case

$$e^{\hat{t}} = e^{A(r)} dt \,, \qquad e^{\hat{r}} = e^{B(r)} dr \,, \qquad e^i = e^{C(r)} \bar{e}^i \,. \tag{C.8}$$

Solving the first Cartan structure equation with the help of the S^n connection, one finds

$$\begin{aligned}
\omega^{\hat{t}\hat{r}} &= A' e^{-B} e^{\hat{t}} \,, \\
\omega^{i\hat{r}} &= C' e^{-B} \bar{e}^{i} \,, \\
\omega^{ij} &= \bar{\omega}^{ij} \,.
\end{aligned} \tag{C.9}$$

This can be used to compute the curvature form

$$\begin{aligned}
\rho^{\hat{t}\hat{r}} &= (A'' + A'^2 - B'A') e^{-2B} e^{\hat{r}} \wedge e^{\hat{t}} \,, \\
\rho^{i\hat{r}} &= (C'' + C'^2 - B'C') e^{-2B} e^{\hat{r}} \wedge e^{i} \,, \\
\rho^{i\hat{t}} &= C'A' e^{-2B} e^{\hat{t}} \wedge e^{i} \,, \\
\rho^{ij} &= (e^{-2C} - C'^2 e^{-2B}) e^{i} \wedge e^{j}
\end{aligned} \tag{C.10}$$

and finally the frame components of the Ricci tensor

$$\begin{aligned}
R^{\hat{t}}_{\hat{t}} &= \left[B'A' - A'' - A'^2 - (D-2)C'A'\right] e^{-2B} \,, \\
R^{\hat{r}}_{\hat{r}} &= \left[B'A' - A'' - A'^2 - (D-2)\left(C'^2 + C'' - B'C'\right)\right] e^{-2B} \,, \\
R^{i}_{j} &= \left[\left(B'C' - C'' - (D-2)C'^2 - C'A'\right) e^{-2B} + (D-3) e^{-2C}\right] \delta^{i}_{j} \,.
\end{aligned} \tag{C.11}$$

Let us conclude by specialising to the case $U \equiv A = -B$ and $C = -U + \ln r + \frac{1}{2}\ln f$, corresponding to the line element

$$ds^2 = -e^{2U} dt^2 + e^{-2U}\left(dr^2 + r^2 f(r) d\Omega_2^2\right) \,. \tag{C.12}$$

To simplify the notation, we will use the short hand

$$\nabla^2 = \partial_r^2 + \frac{2}{r}\partial_r \,, \tag{C.13}$$

to denote the flat space Laplacian acting on isotropic functions. The Ricci tensor is

$$\begin{aligned}
R^{\hat{t}}_{\hat{t}} &= -\left(\nabla^2 U + U'\frac{f'}{f}\right) e^{2U} \,, \\
R^{\hat{r}}_{\hat{r}} &= \left(\nabla^2 U - 2U'^2 + U'\frac{f'}{f} - \frac{\nabla^2 f}{f} + \frac{1}{2}\left(\frac{f'}{f}\right)^2\right) e^{2U} \,, \\
R^{i}_{j} &= \delta^{i}_{j}\left(\nabla^2 U + \left(U' - \frac{1}{r}\right)\frac{f'}{f} - \frac{1}{2}\frac{\nabla^2 f}{f} + \frac{1}{r^2}\left(\frac{1}{f} - 1\right)\right) e^{2U} \,.
\end{aligned} \tag{C.14}$$

D Random Matrices

In building supergravity models with many scalar fields one is often interested in *generic* properties of the Kähler metric on the scalar field target space. In the limit of large N, it turns out that its properties are universally described by random matrix theory. For references for the use of random matrix theory, see [24, 34]. An important ensemble of random hermitian matrices is the Wishart ensemble. It is obtained by first drawing the entries of an auxiliary matrix $A \in \mathbb{R}^{N\times M}$ independently from a distribution of variance σ^2 (e.g. normal distribution) and then computing

