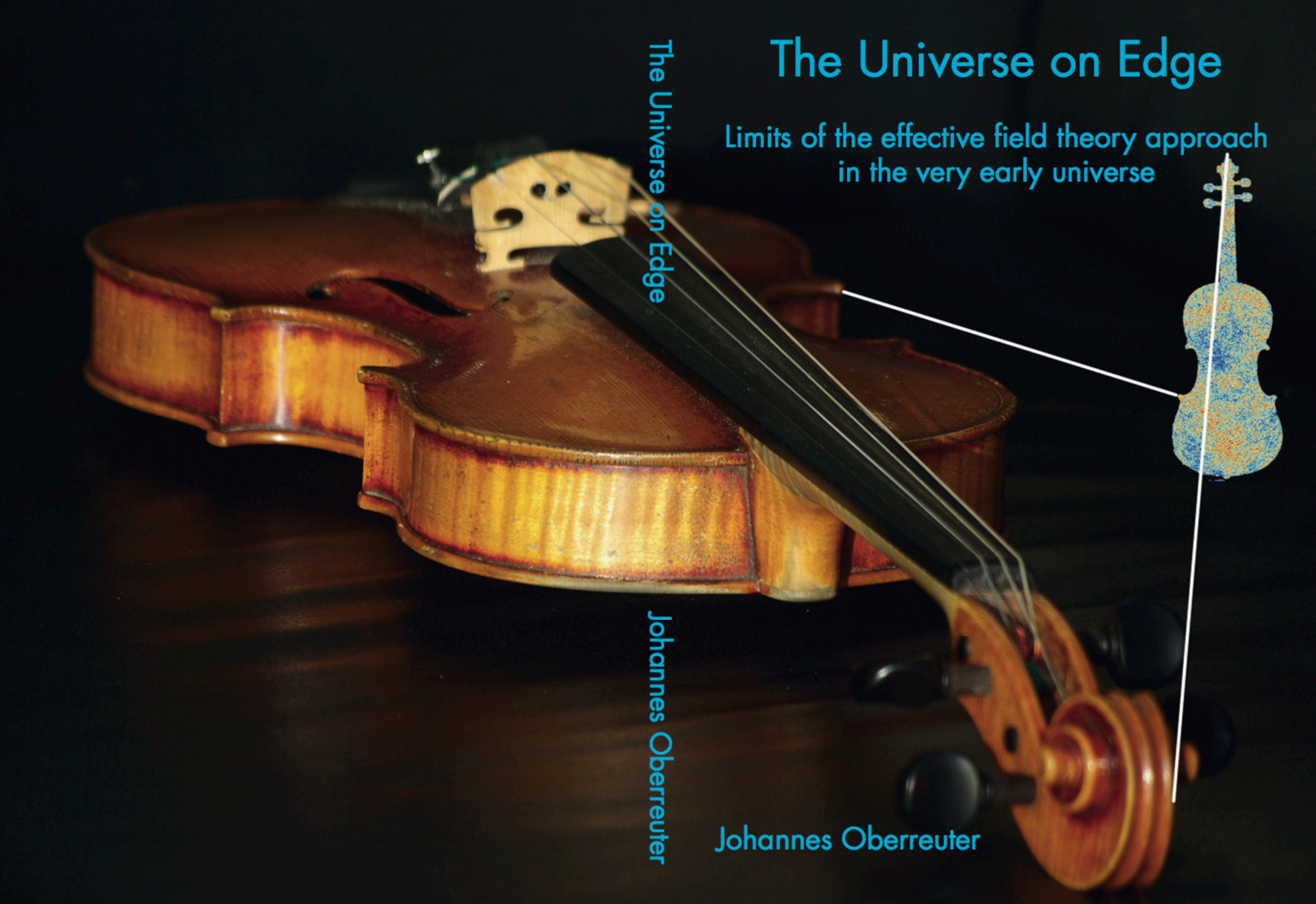
File ID	485112
Filename	Thesis

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	Dissertation
Title	The universe on edge: Limits of the effective field theory approach in the very early universe
Author	J.M. Oberreuter
Faculty	Faculty of Science
Year	2013
Pages	xii, 172
ISBN	9789461917485

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/446399


The Universe on Edge
Johannes Oberreuter
The Universe on Edge
Limits of the effective field theory approach
in the very early universe
Johannes Oberreuter

The Universe on Edge

Limits of the effective field theory approach in the very early universe

The Universe on Edge

Limits of the effective field theory approach in the very early universe

Academisch Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. D. C. van den Boom

ten overstaan van een door het college voor promoties ingestelde

commissie, in het openbaar te verdedigen in de Agnietenkapel

op woensdag 29 mei 2013, te 12:00 uur

door

Johannes Michael Oberreuter

geboren te Gräfelfing, Duitsland.

Promotiecommissie

Promotor

prof. dr. E. P. Verlinde

Co-Promotores

dr. K. E. Schalm

dr. J. P. van der Schaar

Overige leden

prof. dr. A. Achúcarro

prof. dr. J. de Boer

prof. dr. B. Craps

dr. D. Roest

prof. dr. K. Skenderis

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

ISBN 978-94-6191-748-5

The cover illustration is inspired by “Der Zeichner der Laute”, woodcut by Albrecht Dürer (1525). The pattern in violin shape is the map of the temperature anisotropies of the cosmic microwave background radiation as measured by the Planck spacecraft (copyright ESA and the Planck Collaboration).

This work is part of the research program of the “Stichting voor Fundamenteel Onderzoek der Materie (FOM)”, which is part of the “Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)”.

Von der inspiration

Nur ein anfänger von engel
fliegt unterhalb der wolken
(noch ist er in sich selbst
nicht weit genug entfernt vom menschen)

Wenn deine stirn ein flügel streift,
ist’s einer von ihnen,

und du stehst am anfang
wie er.

Reiner Kunze [6]

PUBLICATIONS

This thesis is based on the following publications:

✏ [1] Sjoerd Hardeman, Johannes. M. Oberreuter, Gonzalo A. Palma, Koenraad E. Schalm and Ted van der Aalst,
"The everpresent eta-problem: knowledge of all hidden sectors required"
JHEP **1104** (2011) 009 [arXiv:1012.5966 [hep-ph]].

✏ [2] Ana Achucarro, Sjoerd Hardeman, Johannes. M. Oberreuter, Koenraad E. Schalm and Ted van der Aalst,
"Decoupling limits in multi-sector supergravities"
JCAP **1303** (2013) 038 [arXiv:1012.5966 [hep-th]].

✏ [3] Johannes. M. Oberreuter, Koenraad E. Schalm and Jan Pieter van der Schaar,
"Resolution of Cosmic Singularities with the AdS/CFT correspondence"
in preparation.

Other publications by the author are:

✏ [4] Gabriel Lopes Cardoso, Anna Ceresole, Gianguido Dall'Agata, Johannes M. Oberreuter, Jan Perz,
"First-order flow equations for extremal black holes in very special geometry."
JHEP **0710** (2007) 063 [arXiv:0706.3373 [hep-th]].

✏ [5] Gabriel Lopes Cardoso, Johannes M. Oberreuter, Jan Perz,
"Entropy function for rotating extremal black holes in very special geometry."
JHEP **0705** (2007) 025 [hep-th/0701176]

CONTENTS

CHAPTER 1

EFFECTIVE THEORIES FOR GRAVITY IN COSMOLOGY

1.1 THE REASONABLE EFFECTIVENESS OF PHYSICS IN THE COSMOLOGICAL SCIENCES

Why does physics work? Why is it a predictive science whose mathematical methods we can use to theoretically produce an observable and experimentally testable result?

The world around us is complex. Most processes involve a multitude of particles which are connected to each other with a multitude of different interactions governed by physics we might not even know yet. Hence, at first sight, it seems as if there was no reason to hope to be able to calculate anything useful, let alone correct.

The reason why this is still possible is that physics has scales. Different physics is relevant at different scales. A scale is a range of, say, energy, momentum, inverse length or time, which are all equivalent for that matter, within which the dynamics can be described to any desired level of accuracy by using a particular theory. There is a well-known procedure how to determine this theory as a simplification of physics at higher scales, or how to devise a theory which works although the physics at higher scales is completely obscure. Such a theory is known as a *(low energy) effective theory*. It describes the essential phenomena for a certain part of the parameter space.

The idea of reducing complexity on lower energy scales originates in statistical physics. Systems, which generically consist of an order of the Avogadro number $N_A = 6.022 \times 10^{23}$ particles and would be far too complicated to describe on the microscopic level of their mutual interactions, can be understood much better on a low energy scale in terms of only a few macroscopic variables like temperature, pressure, entropy, chemical potential or magnetization. In this case, the low energy approximation simplifies the description of the system. We use these thermodynamic variables instead of the fundamental ones like momenta and positions of the individual particles. What makes it somewhat natural to perform this limit is the fact that thermodynamical systems very often have an inherent upper bound on the energy scale. When describing a gas or a fluid, the low energy physics is governed by the inter-atomic interactions and atomic physics itself can be safely discarded. Therefore, the size of an atom can be taken as the *cut-off*, the inverse of the maximum energy scale. Of course, this cut-off represents a limit of the validity of the approximation. The closer the energy scale of interest lies to the cut-off, the more the effects of atomic physics will play a rôle and can no longer be ignored. In other words, at the cut-off scale, *new physics* enters the description.

Very interesting phenomena that can be studied in the low energy effective theory are *phase transitions*. At a generic point, thermal fluctuations of a continuous field are correlated only over a few atomic distances. At a critical point, however, the correlation length diverges and becomes of the order of magnitude of the material sample. A critical point is the point in parameter space where the order parameter, which emerges due to a first order phase transition, becomes continuous and the difference between the two adjacent phases vanishes. In other words, at the critical point, a single thermodynamic state bifurcates into two distinct phases. This leads to long-range thermal fluctuations, which characterize a second order phase transition. It turns out that, remarkably, very different systems can share essential properties around a critical point, e.g. in the vicinity of the critical temperature. Assuming just a few general symmetries, Landau theory predicts, that a wide range of system show *universal behavior* of characteristic quantities like the order parameter, the correlation length and other thermodynamic quantities, which universally depend on the temperature exponentially like $(T - T_C)^{\alpha}$. This *universality* is also remarkable, because it shows that most of the relevant physics at low energies does not depend on the high energy behavior. However, the power with which the thermodynamic quantities depend on the temperature, the *critical exponent* α, can differ between different systems. Systems with the same critical exponent form a *universality class*. It should be noted that at the critical point, those quantities do not scale at all, i.e. they are scale independent.

When turning from a classical to a quantum theory, we can replace the statistical fluctuations with quantum fluctuations and describe the statistical system near the critical point with a quantum field theory. The mass of the quantum field is taken to be well below the cut-off scale and must vanish at the critical point. In this case, the field theory at the critical point has no scales and is a *conformal field theory*. The universal behavior of the multitude of phase transitions in nature is caused by the low energy physics being independent of the physics at the cut-off. The quantum field theory is *renormalizable* and predictive, after determining a finite number of parameters from experiment. Renormalizability and universality are the two sides of one coin. The parameters which cannot be determined theoretically are linked to the infamous ultra-violet divergences of quantum field theories. They signal that an infra-red quantity depends on the UV physics. If physics above the cutoff is unknown, as is generically the case in a quantum field theory, the value of such quantities, like the charge and mass of an elementary particle, cannot be derived but only measured. On the other hand, if a specific interaction in the UV involving heavy fields has been integrated out, the interaction can enter the effective theory as non-local.

It is a remarkable feature that we can write down a theory which accurately describes the physics at accessible scales, without the need to resort to knowledge about fundamental physics. Though it is possible to simplify physics by distilling the relevant physics from a UV complete theory, it is impossible to infer the fundamental degrees of freedom from the low energy physics. Yet, we know that new physics must enter at several scales. The aim of this thesis is to contribute to the quest for such new physics.

Since the standard model of particle physics, as a very successful application of these ideas to high-energy particle physics, unifies three of the fundamental forces of nature, the electromagnetic, strong and weak interactions, into one framework of quantum gauge field theory, new physics is expected to come mainly from gravity. It turns out that gravity cannot be described by the same framework which we used above because it is *non-renormalizable*. The scale at which new degrees of freedom of gravity enter the description is conjectured to be the *Planck-scale*, which is at 1.22×10^{19} GeV. In terms of a length scale, this would be 1.612×10^{-35} m, which is about 10^{20} times smaller than a proton. At this size, it is not even clear if it makes sense to talk about space-time as a geometric concept, let alone define a quantum field theory. This scale seems to be nearly inaccessible experimentally, too. As a comparison, the *Large Hadron Collider (LHC)*, which seems to be on its way to measure the last parameters needed to complete the standard model, reaches a top energy of 14 TeV, still an order of 10^{16} short of probing this scale.[1]

We have one system at hand, which at some time in the past most certainly has been governed by physics at the highest possible energy range: namely the universe. In our current understanding the universe is now expanding and has been doing so since its beginning, the *big bang*. This means that its density is becoming smaller and smaller, whereas if we return in time towards the big bang, its density grows and with it the temperature and the energy. The furthest back in time we can experimentally look is observing the *cosmic microwave background radiation (CMB)*, which was formed approximately 360.000 years after the big bang. This was the time when, due to *recombination* of hydrogen and helium ions with free electrons, the universe became transparent for photons. At this point, the universe had a temperature of about 3000 K or 0.2 eV, much less than the Planck scale. However, the gravitational physics governing the behavior of the plasma descends from Planck scale physics, which is thus imprinted on the CMB. It has been measured to very high accuracy over the past years by the COBE and WMAP missions and by the PLANCK satellite, whose results are expected to come forward in the coming year at the time of writing. This means, that there is a good chance to learn something about new physics at the Planck scale from such high-precision measurements of the CMB.

[1] However, it might be possible to explore the Planck-scale via the effects its physics has on the physics at accessible scales. The Alpha-experiment might be able to discover a breaking of the Lorentz- or CPT-symmetry in comparing the spectra of hydrogen and anti-hydrogen which might come from Planck scale effects [7].

This thesis is not so much concerned with the observational wealth of the early universe but with the theoretical puzzles which it imposes on us. For only if we understand the early universe well enough such that our model of it is without any conundrums can we hope to interpret the data in a useful way. There are quite a few problems in the very early universe, which we can not understand very well with the technology of effective field theory as explained above. Instead, we have to resort to techniques coming from candidates for a quantum theory of gravity to tackle them. String theory has for a long time been such a candidate, which changes the rules of the effective field theory game quite a bit. While in this framework it is possible to solve the renormalizability issues of gravity, it has not yet been predictive to the point where it could have been supported by experimental evidence.