$$K = AA^\dagger \sim W_N(M) \tag{D.1}$$

The matrix A might have real or complex entries and it is convenient to define the parameter β, which counts the real degrees of freedom per matrix entry. In the limit $N \to \infty$ the behaviour of the spectrum becomes universal and essentially independent of the particular properties of the distribution of the matrix entries. Crucially, the Wishart ensemble is invariant under conjugation by $SO(N)$ in the real case and $SU(N)$ in the complex case. The Wishart ensemble realises eigenvalue repulsion – it is highly unlikely to find two eigenvalues close to each other. In fact, the joint eigenvalue distribution is given by

$$\rho(\lambda) = \frac{1}{Z}e^{-\beta H} \qquad H = \frac{1}{2\sigma}\sum_{i=1}^{N}\lambda_i - \sum_{i<j}\log|\lambda_i - \lambda_j| - \xi\sum_{i=1}^{N}\log(\lambda_i), \tag{D.2}$$

with $\xi = M - N + 1 - 2/\beta$. The second sum describes repulsive interactions between the eigenvalues and evidently diverges in the case of a degenerate spectrum, assigning zero probability. Also of interest is the eigenvalue spectrum

$$\rho(\lambda) = \frac{1}{2\pi N\sigma^2\lambda}\sqrt{(4N\sigma^2 - \lambda)\lambda}, \tag{D.3}$$

and the distribution of the smallest eigenvalue (for $\beta = 1$, at large N)

$$\rho_{\lambda_{\min}}(\lambda) = \frac{1}{2\sigma^2}\left(\sqrt{\frac{N\sigma^2}{\lambda}} + N\right)\exp\left(-\sqrt{\frac{N\lambda}{\sigma^2}} - \frac{N\lambda}{2\sigma^2}\right). \tag{D.4}$$

It will be useful to know the typical scale of the smallest and largest eigenvalues of a Wishart matrix in the large N limit. The scale of the largest eigenvalue is given by $\lambda_N = 4\sigma^2 N$, while the median size of the smallest eigenvalue is $\bar{\lambda}_1 = C\sigma^2/N$, with C an order one constant [32]. An interesting property follows from the rotational invariance of the Wishart ensemble. For large N, the eigenvectors are uniformly distributed on the $(N-1)$-sphere. This phenomenon is known as eigenvector delocalisation. The entries of a normalised, delocalised eigenvector are then distributed according to the normal distribution $\mathscr{N}(0,1/\sqrt{N})$. One can then see that the median size of the largest magnitude entry of such an eigenvector ψ scales like

$$\mathrm{Max}(|\psi_i|) = \frac{\ell_N}{\sqrt{N}}, \tag{D.5}$$

where ℓ_N captures logarithmic corrections, see [32] for details. This implies that eigenvectors typically do not point along the coordinate axes, but are much more likely to be approximately aligned with a diagonal. The intuition behind this is that an N hypercube has $\mathscr{O}(N)$ faces but $\mathscr{O}(\exp(N))$ diagonals.

E The STU and Dilaton Black Holes from Type IIA String Theory

In this appendix we derive the Dilaton black hole from chapter 6 from string theory and give a microscopic description of the WGC states. The setup is that of type IIA string theory on a factorisable $T^6 = (T^2)^3$. We start with the type IIA SUGRA action (3.5). After a Weyl rescaling to Einstein frame $G = \exp(\Phi/2) G_E$, we arrive at

$$
\begin{aligned}
S^E_{\mathrm{IIA}} = \frac{1}{(2\pi)^7 \ell_s^8} \int d^{10}x \sqrt{-G_E} \left[R - \tfrac{1}{2}(\partial\Phi)^2 - \tfrac{1}{2}\left(e^{-\Phi}|H_3|^2 - e^{\frac{3}{2}\Phi}|F_2|^2 + e^{\frac{1}{2}\Phi}|\tilde{F}_4|^2 \right) \right] \\
+ \frac{1}{2(2\pi)^7 \ell_s^8} \int B^2 \wedge F_4 \wedge F_4 \; .
\end{aligned}
\tag{E.1}
$$