I have picked two topics which shed light on the problems which arise when combining novel with established techniques in the early universe. The first problem is connected to its very beginning, the big bang itself. In the standard cosmological model it is viewed as a singularity, because as the contraction of the universe is extrapolated into the past, its density diverges and so does the curvature. It is important to notice, though, that this singularity is an artefact of using an effective theory, namely *general relativity*, outside of its range of validity. It is a theory which is in particular only valid at large distances and has in fact not been tested below about 55 μm [8]. At the Planck scale, new physics is expected to *resolve the initial singularity*. For instance, in string theory, gravitons are oscillations of the string, which itself shrinks to a point at low energies and the effective theory treats it as a conventional particle. I am presenting an attempt to understand this using the *Anti-de Sitter/conformal field theory (AdS/CFT) correspondence* (chapter 2). With the help of this duality, new physics originating from string theory becomes tractable in the dual conformal field theory. There, a space-time at strong coupling, i.e. at high curvature, is described as a field theory at weak coupling. Briefly after the big bang at about 10^{-36} to 10^{-33} s, current cosmological models assume en exponentially accelerated expansion of the universe. This so-called *inflation* is driven by the potential energy of some new degree of freedom. Although the universe is already below the Planck energy at this time, this potential must descend from a fundamental theory which includes gravity. Thus, inflation is a prototypical example of the effective field theory approach: All the physics above some cut-off, which is assumed to be close to the Planck scale, gets integrated out to obtain an effective theory, which has to contain at least one scalar degree of freedom that can serve as the inflaton. The shape of its potential, which is subject to conditions ensuring that inflation works, descends from the UV physics. While this is currently a matter of taste, there are generic operators which any UV safe theory of gravity will deliver. And from these, general conclusions for the physics of inflation can be drawn. *Supergravity* is the low energy effective theory of string theory. String theory has numerous scalar fields, which are assumed not to take part in the inflationary dynamics. I present an analysis in supergravity, in which I examine to what extent the new inflationary degrees of freedom can be separated from all the other degrees of freedom that are usually silently assumed not to participate in the dynamics of the inflationary period (chapter 3). It turns out that strong restrictions apply to building such models, even stronger than usually assumed.

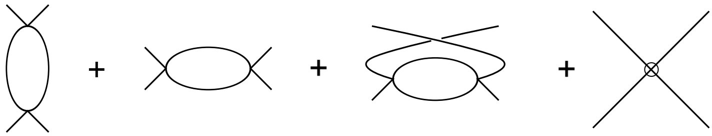


Figure 1.1: *The one-loop structure of ϕ^4 theory. The loop diagrams are UV-divergent. The divergence is cancelled by the counterterm on the right.*

1.2 Renormalization and low energy effective theories

I will begin this exposition by linking the problems of quantum field theories with the ideas of an effective field theory. A quantum field theory is generically defined perturbatively in a small coupling constant λ around some free point. The perturbative series is conveniently represented in terms of Feynman diagrams. To calculate the amplitude of a scattering process or interaction with specified in- and out-states we draw, evaluate and sum all Feynman diagrams with these states as external legs. The perturbative order in the coupling at which each diagram contributes depends on the number of loops and calculating all diagrams containing a particular number of loops means approximating the theory to that level of accuracy.

A particularly simple case is the so-called ϕ^4-theory of one scalar field with Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4 \,. \tag{1.1}$$

Like any quantum field theory, Feynman diagrams containing loops are UV-divergent. This comes from the fact that particles can go around a closed loop with any momentum, which needs to be integrated over. This can be avoided by introducing a UV cut-off, like the spacing of a lattice. If there is no natural cut-off, those divergences can be cancelled by counter-terms to obtain a finite answer. They are introduced by defining the parameters in the theory to get

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4}\phi^4 + \frac{1}{2}\delta_Z(\partial_\mu \phi)^2 - \frac{1}{2}\delta_{m^2}\phi^2 - \frac{\delta_\lambda}{4}\phi^4 \,. \tag{1.2}$$

For the one loop order, the structure of the four-field interactions is depicted in figure 1.1. These *radiative corrections* will change the value of the parameters like masses and couplings depending on the scale, at which they are measured by a specific experiment. They are called the *physical couplings*, whereas the parameters λ, m in the Lagrangian are the bare couplings, which have no real meaning. Therefore, those parameters and with them the theory needs to be defined at a specific scale. This needs to be done for every parameter and every divergence. A useful set of *renormalization conditions* is, for instance, to define

- m^2 to be the pole of the renormalized propagator and
- λ to be the 4-point amplitude of the scattering amplitude at zero momentum.

Now, summing the Feynman diagrams to any desired order while maintaining the renormalization conditions will produce a finite result which is independent of the regulator. In that way, the counterterms transform the UV divergences into scale dependence.

This procedure can be performed if the theory is (super-)*renormalizable*. This is the case if all the coupling constants of a theory have non-negative mass dimension. If there is an interaction which has a coupling constant with negative mass dimension, it is *non-renormalizable*.

When imposing a cutoff, we basically discard all the dynamics of the higher momentum contributions. Consider the generating function (or for that matter the action, the Hamiltonian or any other object which describes the full theory)

$$Z[J] = \int \mathcal{D}\phi \mathrm{e}^{\mathrm{i}\int(\mathcal{L}+J\phi)} = \left(\prod_k \int \mathrm{d}\phi(k)\right) \mathrm{e}^{\mathrm{i}\int(\mathcal{L}+J\phi)} \ . \tag{1.3}$$

The effect of a UV cutoff Λ is to set $\phi(k) = 0$ for $|k| > \Lambda$. The difference between this and the full theory is precisely the integral over the Fourier components with momenta higher than the cutoff. These modes are being *integrated out*. This can be done in several steps, lowering the cutoff more and more. If the steps are taken to be infinitesimally small, integrating out high momentum modes leads to a continuous transformation of the parameters of the theory. Thus, going from higher to lower energies introduces a flow of the coupling constants. Since the coupling constants define the theory, one can perceive this flow as a trajectory in the space of possible theories. This idea has become known as the *renormalization group (RG)* [9].

Of course, around a point where all coupling constants vanish $m^2 = \lambda = \cdots = 0$, the theory does not change any more under a scale transformation. This point is called a free *fixed point*, where the theory is scale invariant or a *conformal theory*. There are also fixed points which are not free, like the Wilson-Fisher fixed point in ϕ^4 theory.

The picture of integrating out higher momentum modes also sheds some new light on the systematics of renormalizability with different kinds of couplings in a theory. Those couplings can be seen as local operators which perturb the fixed-point Lagrangian. Operators whose coefficients grow while going down the energy scale are called *relevant* operators, because in the statistical picture they determine the low energy physics, whereas operators whose coefficients diminish are called *irrelevant* operators. An operator is relevant if its mass dimension

$$d_i = N\frac{d-2}{2} + M \ , \tag{1.4}$$

with N the number of scalar fields, M the number of derivatives in the operator and d the number of space-time dimensionsis smaller than the space-time dimension, $d_i < d$, and irrelevant if it is larger, $d_i > d$. If the mass dimension is the same as the space-time dimension, $d = d_i$, it is called *marginal*, which means that its relevance is determined by quantum corrections. Operators which are exactly marginal to all orders of perturbation theory do not perturb the theory away from a conformal point. An example of such a case is $\mathcal{N} = 4$ Super Yang-Mills theory. Note that the RG transformation is lossy and works only one way to lower energies.

It now becomes clear why a low energy theory is always fairly simple irrespective of how complicated it was in the UV. The cutoff Λ, which was originally introduced as an artificial regulator, now plays the rôle of a physical scale at which a theory contains some rich physics. As momenta k become smaller with respect to this cutoff, the bigger part of this physics scales away as $(k/\Lambda)^{d_i-d}$. At every order in $1/\Lambda$, non-renormalizable operators are introduced into the effective theory. However, since it is only valid for small energies $k \ll \Lambda$ only renormalizable operators play a rôle. Furthermore, theories within one universality class are distinguished only by irrelevant operators.

Now, a theory could start at a UV fixed point and be perturbed by some relevant deformation which, as we go down in energy, would make the theory flow away to a new IR fixed point. Again, it is not possible to go the other way round, because it is not a priori clear which of the infinitely many possible irrelevant operators to add, unless some UV symmetry restricts them. A theory is called *UV complete* or *UV safe*, if there are not any further irrelevant operators entering at some scale. The theory then captures all physics. If this is connected with a UV fixed point, the theory can either be *asymptotically free* like QCD, which means that the coupling becomes arbitrarily small at high energies, or *asymptotically safe* if the UV fixed point is not free. It is also possible that the coupling becomes infinite at a finite energy as in QED. In that case, it is clear that the theory is not UV complete but if it was, such a pole would signal that the perturbative approximation breaks down.

The way to find fixed points and determine the precise trajectory of the RG-flow in the space of possible Lagrangians is to exploit the properties of the UV divergences of the theory. Having removed such divergences by introducing counterterms and adjusting the amplitudes to match the renormalization conditions, the result is dependent on the *renormalization scale*, the momentum scale at which the conditions are applied. This dependence encodes the information of the renormalization group flow. Around a critical point with $m^2 = 0$, the renormalization conditions we have been using earlier would lead to singular counterterms. We avoid this by imposing the renormalization conditions at arbitrary space-like momenta $p^2 = -M^2$, namely

- the 2-point function vanishes at $p^2 = -M^2$,
- the derivative of the 2-point function vanishes at $p^2 = -M^2$ and
- the 4-point function is $-\mathrm{i}\lambda$ at $s = t = u = -M^2$.

Thus, the Green's functions are fixed at a certain point and UV-divergences are removed. The theory is defined *at some scale* M.

From these conditions we can now work out the flow equation of all the couplings, the so-called *Callan-Symanzik equation* [10, 11]. There is no preferred scale to define the theory and we could have just as well used $M' \neq M$ as our renormalization scale. This change of scale would only affect the renormalized Green's functions, whereas the bare theory would not see it at all. The connected n-point function in renormalized perturbation theory is

$$G^{(n)}(x_1, \dots, x_n) = \langle \psi_0 | T\phi(x_1) \dots \phi(x_n) | \psi_0 \rangle_{\text{connected}} \, . \tag{1.5}$$

Under an infinitesimal shift $M \to M + \delta M$, the couplings and the fields have to scale as

$$\lambda \to \lambda + \delta\lambda \ , \tag{1.6}$$

$$\phi \to (1 + \delta\eta)\phi \tag{1.7}$$

to keep the bare Green's function invariant. The rescaling of the fields will then introduce a shift in the renormalized Green's function

$$G^{(n)} \to (1 + n\delta\eta)G^{(n)} \ . \tag{1.8}$$

This can be written as a differential

$$\mathrm{d}G^{(n)} = \frac{\partial\, G^{(n)}}{\partial M}\delta M + \frac{\partial\, G^{(n)}}{\partial \lambda}\delta\lambda = n\delta\eta G^{(n)} \ . \tag{1.9}$$

Introducing the dimensionless parameters

$$\beta = \frac{M}{\delta M}\delta\lambda \ , \qquad \gamma = -\frac{M}{\delta M}\delta\eta \ , \tag{1.10}$$

we can write this relation as

$$\left(M\frac{\partial}{\partial M} + \beta\frac{\partial}{\partial \lambda} + n\gamma \right) G^{(n)}(x_1, \ldots, x_n, M, \lambda) = 0 \ . \tag{1.11}$$

The β-function and the anomalous scaling γ are universal for all n and independent of the coordinates. Both depend on the coupling λ. The β-function describes the running of the coupling and the anomalous dimension γ the shift in the scaling dimension. Both relate the shift of the couplings, which compensates for the shift in the renormalization scale. A vanishing β-function means that a theory is conformal, while a negative β-function means that it is asymptotically free.

A complete theory should include all fundamental forces of nature. The standard model, albeit very successful, does not contain gravity because gravity is non-renormalizable (cf. e.g. [12]). The Einstein-Hilbert action

$$\mathcal{S} = \frac{1}{16\pi G_N}\int \mathrm{d}^d x\sqrt{g}\mathcal{R}[g] \tag{1.12}$$

is derived from general coordinate covariance as a symmetry principle. The mass dimension of Newton's constant is

$$[G_E] = 2 - d \ . \tag{1.13}$$

With the redefinition $2\kappa^2 = \frac{1}{16\pi G_N}$ using linearized gravity $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$, this yields

$$\mathcal{S} = \frac{1}{2}\int \mathrm{d}^d x \left[(\partial h)^2 + \kappa(\partial h)^2 h + \ldots\right] \ , \tag{1.14}$$

which looks like a perturbation around a fixed point. Now we see that the gravitational interaction is irrelevant for $d > 2$. Therefore, gravity is non-renormalizable, which means that it can only be treated with the use of an effective theory. We will return to the question whether there might be a UV fixed point in the next section. Since the cutoff scale for gravity is the unimaginably large Planck scale M_{pl}, this almost never produces any problems. Cosmology, however, probes this energy scale and the effective field theory description will break down at the beginning of the cosmological evolution. This breakdown, when carefully approached, can teach us lessons about the new physics, which we want to discover at the Planck scale.

1.3 The Weinberg-Witten no-go theorem and holography as a way out

I have argued that in general, an effective field theory containing gravity is not UV complete, because it is non-renormalizable but that the early universe is sensitive to physics at the UV. We observe in gauge theories, though, that a gauge symmetry does not necessarily have to be present at the fundamental level in the UV. We rather see in the case of spin 1 particles, that a gauge symmetry can be an emergent phenomenon, which we only perceive in the IR at large distances. A natural idea is then that gravity could be an "emergent phenomenon", in which the graviton is a composite particle at low energies, which only appears elementary. However, it has been shown in the Weinberg-Witten theorem [13] that this is not possible. I am going to briefly outline below, why gravitons cannot emerge from a quantum field theory and how we can circumvent this no-go theorem.

We know that the graviton must be

- a massless particle, because gravity is the longest range force that we know in nature,
- a spin 2 particle, because it is sourced by the stress-energy tensor, which is second rank. Besides, it can been shown that any massless spin-2 field must interact with the stress-energy tensor just as the gravitational field and would be indistinguishable from gravity.

The Weinberg-Witten theorem states, that the assumption of a $3+1$-dimensional local QFT with a conserved and Poincaré-covariant stress-energy tensor does not admit a massless (composite) particle with helicity $|h| > 1$ and thus excludes gravitons.

The reason for this restriction lies in the nature of the stress-energy tensor. If a particle or a state with spin two or higher interacts with the stress-energy operator in its rest-frame its helicity would have to change from positive to negative, which is a total change of 4. To allow for a Lorentz-invariant spin 2 state, the stress-energy tensor would have to be a spin-4 state, but it is bounded to be maximally spin 2.

In the following, I sketch the proof of the theorem. We look at a process which measures the energy of the particle, i.e. the interaction of its state with the stress-energy operator. With p^μ the value of the null-component of the stress energy tensor $p^\mu = \int \mathrm{d}^3x T^{\mu 0}$ and E the eigenvalue for the energy operator p^0, we have

$$E\delta^3(p' - p) = \langle p', h|p^0|p, h\rangle = \int \mathrm{d}^3x \langle p', h|T^{00}(0, \vec{x})|p, h\rangle \tag{1.15}$$

$$= \int \mathrm{d}^3x \; \mathrm{e}^{\mathrm{i}(\vec{p} - \vec{p}\,') \cdot \vec{x}} \langle p', h|T^{00}(0, 0)|p, h\rangle \tag{1.16}$$

$$= (2\pi)^3 \delta^3(\vec{p}\,' - \vec{p}) \langle p', h|T^{00}(0, 0)|p, h\rangle \tag{1.17}$$

for single-particle states with momenta p, p', respectively and helicity h, and hence we have