Now we compactify on $T^6 = (T^2)^3$. Keeping in mind the fact that the complex structure moduli of the tori decouple in the end, we can set them to zero from the beginning for simplicity. Hence we take G_E to be of the form

$$
G_E = \tilde{g}_4 + \sum_{i=1}^{3} v_i dz^i \otimes d\bar{z}^i \; , \tag{E.2}
$$

where $V = v_1 v_2 v_3$ is the overall compactification volume. First we concentrate on the curvature/Dilaton sector and omit the form fields from the discussion. After a partial integration we arrive at

$$
S = \int d^4x \sqrt{-\tilde{g}} V \left(R_4 + \tfrac{1}{2}\sum_i (\partial \ln v_i)^2 + 2\sum_{i<j} (\partial \ln v_i \cdot \partial \ln v_j) - \tfrac{1}{2}(\partial\Phi)^2 \right) \; . \tag{E.3}
$$

After a further Weyl rescaling with respect to V, the result is

$$
S = \int d^4x \sqrt{-\tilde{g}} V \left(R - \tfrac{1}{2}\sum_i (\partial \ln v_i)^2 - \tfrac{1}{2}(\partial\Phi)^2 - \tfrac{1}{2}(\partial \ln V)^2 \right) \; . \tag{E.4}
$$

in order to decouple the overall volume V, one can further rescale the volume moduli and define the 4d Dilaton by

$$
v_i = e^{-\frac{1}{2}\Phi} \tilde{v}_i \qquad \Rightarrow \qquad \varphi = \Phi - \frac{1}{2} \ln \tilde{V} \; , \tag{E.5}
$$

in order to arrive at

$$
S = \int d^4x \sqrt{-g} \left(R - \tfrac{1}{2}\sum_i \left(\frac{\partial \tilde{v}_i}{\tilde{v}_i} \right)^2 - 2(\partial\varphi)^2 \right) \; . \tag{E.6}
$$

Since the 4d Dilaton does not mix with the Kähler moduli, along with the complex structure moduli, we omit it from now. We introduce the following basis of harmonic 2-forms on $(T^2)^3$

$$\omega^i = \frac{i}{2} dz^i \wedge d\bar{z}^i \,, \qquad \int_{T^6} \omega^i \wedge \star\omega^j = \frac{V}{v_i^2} \delta^{ij} \,. \tag{E.7}$$

The IIAnform fields are expanded according to

$$C_1 = A_0 \,, \qquad C_3 = \sum_i A_i \wedge \omega^i \,, \qquad B_2 = \sum_i b_i \omega^i \,. \tag{E.8}$$

The 4d 1-form gauge fields A_α with field strengths G_α are indexed by $\alpha = 0, i$. Here we omitted the non-dynamical 4d 3-form descending from C_3 and the C_3-axions as well as the universal B_2 axion, since they only contribute to the decoupling complex structure sector. The form field kinetic terms reduce to

$$\begin{aligned} &\int_{T^6} H_3 \wedge \star_{10} H_3 = V \sum_i \frac{db_i \wedge \star_4 db_i}{v_i^2} \,, \qquad \int_{T^6} F_2 \wedge \star_{10} F_2 = V G_0 \wedge \star_4 G_0 \,, \\ &\int_{T^6} \tilde{F}_4 \wedge \star_{10} \tilde{F}_4 = V \sum_i \frac{1}{v_i^2} (G_i - db_i \wedge A_0) \wedge \star_4 (G_i - db_i \wedge A_0) \,. \end{aligned} \tag{E.9}$$

After a Weyl rescaling with respect to V, and the same rescaling of the volumes with respect to Φ, the full action, including the reduced Chern-Simons term is

$$\begin{aligned} S_{\text{IIA}}^{\text{red.}} = \frac{1}{2\kappa^2} \int \Bigg(&R \star 1 - \tfrac{1}{2} \sum_i \frac{1}{\tilde{v}_i^2} \left(d\tilde{v}_i \wedge \star d\tilde{v}_i + db_i \wedge \star db_i \right) \\ &- \tfrac{1}{2} \tilde{v}_1 \tilde{v}_2 \tilde{v}_3 G_0 \wedge \star G_0 - \tfrac{1}{2} \sum_i \frac{\tilde{v}_1 \tilde{v}_2 \tilde{v}_3}{\tilde{v}_i^2} (G_i - db_i \wedge A_0) \wedge \star (G_i - db_i \wedge A_0) \Bigg) \\ &- \frac{1}{4\kappa^2} \int \sum_{ijk} |\varepsilon_{ijk}| b_i (G_j - db_j \wedge A_0) \wedge (G_k - db_k \wedge A_0) \,. \end{aligned} \tag{E.10}$$