$$\langle p, h|T^{00}(0, 0)|p, h\rangle = \frac{E}{(2\pi)^3} \; , \tag{1.18}$$

where $E \neq 0$. If the momentum transfer between two states $|p^\mu\rangle$ and $|p'^\mu\rangle$ is assumed to be space-like, i.e. such that $(p - p')$ is not null, we can always transform to a reference frame, where $p' + p$ is along the time direction such that the momentum of the ingoing particle is $(\frac{q}{2}, 0, 0, -\frac{q}{2})$ and $p' - p$ along the space direction of motion, z say, such that the momentum of the outgoing particle is $(\frac{q}{2}, 0, 0, \frac{q}{2})$. Under a rotation of an angle θ about the direction of motion, the single-particle states transform as $|p, \pm h\rangle \to \mathrm{e}^{\pm\mathrm{i}\theta h}$ and $|p', \pm h\rangle \to \mathrm{e}^{\mp\mathrm{i}\theta h}$ and thus, the left-hand side of (1.18), $\langle p, h|T^{00}(0,0)|p, h\rangle$ transforms as $\mathrm{e}^{2\mathrm{i}\theta h)\theta}$. To preserve rotational invariance, the matrix elements must transform to

$$R^\mu{}_\rho(\theta) R^\nu{}_\sigma(\theta) \langle p', \pm h|T^{\rho\sigma}|p, \pm h\rangle \; , \tag{1.19}$$

where $R(\theta)$ is the rotation matrix, which has Fourier components $\mathrm{e}^{\pm\mathrm{i}\theta}$, only. Hence, to preserve rotational invariance, the matrix elements of $T^{\mu\nu}$ must vanish unless $|h| = 0, \frac{1}{2}, 1$. For these values, the rotation coincides with the and Lorentz-invariance can be preserved. In particular, there is no spin-2 state. Therefore it is proven that gravitons cannot be described in and gravity cannot emerge from a local quantum field theory.

Another indication for this fact comes from the observation that in a local quantum field theory, the entropy scales with the volume of a system. In a gravitational theory, the bound is stronger and the entropy can only grow with the area of the boundary of the system. This is based on the argument that the entropy of a black hole is the area of its horizon measured in Planck units [14–16]

$$\mathcal{S} = \frac{A_{\text{horizon}}}{4 l_{\text{Planck}}^2} \; . \tag{1.20}$$

This as well hints to the fact that gravity cannot be obtained from local degrees of freedom.

It should be noted that a gauge symmetry never comes from a global symmetry. The gauge symmetry only arises in the IR. In gravity, space-time points themselves are not gauge invariant and are changed under a gauge transformation. Therefore, we can conclude that in any theory, in which gravity emerges, the geometry of space-time must emerge with it, just as everything has to emerge, that lives on this space-time like gauge and matter fields.

This last observation points to a possible solution of the problem: Instead of having gravity emerge from a field theory defined on the same space-time, gravity can emerge from a quantum field theory in a dimension less. The symmetries of the field theory in such a setup give rise to the isometries of the emerging space-time. In the case of a conformal field theory, which has a scaling symmetry, the emergent space-time is an anti-de Sitter space with a metric

$$\mathrm{d}s^2 = R_{\text{AdS}} \frac{\mathrm{d}z^2 + \mathrm{d}x_\mu \mathrm{d}x^\mu}{z^2} \; , \tag{1.21}$$

where R_{AdS} is the AdS radius. The scaling symmetry under $x \to \lambda x$ of the CFT leaves the metric invariant, if also $z \to \lambda z$. Since gravitational interactions decay with the distance between two particles, the Hilbert space of the theory should allow for a Fock space structure. Gauge theories with SU(N) gauge group have such a structure in their large-N limit. It has been speculated for a long time that such theories are related to string theory as a theory of gravity ([17], see [18] for a review).

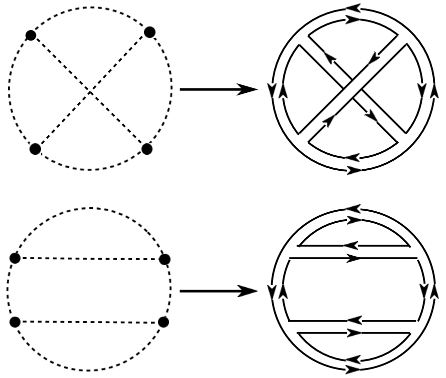


Figure 1.2: *The large-N limit in double line notation. Each line represents a fundamental or anti-fundamental index, respectively, of one gauge field. Diagrams which are non-planar have generically more factors of N for each closed index loop. In the limit $N \to \infty$, those diagrams are sub-leading and can be ignored for most purposes.*

The large-N limit uses the observation, that a gauge field in the adjoint representation of SU(N) can be recast as a direct product of a fundamental and an anti-fundamental field with separate indices. Those are depicted as separate lines adjacent to each other in a Feynman diagram. Then we observe that the Feynman diagrams of the theory fall into two separate classes (cf. figure 1.2):

1. *planar diagrams* which can be drawn on a flat surface such that none of the lines cross and
2. *non-planar diagrams* which can only be drawn on a surface with a higher genus to avoid crossing of lines.

In every closed loop, the gauge indices are not fixed and are summed over all N possibilities. Thus, such a loop contributes a factor N to the value of the diagram. Carefully counting the different combinations of powers of the gauge coupling g and of N for each closed index loop shows that we can split off an effective coupling of $\lambda_{\text{'t Hooft}} = g^2 N$, the so-called *'t Hooft coupling*. If this coupling is held fixed in a limit

$$N \to \infty \,, \qquad g^2 N = \text{constant} \,, \tag{1.22}$$

we see that

1. planar diagrams survive the large-N limit, whereas
2. non-planar diagrams are sub-bleeding in $\frac{1}{N^{2\gamma}}$ and die in this limit, where $\gamma > 0$ is the genus of the surface on which they are drawn.

The field theory is now expanded in a double series in g^2N and in $\frac{1}{N}$. The latter corresponds effectively to an expansion in the genus of the corresponding surface.

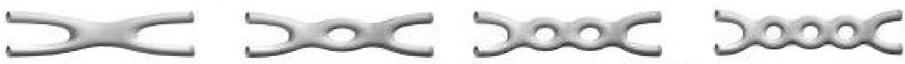


Figure 1.3: *One-dimensional strings trace out two-dimensional world-sheets in time. The scattering of closed strings is then described by the joining and separation of world tubes. A perturbative expansion in the string coupling is then a series of surfaces with higher genus. Figure adapted from [19].*

This is very resemblant of string theory, where particles are replaced by strings with an inner degree of freedom and the Feynman diagrams of QFT are replaced by world-sheet diagrams of connecting strings (see fig. 1.3). Now, if we consider a closed chain made up of particles which each carry two of the N different colors such that two adjacent particles always share a color, the probability of two such chains with the same color combination passing by each other is $\frac{1}{N}$. Only such strings would interact and the string coupling would be

$$g_S \sim \frac{1}{N} \; . \tag{1.23}$$

In the strict large-N limit of an infinite chain, this coupling vanishes and the corresponding string theory is free. This means that all the loop string scattering diagrams are suppressed and all the higher genus surfaces in fig. 1.3 do not contribute.

If we apply this observation to the interpretation of a surface with a specific genus to represent the diagrams at a specific order in a $\frac{1}{N}$ expansion of a field theory, this means that the contribution of all the non-planar diagrams vanishes. The field theory is expanded only in the finite, weak ’t Hooft coupling and its Hilbert space automatically has the Fock space structure we have required earlier. The strings, conversely, are effectively treated as point particles, whose interactions are suppressed by $\frac{1}{N}$. Of course, up to this point we do not know at all, what this string theory would be and we will have to look for a concrete realization of this idea as a quantum field theory, from which a known string theory emerges.

The best known example is the duality between $\mathcal{N} = 4$ Super-Yang-Mills theory in 4 dimensions and type IIB string theory on AdS$_5 \times S^5$ [20–22]. Here, we examine the field theory around its conformal fixed point because $\mathcal{N} = 4$ SYM is a conformal theory in four dimensions. The relation between the geometric parameters and the field theory quantities is

$$\lambda_{\text{'t Hooft}} = \left(\frac{R_{\text{AdS}}}{l_{\text{string}}}\right)^4 \tag{1.24}$$

$$g^2_{\text{YM}} = g_s \; . \tag{1.25}$$

We see that this correspondence is, indeed a duality, because in a region where the AdS curvature radius R_{AdS} is small as compared to the string length l_{string}, which means that gravity is strongly coupled, the ’t Hooft coupling $\lambda_{\text{'t Hooft}}$ is small and vice versa. Note that taking the ’t Hooft limit is essential to that observation, because a large Yang-Mills coupling g_{YM} would lead to a large string coupling g_s. In turn, if the ’t Hooft coupling is fixed, the string coupling, indeed, scales inversely with N.

This so-called *AdS/CFT correspondence* allows to describe the dynamics in a strongly coupled gravitational *bulk*, where quantum gravity effects are important, by a perturbative quantum field theory and vice versa, if the fields in the bulk are related to the operators in the boundary field theory. To perturb the theory around the conformal point, we can add single trace operators to the field theory, i.e. operators of the form

$$\mathcal{O} = g^2 \mathrm{tr}\, \phi^4 = \phi^i_j \phi^j_k \phi^k_l \phi^l_i \;, \tag{1.26}$$

where upper and lower indices are fundamental and anti-fundamental, respectively. The prescription is now, that the source of such an operator is the boundary condition for a field in the bulk.

A very intriguing feature of the AdS/CFT correspondence is that it embeds the ideas of the renormalization group (see section 1.2) in a geometric way [23–26]. It turns out that the Hamilton-Jacobi equations of supergravity take the form of the Callan-Symanzik equations of the field theory

$$\frac{1}{\sqrt{g}} \left(g^{\mu\nu} \frac{\delta}{\delta g^{\mu\nu}} + \beta^I(\phi) \frac{\delta}{\delta \phi^I} \right) \Gamma[\phi, g] = \text{4-derivative terms} \;, \tag{1.27}$$

where Γ is the non-local part of the action. The 4-derivative terms on the right hand side of this equation stem from cross-terms of the potential, functional derivatives thereof and the non-local effective action Γ, as well as curvature squared terms and products of the curvature with space-time derivatives of the scalar fields. These terms drop out upon variation of the action and do not play a rôle. With the metric $g_{\mu\nu} = a^2 \eta_{\mu\nu}$ this yields the Callan-Symanzik equation upon replacing the functional derivatives with ordinary ones by virtue of

$$\int g^{\mu\nu} \frac{\delta}{\delta g^{mn}} = a \frac{\partial}{\partial a} \;, \qquad \int \frac{\delta}{\delta \phi^I} = \frac{\partial}{\partial \phi^I} \;. \tag{1.28}$$

This is depicted in figure 1.4. If the field theory is not taken to be at the boundary but at a finite distance $z = z_{\mathrm{cutoff}}$, this theory would correspond to a renormalized version of the boundary field theory, which has some multi-trace operators added [28]. Note that the stress-energy tensor, which measures the energy, should never be renormalized.

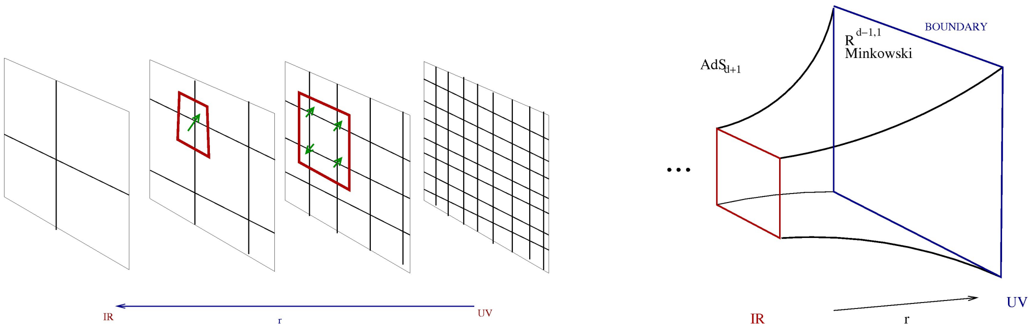


Figure 1.4: *The AdS/CFT correspondence provides a geometric picture of Wilson RG flow. The theory at the boundary is the UV version. Defining the theory at a finite distance from it corresponds to integrating out UV degrees of freedom just like in a series of block spin transformations labeled by a parameter r on the left. Figure from [27].*

CHAPTER 2

THE RESOLUTION OF COSMIC SINGULARITIES WITH THE HELP OF THE GAUGE/GRAVITY DUALITY

Our whole universe was in a hot dense state,
Then nearly fourteen billion years ago
expansion started. Wait...
The Earth began to cool,
The autotrophs began to drool,
Neanderthals developed tools,
We built a wall (we built the pyramids),
Math, science, history, unraveling the mysteries,
That all started with the big bang!

Barenaked ladies,
The Big Bang theory theme song

In this chapter, I am going to investigate the effects of quantum gravity on the *big bang singularity*. That there is a singularity at the beginning of our universe poses a severe problem to our understanding of the cosmos and of gravity. A singularity means that an important quantity of the theory, in this case the Ricci scalar, becomes infinite. Not only does the theory at hand lose its predictability, it is also impossible to impose initial conditions. In general, a singularity signals the breakdown of the approximation used for the specific problem under investigation, in this case general relativity [29, 30] as the low energy theory of UV complete gravity.

If a theory of gravity contains a singularity, there are two possible scenarios. Either, they are resolved in the full quantum theory. If string theory is such a theory, this means including higher curvature and string coupling corrections. Or the theory was ill-defined to start with, meaning that even the UV complete theory has a singularity. No consistent theory should have singularities. However, there might be different types of singularities some of which a quantum theory is not required to resolve.

Usually, unphysical solutions to the equations of motion of a theory can be ruled out by analyzing their stability. In general relativity, the potential energy is only bounded if an energy condition is applied. However, this still does not rule out singularities, such as for instance black holes or a big crunch, to form. Such singularities occur in a well-defined classical theory and therefore need to be resolved by its quantum version, whereas, if a solution is perturbatively unstable, we might have to discard it as a whole. Still, in gravity, there is no general relation between the stability of a theory and the occurrence of singularities.

The gravitational potential considered in the following is perturbatively stable. Therefore it appears as if the situation is not fundamentally flawed and we think that quantum effects will play an important rôle. Such quantum effects can be perturbative, but do not necessarily show up at the lowest order in perturbation theory, or they can even be non-perturbative.

2.1 A singularity at strongly coupled gravity

I will first present the reason for having a singularity at the beginning of the universe. Already in 1929, Edwin Hubble and Milton L. Humason realized that the universe was expanding when discovering the proportionality of the red-shift of the spectra of distant galaxies to their distance [31, 32]. They thus confirmed the conjecture, which was put forward a couple of years earlier by Georges Lemaître [33, 34], that the universe was expanding and started from a "unique atom, the atomic weight of which is the total mass of the universe". He was building his model on the solutions of Albert Einstein [35] and Willem de Sitter [36–38] to the theory of general relativity. Their solutions suffered from being unstable and only allowed for a universe expanding at a declining rate or contracting increasingly fast.