The last step is then to introduce the shifted gauge fields $\tilde{A}_i = A_i - b_i A_0$ and canonically normalise the scalars via $\phi_i = \ln \tilde{v}_i$.

$$\begin{aligned} S_{\text{IIA}}^{\text{red.}} = \frac{1}{2\kappa^2} \int \big(&R \star 1 - \tfrac{1}{2} \sum d\varphi_i \wedge \star d\varphi_i - \tfrac{1}{2} \sum e^{-2\varphi_i} db_i \wedge \star db_i \\ &- \tfrac{1}{2} e^{+\varphi_1 + \varphi_2 + \varphi_3} G_0 \wedge \star G_0 - \tfrac{1}{2} e^{-\varphi_1 + \varphi_2 + \varphi_3} \hat{G}_1 \wedge \star \hat{G}_1 \\ &- \tfrac{1}{2} e^{+\varphi_1 - \varphi_2 + \varphi_3} \hat{G}_2 \wedge \star \hat{G}_2 - \tfrac{1}{2} e^{+\varphi_1 + \varphi_2 - \varphi_3} \hat{G}_3 \wedge \star \hat{G}_3 \\ &- b_1 \hat{G}_2 \wedge \hat{G}_3 - b_2 \hat{G}_1 \wedge \hat{G}_3 - b_3 \hat{G}_1 \wedge \hat{G}_2 \big) \,, \end{aligned} \tag{E.11}$$

where the hatted field strengths are $\hat{G}_i = d\tilde{A}_i + b_i G_0 \equiv \tilde{G}_i + b_i G_0$.

$$0 \stackrel{!}{=} \left. \frac{\delta \mathscr{L}}{\delta b_i} \right|_{b_i = 0} \,. \tag{E.12}$$

Here it suffices to look at the derivative of the potential since the axion kinetic terms vanish upon imposing $b_i = 0$. The condition that one will get is the vanishing of some quadratic polynomial in the field strengths with exponentials of the Kähler moduli as coefficients. We will not state the equations of motions in full generality.

One can easily see that a charge configuration of type (Q,P,P,P) or (P,Q,Q,Q) will solve the axion equation of motions. This is because in the Chern-Simons action the only terms linear in the axions are of type $G_i \wedge G_j$, which all vanish because the G_i are parallel. In the kinetic terms, the only terms linear in the axions are of form $G_i \wedge \star G_0$, which vanish aswell because G_i and $\star G_0$ are also parallel. The derivatives of all the terms with higher degree than 1 in the axions will vanish upon imposing $b_i = 0$.

Now the Lagrangian reduces to the very simple form

$$\begin{aligned} S_{\text{IIA}}^{\text{red.}} = \frac{1}{2\kappa^2} \int \Bigg(& R \star 1 - \tfrac{1}{2} \sum_i d\varphi_i \wedge \star d\varphi_i \\ & - \tfrac{1}{2} e^{+\varphi_1+\varphi_2+\varphi_3} G_0 \wedge \star G_0 - \tfrac{1}{2} e^{-\varphi_1+\varphi_2+\varphi_3} G_1 \wedge \star G_1 \\ & - \tfrac{1}{2} e^{+\varphi_1-\varphi_2+\varphi_3} G_2 \wedge \star G_2 - \tfrac{1}{2} e^{+\varphi_1+\varphi_2-\varphi_3} G_3 \wedge \star G_3 \Bigg) \ . \end{aligned} \tag{E.13}$$

This is the so-called STU model of $\mathcal{N} = 2$ SUGRA. We will choose the charge configuration with mostly magnetic charges[1] We look at spherically symmetric extremal black hole solutions using the metric ansatz (6.50). The gauge field background is

$$G_0 = g_0^2 \frac{q_0}{r^2} e^{2U} e^{\hat{t}} \wedge e^{\hat{r}} \ , \qquad G_i = \frac{p_i}{r^2} e^{2U} e^{\hat{\theta}} \wedge e^{\hat{\phi}} \ . \tag{E.14}$$