This problem was turned into what later should become the *standard model of cosmology* by Alexander Friedmann, Georges Lemaître, Howard Robertson and Arthur Walker [33, 39–42]. It is built on the *cosmological principle*, that no observer is at the center of the universe, which looks the same viewed from any point, i.e. it is isotropic and homogeneous. This leads to a maximally symmetric metric, in which a scale factor $a(t)$ accounts for the expansion or the collapse of the universe

$$\mathrm{d}s^2 = \mathrm{d}t^2 + a(t)^2 \left[\frac{\mathrm{d}r^2}{(1-kr^2)} + r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2) \right] . \tag{2.1}$$

The parameter k encodes the spatial curvature of the universe as

$$k = \begin{cases} -1 & \text{for negatively curved} \\ 0 & \text{for flat} \\ 1 & \text{for positively curved} \end{cases} \tag{2.2}$$

spatial hyper-surfaces. With this ansatz, the Einstein field equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi \mathrm{G_N} T_{\mu\nu} \ , \tag{2.3}$$

yield the Friedmann equations. The first one describes the evolution of the scale factor, the so-called *Hubble parameter*

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi \mathrm{G_N}}{3}\rho(t) - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \ , \tag{2.4}$$

where c is the speed of light. Besides, Λ denotes the *cosmological constant* related to the vacuum energy. The energy density of matter is indicated with ρ.

The Hubble parameter is currently measured to be $73.8 \pm 2.4 \frac{\mathrm{km/s}}{\mathrm{Mpc}}$ using the Wide Field Camera 3 on the Hubble space telescope [43] or $67.0 \pm 3.2 \frac{\mathrm{km/s}}{\mathrm{Mpc}}$ using the 6dF Galaxy Survey [44]. Hubble himself obtained a value of around $500 \frac{\mathrm{km/s}}{\mathrm{Mpc}}$. The Hubble time, which is H^{-1}, is the approximate age of the universe. The cosmic expansion drives anything which is further than the *Hubble radius* $\frac{c}{H}$ apart from a given observer away from it faster than the speed of light. Hence, the Hubble radius is the size of the observable universe. However, if the Hubble parameter H is not constant, the observable region changes, which is an important feature for cosmological model building.

The second Friedman equation or Raychaudhuri equation describes the acceleration of the scale factor and is derived from the spatial components of the Einstein equations

$$\frac{\ddot{a}}{a} = -\frac{4\pi \mathrm{G_N}}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \ . \tag{2.5}$$

On top, conservation of energy yields the continuity equation

$$\dot{\rho} = -3H\left(\frac{p}{c^2} + \rho\right) \ , \tag{2.6}$$

which describes the dilution of energy during the expansion. To solve the above set of cosmological equations, we need to specify the relation between the pressure p and the energy density ρ, which is the so-called *equation of state*. The different contributions to the energy in the universe are well-described by a linear equation of state $p = \mathrm{c}^2 w\rho$, which can also describe a cosmological constant as dark energy and curvature with effective pressure and energy density. Different values of w describe different such components, namely

- $w = 0$: cold matter
- $w = \frac{1}{3}$: radiation
- $w = -\frac{1}{3}$: e.g. (negative) curvature
- $w = -1$: cosmological constant/dark energy.

We can then describe a universe filled with different kinds of energy by summing up the different contributions. Thus, we rewrite the first Friedman equation in terms of the *critical density* of a flat universe

$$\rho_c = \frac{3H^2}{8\pi \mathrm{G_N}} \ , \tag{2.7}$$

such that all the contributions need to sum up to

$$1 = \sum_{\substack{i=\text{matter},\\ \text{radiation},\\ k,\Lambda}} \frac{\rho_i}{\rho_c} \; . \tag{2.8}$$

The evolution of the energy density can be determined from integrating (2.6)

$$\frac{\mathrm{d}\rho}{\rho} = -3(1+w)\frac{\mathrm{d}a}{a} \quad \Rightarrow \rho \propto a^{-3(1+w)} \quad \text{for} \quad w \neq -1 \; . \tag{2.9}$$

We see that the density will generally be diluted as the scale factor grows at later times. Note that $\rho = \text{constant}$ for $w = -1$, which means that the cosmological constant is, indeed, not diluted by the expansion as the name suggests.

In turn, when evolving the density backwards in time as the scale factor shrinks, we see that the density is bigger for earlier times, which leads to a divergence of the values in the stress-energy tensor. Correspondingly, from (2.3), also the Riemann scalar R has to diverge and we arrive at a (space-like) curvature singularity. Not only at the singularity itself, but also in its neighborhood is general relativity unpredictive. This can be seen in the linearized approach (1.14), where the coupling constant κ would have to become strong, if $h_{\mu\nu}$ is still to be considered a small perturbation. Therefore, gravity is strongly coupled around the singularity and the perturbative approximation breaks down.

The red-shifting of galactic spectra is not the only evidence for an expanding universe. The FLRW model got much more convincing support from the accidental discovery of the *Cosmic Microwave Background* (CMB) by Arno Penzias and Robert Woodrow Wilson in 1965 [45, 46]. This radiation was just around the same time discussed to be a left-over from the big bang by Robert H. Dicke, Jim Peebles and David Wilkinson [47] but had a history of being conjectured by George Gamow [48, 49], Gamow, "Hans Bethe" and Ralph Alpher [50] and the latter with Robert Herman [51]. In a theory of an expanding universe, this radiation was created about 380,000 years after the big bang. Before that time, the temperature of the universe would be too high for neutral hydrogen to exist and free electrons scatter photons very efficiently such that the early plasma was opaque. As the temperature dropped with the expansion of the universe, neutral hydrogen formed and the universe became transparent with a mean free path of photons larger than the Hubble radius. These photons were since redshifted by a factor $\frac{T_{\text{now}}}{T_{\text{recombination}}} \sim 1100$ to a temperature of about 2.725 K [52] and now form the CMB. This is one of the features of big bang cosmology, which is the most difficult to attain by alternative models.

So far, I have laid out good and generally accepted arguments supporting the idea that the beginning of the universe is a strange singular state, the big bang, which marks the beginning of what we describe as the evolution of our universe in the paradigm of general relativity. This is a frustrating situation, since this very beginning is a point of utter interest. It is a very natural question, what happened "before" the big bang and which dynamics led to the onset of the expansion at a certain point. It is here, where we need to impose boundary conditions, if there is no dynamics before to ensure, e.g. that the universe starts out in a low, even zero entropy state, to satisfy the second law of thermodynamics (Penrose in [53]). Of course, we would rather have this special state to be selected dynamically.

To understand what a low entropy state means, we first need to understand the micro-states of any theory at hand. In the case of the very early universe, where gravity is the dominating force, those would mainly be the micro-states of gravity. The notion of gravitational micro-states comes from the aforementioned observation that black holes have a finite entropy, given by the area of their horizon in Planck units as in equation (1.20). The micro-states of a system are accessible in the UV regime. In the case at hand this is a theory of quantum gravity. The observation that the effective description of gravity becomes strongly coupled and breaks down close to the singularity is another fact supporting the idea that a theory of quantum gravity should naturally solve the problems associated with a big bang singularity.

A number of different approaches to quantum gravity seem to produce the correct density of micro-states for black holes [54–56]. Under the assumption, that only the near-horizon geometry accounts for the degrees of freedom, the conditions, which need to be imposed on any such theory of quantum gravity to produce a black hole seem to enforce a conformal symmetry there. A situation, in which we understand the rôle which is played by conformal field theory and in particular the Cardy-Verlinde formula [57–59] in the counting of gravitational micro-states [60] is the AdS/CFT duality within string theory mentioned in section 1.3. It suggests holography as a way to define quantum gravity. Being a duality, it relates a strongly coupled theory on one side to a weakly coupled theory on the other. This appears very useful in our case at hand, since gravity, whose microscopic description we are after, is strongly coupled around the singularity. The corresponding quantum field theory is weakly coupled, which is very well understood perturbatively. Intuitively, general relativity is a theory valid at large scales, but close to the big bang the scales are small and a quantum field theory is a more suitable theory, there. The AdS/CFT correspondence is only well-understood for the very specific case of an $\mathrm{AdS}_5 \times S^5$ dual to an $\mathcal{N} = 4$ Super-Yang-Mills theory. However, we expect that the *holographic principle*, derived from the scaling of black hole micro-states with the area of the horizon, holds universally. If the correspondence is taken to have a universal meaning as the "gauge-gravity-duality", every well-defined gravity background should be described by some QFT. A breakdown of one side should be reflected in a breakdown of the other. Since quantum gravity needs to remove the big bang singularity from the gravitational theory, according to the duality it should be possible to describe it by some well understood (deformation of a) conformal field theory.