The solution is given in terms of four harmonic functions (see for example [86])

$$H_0 = \frac{1}{g_{0,\infty}} + \frac{|q_0|}{r} \ , \qquad H_i = g_{i,\infty} + \frac{|p_i|}{r} \ , \qquad U = -\frac{1}{4} \ln \left(\prod_i H_i \right) \ , \tag{E.15}$$

$$\varphi_1 = -\frac{1}{2} \ln \left(\frac{H_2 H_3}{H_0 H_1} \right) \ , \qquad \varphi_2 = -\frac{1}{2} \ln \left(\frac{H_1 H_3}{H_0 H_2} \right) \ , \qquad \varphi_3 = -\frac{1}{2} \ln \left(\frac{H_1 H_2}{H_0 H_3} \right) \ . \tag{E.16}$$

Different consistent ways to reduce this to the case of two gauge fields and a single scalar, i.e. the $\mathcal{N} = 2$ SUGRA with a single gauge multiplet, give different versions of the α deformed Dilaton black hole (6.61). For example, dualising G_2 we get a charge configuration of type (Q_0, P_1, Q_2, P_3). Upon identifying $Q_0 = Q_2$, $P_1 = P_3$ and switching off two of the scalars $\varphi_1 = \varphi_3 = 0$, we obtain the $\alpha = 1$ Dilaton black hole. If we instead choose to stay in the (Q,P,P,P) duality frame and identify the magnetic charges and also all three scalars, we arrive at the $\alpha = \sqrt{3}$ Dilaton black hole. Note that the factor of $\sqrt{3}$ in α arises as the number of scalar fields that are being collectively displaced. This is reminiscent of the $\sqrt{N}$ enhancement for diagonal N-flation. In fact if this relation generally holds true, we could put an upper bound on α by bounding the number of scalars in string compactifications.

[1] This particular charge configuration corresponds to localised $D0$ branes and $D4$ branes wrapping the three four-cycles dual to the two-cycles defined by the T^2 factors in T^6.

Finally we would like to derive the WGC states from BPS-branes wrapped on the various compactification cycles. To do so we start with the DBI action (3.13) for a p-brane wrapped on a p-cycle Σ_p. We use the fact that the metric determinant factorises upon compactification and perform the same steps as for the IIA SUGRA action. After this we end up with

$$S_{\mathrm{DBI}} = -T_p \frac{\mathrm{Vol}(\Sigma_p)}{\sqrt{V}} e^{\frac{p-3}{4}\Phi} \int d\tau \sqrt{-\tilde{g}_{\mu\nu}\dot{X}^\mu \dot{X}^\nu} \, . \tag{E.17}$$

This can be identified with the action of a point particle of mass

$$m_p = T_p \frac{\mathrm{Vol}(\Sigma_p)}{\sqrt{V}} e^{\frac{p-3}{4}\Phi} \, . \tag{E.18}$$

We note that the branes couple to the gauge fields through the Chern-Simons action (3.14), which includes a factor of T_p, so the gauge couplings are

$$g_0 M_p = T_0 \frac{1}{\sqrt{V}} \, , \qquad g_i M_p = T_4 \frac{\frac{1}{2}|\varepsilon_{ijk}| v_i v_j}{\sqrt{V}} \, . \tag{E.19}$$

The IIA theory includes branes of dimension $p = 0,2,4,6$. The masses of the corresponding particles can be read off from (E.18)

$$\begin{aligned} m_0 &= T_0 \frac{1}{\sqrt{V}} \, , \qquad & m_4^{(i)} &= T_4 \frac{\frac{1}{2}|\varepsilon_{ijk}| v_i v_j}{\sqrt{V}} \, , \\ m_6 &= T_6 \frac{V}{\sqrt{V}} \, , \qquad & m_2^{(i)} &= T_2 \frac{v_i}{\sqrt{V}} \, . \end{aligned} \tag{E.20}$$

As we see these provide the threshold Weak Gravity Conjecture states for the corresponding gauge fields. In the duality frame of (E.13), the Weak Gravity Conjecture states corresponding to the gauge fields A_0 and A_i are the D0 branes and D4-branes respectively. Upon dualising the gauge fields, the corresponding Weak Gravity Conjecture states are the dual branes, namely the D6 and D2 ones.

Bibliography

[1]B. P. Abbott et al., Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[2]B. P. Abbott et al., Phys. Rev. Lett. **116**, 241103 (2016), arXiv:1606.04855 [gr-qc].

[3]R. A. Hulse and J. H. Taylor, Astrophys. J. **195**, L51–L53 (1975).

[4]LIGO and Virgo Collaborations, *Data release for event GW150914*, https://losc.ligo.org/events/GW150914/ (visited on 11/14/2016).

[5]P. A. R. Ade et al., Phys. Rev. Lett. **112**, 241101 (2014), arXiv:1403.3985 [astro-ph.CO].

[6]P. A. R. Ade et al., Phys. Rev. Lett. **114**, 101301 (2015), arXiv:1502.00612 [astro-ph.CO].

[7]The BICEP2 Collaboration, *BICEP2 E- and B-mode maps*, BICEP2 2014 Release Figures from Papers, (2014) http://bicepkeck.org/bicep2_2014_release.html#figures (visited on 12/19/2016).

[8]D. H. Lyth, Phys. Rev. Lett. **78**, 1861–1863 (1997), arXiv:hep-ph/9606387 [hep-ph].

[9]N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, JHEP **06**, 060 (2007), arXiv:hep-th/0601001 [hep-th].

[10]T. Rudelius, JCAP **1509**, 020 (2015), arXiv:1503.00795 [hep-th].

[11]J. Brown, W. Cottrell, G. Shiu, and P. Soler, JHEP **10**, 023 (2015), arXiv:1503.04783 [hep-th].

[12]M. Montero, A. M. Uranga, and I. Valenzuela, JHEP **08**, 032 (2015), arXiv:1503.03886 [hep-th].

[13]A. Hebecker, P. Mangat, F. Rompineve, and L. T. Witkowski, Phys. Lett. **B748**, 455–462 (2015), arXiv:1503.07912 [hep-th].

[14]J. Brown, W. Cottrell, G. Shiu, and P. Soler, JHEP **04**, 017 (2016), arXiv:1504.00659 [hep-th].

[15]B. Heidenreich, M. Reece, and T. Rudelius, JHEP **12**, 108 (2015), arXiv:1506.03447 [hep-th].

[16]D. Junghans, JHEP **02**, 128 (2016), arXiv:1504.03566 [hep-th].

[17]D. Klaewer and E. Palti, (2016), arXiv:1610.00010 [hep-th].

[18]H. Ooguri and C. Vafa, Nucl. Phys. **B766**, 21–33 (2007), arXiv:hep-th/0605264 [hep-th].

[19]T. Banks, M. Dine, P. J. Fox, and E. Gorbatov, JCAP **0306**, 001 (2003), arXiv:hep-th/0303252 [hep-th].

[20]J. Polchinski, Int. J. Mod. Phys. **A19S1**, 145–156 (2004), arXiv:hep-th/0304042 [hep-th].

[21]T. Banks and N. Seiberg, Phys. Rev. **D83**, 084019 (2011), arXiv:1011.5120 [hep-th].

[22]A. H. Guth, Phys. Rev. **D23**, 347–356 (1981).

[23]S. Weinberg, *Cosmology* (Oxford Univ. Pr., Oxford, UK, 2008).

[24]D. Baumann and L. McAllister, *Inflation and String Theory* (Cambridge University Press, 2015), arXiv:1404.2601 [hep-th].

[25]K. K. S. Wu, O. Lahav, and M. J. Rees, Nature **397**, 225–230 (1999), arXiv:astro-ph/9804062 [astro-ph].

[26]D. Baumann, *Lecture Notes on Cosmology, Part III Mathematical Tripos*, http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf (visited on 10/17/2016).

[27]ESA, Planck Collaboration, *Cosmic Microwave Background seen by Planck*, (2013) http://sci.esa.int/planck/51553-cosmic-microwave-background-seen-by-planck/ (visited on 12/19/2016).

[28]P. A. R. Ade et al., (2015), arXiv:1502.01589 [astro-ph.CO].

[29]J. Preskill, Ann. Rev. Nucl. Part. Sci. **34**, 461–530 (1984).

[30]P. A. R. Ade et al., (2015), arXiv:1502.02114 [astro-ph.CO].

[31]K. Freese, J. A. Frieman, and A. V. Olinto, Phys. Rev. Lett. **65**, 3233–3236 (1990).

[32]T. C. Bachlechner, C. Long, and L. McAllister, JHEP **12**, 042 (2015), arXiv:1412.1093 [hep-th].

[33]J. E. Kim, H. P. Nilles, and M. Peloso, JCAP **0501**, 005 (2005), arXiv:hep-ph/0409138 [hep-ph].

[34]R. Easther and L. McAllister, JCAP **0605**, 018 (2006), arXiv:hep-th/0512102 [hep-th].

[35]S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, JCAP **0808**, 003 (2008), arXiv:hep-th/0507205 [hep-th].

[36]T. C. Bachlechner, M. Dias, J. Frazer, and L. McAllister, Phys. Rev. **D91**, 023520 (2015), arXiv:1404.7496 [hep-th].

[37]J. Blair, C. Edwards, and J. Johnson, Mathematics of Computation **30**, 827–830 (1976).

[38]L. McAllister, E. Silverstein, and A. Westphal, Phys. Rev. **D82**, 046003 (2010), arXiv:0808.0706 [hep-th].

[39]N. Kaloper and L. Sorbo, Phys. Rev. Lett. **102**, 121301 (2009), arXiv:0811.1989 [hep-th].

[40]R. Flauger, L. McAllister, E. Pajer, A. Westphal, and G. Xu, JCAP **1006**, 009 (2010), arXiv:0907.2916 [hep-th].

[41]F. Marchesano, G. Shiu, and A. M. Uranga, JHEP **09**, 184 (2014), arXiv:1404.3040 [hep-th].

[42]A. Hebecker, P. Mangat, F. Rompineve, and L. T. Witkowski, Nucl. Phys. **B894**, 456–495 (2015), arXiv:1411.2032 [hep-th].

[43]A. Hebecker, F. Rompineve, and A. Westphal, JHEP **04**, 157 (2016), arXiv:1512.03768 [hep-th].

[44] A. Hebecker, J. Moritz, A. Westphal, and L. T. Witkowski, Phys. Lett. **B754**, 328–334 (2016), arXiv:1512.04463 [hep-th].

[45] A. Hebecker, J. Jaeckel, F. Rompineve, and L. T. Witkowski, (2016), arXiv:1606.07812 [hep-ph].

[46] L. McAllister, P. Schwaller, G. Servant, J. Stout, and A. Westphal, (2016), arXiv:1610.05320 [hep-th].

[47] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string* (Cambridge University Press, 2007).

[48] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond* (Cambridge University Press, 2007).

[49] R. Blumenhagen, D. Lst, and S. Theisen, *Basic concepts of string theory*, Theoretical and Mathematical Physics (Springer, Heidelberg, Germany, 2013).

[50] K. Becker, M. Becker, and J. H. Schwarz, *String theory and M-theory: A modern introduction* (Cambridge University Press, 2006).

[51] B. R. Greene, K. Schalm, and G. Shiu, Nucl. Phys. **B584**, 480–508 (2000), arXiv:hep-th/0004103 [hep-th].

[52] O. DeWolfe and S. B. Giddings, Phys. Rev. **D67**, 066008 (2003), arXiv:hep-th/0208123 [hep-th].

[53] R. Blumenhagen, B. Kors, D. Lust, and S. Stieberger, Phys. Rept. **445**, 1–193 (2007), arXiv:hep-th/0610327 [hep-th].

[54] D. Z. Freedman and A. Van Proeyen, *Supergravity* (Cambridge Univ. Press, Cambridge, UK, 2012).

[55] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, Phys. Rev. **D68**, 046005 (2003), arXiv:hep-th/0301240 [hep-th].

[56] V. Balasubramanian, P. Berglund, J. P. Conlon, and F. Quevedo, JHEP **03**, 007 (2005), arXiv:hep-th/0502058 [hep-th].

[57] O. DeWolfe, A. Giryavets, S. Kachru, and W. Taylor, JHEP **07**, 066 (2005), arXiv:hep-th/0505160 [hep-th].

[58] J. Maltz and L. Susskind, (2016), arXiv:1611.00360 [hep-th].

[59] J. Maltz, (2016), arXiv:1611.03491 [hep-th].

[60] P. Svrcek and E. Witten, JHEP **06**, 051 (2006), arXiv:hep-th/0605206 [hep-th].

[61] E. Palti, JHEP **10**, 188 (2015), arXiv:1508.00009 [hep-th].

[62] T. W. Grimm and J. Louis, Nucl. Phys. **B718**, 153–202 (2005), arXiv:hep-th/0412277 [hep-th].

[63] C. Long, L. McAllister, and P. McGuirk, Phys. Rev. **D90**, 023501 (2014), arXiv:1404.7852 [hep-th].

[64] T. W. Grimm, Phys. Rev. **D77**, 126007 (2008), arXiv:0710.3883 [hep-th].

[65]M. Cicoli, K. Dutta, and A. Maharana, JCAP **1408**, 012 (2014), arXiv:1401.2579 [hep-th].

[66]F. Denef, M. R. Douglas, and B. Florea, JHEP **06**, 034 (2004), arXiv:hep-th/0404257 [hep-th].

[67]A. Strominger, *Black Holes: The Harmonic Oscillators of the 21st Century*, Talk given at Perimeter Institute Colloquium, (2014) http://www.perimeterinstitute.ca/videos/black-holes-harmonic-oscillators-21st-century (visited on 11/15/2016).

[68]C. Vafa, (2005), arXiv:hep-th/0509212 [hep-th].

[69]J. D. Bekenstein, Phys. Rev. **D23**, 287 (1981).

[70]R. Bousso, JHEP **07**, 004 (1999), arXiv:hep-th/9905177 [hep-th].

[71]L. Susskind, (1995), arXiv:hep-th/9501106 [hep-th].

[72]T. Banks and L. J. Dixon, Nucl. Phys. **B307**, 93–108 (1988).

[73]X. G. Wen and E. Witten, Phys. Lett. **B166**, 397–401 (1986).

[74]C. Cheung and G. N. Remmen, Phys. Rev. Lett. **113**, 051601 (2014), arXiv:1402.2287 [hep-ph].

[75]H. Ooguri and C. Vafa, (2016), arXiv:1610.01533 [hep-th].

[76]P. Saraswat, (2016), arXiv:1608.06951 [hep-th].

[77]B. Heidenreich, M. Reece, and T. Rudelius, JHEP **02**, 140 (2016), arXiv:1509.06374 [hep-th].

[78]A. Hebecker, F. Rompineve, and A. Westphal, (2015), arXiv:1512.03768 [hep-th].

[79]B. Heidenreich, M. Reece, and T. Rudelius, (2016), arXiv:1606.08437 [hep-th].

[80]J. Brown, W. Cottrell, G. Shiu, and P. Soler, (2015), arXiv:1504.00659 [hep-th].

[81]W. Cottrell, G. Shiu, and P. Soler, (2016), arXiv:1611.06270 [hep-th].

[82]P. Bizon, Phys. Rev. **D47**, 1656–1663 (1993), arXiv:hep-th/9209106 [hep-th].

[83]F. Baume and E. Palti, JHEP **08**, 043 (2016), arXiv:1602.06517 [hep-th].

[84]G. Dall'Agata, Springer Proc. Phys. **142**, 1–45 (2013), arXiv:1106.2611 [hep-th].

[85]K. Goldstein, N. Iizuka, R. P. Jena, and S. P. Trivedi, Phys. Rev. **D72**, 124021 (2005), arXiv:hep-th/0507096 [hep-th].

[86]S. Ferrara and R. Kallosh, Phys. Rev. **D54**, 1514–1524 (1996), arXiv:hep-th/9602136 [hep-th].

[87]L. Andrianopoli, R. D'Auria, S. Ferrara, and M. Trigiante, Lect. Notes Phys. **737**, 661–727 (2008), arXiv:hep-th/0611345 [hep-th].

[88]A. Nicolis, JHEP **07**, 023 (2008), arXiv:0802.3923 [hep-th].

[89]R. Kallosh, A. D. Linde, T. Ortin, A. W. Peet, and A. Van Proeyen, Phys. Rev. **D46**, 5278–5302 (1992), arXiv:hep-th/9205027 [hep-th].

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 23.12.2016 ....