In the following I am first going to introduce some technical aspects of the AdS/CFT correspondence before applying it to a specific model of a space-like gravitational singularity as a model of the big bang.

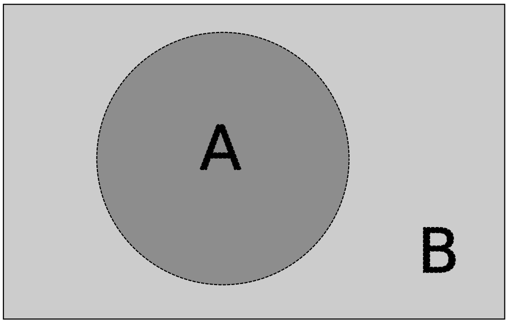


Figure 2.1: *When reducing to a subsystem, the fields become entangled over the boundary. This resembles the situation of a black hole, where the degrees of freedom inside the horizon are inaccessible to an outside observer.*

2.2 THE ADS/CFT CORRESPONDENCE

The idea that there is a duality between the microscopic theory of gravity and a quantum field theory on the boundary of the space under consideration is conceptually built on the holographic principle [61, 62]. This in turn is a conjecture which is fed by the observation that the entropy of a black hole doesn't scale with its volume as one would assume for a quantum field theory but with the area of its horizon. The horizon is a hyper-surface with one dimension less than the black hole. The entropy which can be contained in a given volume Γ is bounded by the entropy of a black hole occupying this volume. For theories of gravity in d dimensions, the entropy is hence bounded by its surface in terms of the Planck length

$$S_{\text{Bekenstein}} \leq \frac{\text{Vol}(\partial\Gamma)}{4 l_{\text{Planck}}^{d-2}} . \tag{2.10}$$

Although surprising from the point of view of statistical mechanics, where we expect entropy as an extensive quantity to grow like the volume of the system, such a scaling is well-known for *entanglement entropy*. The entanglement entropy quantifies the amount of information observers loose about a system, if they cannot access a part of it any more. It arises, because quantum states are defined globally. When projecting it onto the subsystems, the eigenstates become entangled over the boundary, which separates the two subsystems, see fig. 2.1. We reduce the system to subsystem B, say, by tracing the density matrix over the degrees of freedom of subsystem A and obtain the reduced density matrix ρ_B. The entropy, associated with this loss of information is

$$S_B = -\text{tr}\left(\rho_B \log \rho_B\right) \sim \partial A = \partial B , \tag{2.11}$$

which scales with the boundary between the two systems [63, 64]. This can be understood heuristically, because for any theory with short-ranged interactions, the biggest contribution to the entanglement comes from the states close to the boundary, the number of which is proportional to it. This also means, that the entanglement entropy is symmetric

$$S_A = S_B \ , \tag{2.12}$$

i.e. that it doesn't matter if we trace out system A or B.

This resembles the situation with a black hole insofar as an observer at the outside also looses information about the states inside the horizon. The vacuum state of all the fields including gravity inside the horizon region become entangled with the fields outside. Thus, entanglement entropy is seen to contribute at least a large part of the black hole entropy [65, 66]. Actually, equation (2.11) is divergent, unless we invoke a UV cutoff. If we interpret the Bekenstein-Hawking entropy as an entanglement entropy, we see that this cutoff is the Planck scale. This suggests, that the Planck scale is the minimal length scale up to which classical gravity can serve as a good approximation to the fundamental theory, which is in line with the expectations. The understanding of how this cutoff arises as a physical length scale is related to understanding the physics is characteristic at this scale and remains to be uncovered. The area law suggests that the horizon of a black hole is the place where the degrees of freedom of quantum gravity live and that that they are described for any system by a quantum field theory on its boundary.

The AdS/CFT correspondence in string theory is the only concrete realization that we know of the holographic principle so far. Here, string theory as a theory of quantum gravity on an Anti-de Sitter background, the *bulk*, is dual to and can be described by a quantum gauge field theory living on its asymptotic *boundary*, a flat Minkowski space of one dimension less. In the original setup, a string theory in ten dimensions is compactified on a five-sphere such that supergravity on an AdS_5 background remains, which is dual to an $\mathcal{N} = 4$ super-Yang Mills theory on four-dimensional Minkowski space [20–22]. Since then, the the correspondence has been extended to other dimensions and spaces with less symmetry [67] and new realizations are much sought after. In particular, the generalization of the duality to de Sitter space presents us with conceptual problems and is still badly understood [68]. This is why I will present an application of the correspondence to cosmology in AdS space, although our universe resembles de Sitter space [69]. The lessons to be learned about the generic behavior of quantum gravity at the beginning of the universe and about the problems of applying the correspondence to gravity at strong coupling are still invaluable.

In the remainder of this section I am going to define Anti-de Sitter space and explain some of its peculiarities, before describing the large-N limit of quantum field theories. Then I am going to sketch how to relate the two and emphasize the importance of boundary conditions for this setup. There are numerous reviews, from which the following material can be extracted [18, 70–74].

2.2.1 Anti-de Sitter space

The Einstein field equations have three classes of maximally symmetric vacuum solutions, namely flat space, positively curved space and negatively curved space (cf. e.g. [75–77]). Anti-de Sitter space (AdS) is the one with negative scalar curvature, which corresponds to having a negative cosmological constant. This is one of the reasons why it is easier to formulate quantum gravity on Anti-de Sitter than on de Sitter space, because a positive vacuum energy necessarily breaks supersymmetry, which complicates e.g. the extrapolation of black hole solutions to strong coupling.

A $d+1$-dimensional Anti-de Sitter space, AdS_{d+1}, is a homogeneous space

$$\mathrm{AdS}_{d+1} \cong \frac{\mathrm{SO}(d,2)}{\mathrm{SO}(d,1)} \tag{2.13}$$

and forms the Lorentzian analogue of a hyperbolic space. Therefore, AdS_{d+1} can be embedded into a $d+2$ dimensional space-time $\mathbb{R}^{2,d}$ as as a quadratic surface

$$-(X^0)^2 + (X^1)^2 + (X^2)^2 + \cdots + (X^d)^2 - (X^{d+1})^2 = -1 \; . \tag{2.14}$$

This quadric is invariant under the $\mathrm{SO}(d,2)$ isometries of the embedding manifold and therefore maximally symmetric. The flat metric on $\mathbb{R}^{2,d}$

$$\mathrm{d}s^2 = -(\mathrm{d}X^0)^2 + (\mathrm{d}X^1)^2 + (\mathrm{d}X^2)^2 + \cdots + (\mathrm{d}X^d)^2 - (\mathrm{d}X^{d+1})^2 \tag{2.15}$$

induces a Lorentzian metric on the hyperboloid with a scalar curvature

$$R = -d(d+1) \; . \tag{2.16}$$

The induced metric solves Einstein's equations with a negative cosmological constant

$$\Lambda = -\frac{d(d-1)}{2} \; . \tag{2.17}$$

For an arbitrary curvature radius, the ambient metric is rescaled to

$$\mathrm{d}s^2 = -R_{\mathrm{AdS}}^2 \left[(\mathrm{d}X^0)^2 + (\mathrm{d}X^1)^2 + (\mathrm{d}X^2)^2 + \cdots + (\mathrm{d}X^d)^2 - (\mathrm{d}X^{d+1})^2\right] \tag{2.18}$$

such that the scalar curvature will change to

$$R = -\frac{d(d+1)}{R_{\mathrm{AdS}}^2} \; . \tag{2.19}$$

In the following it will thus be sufficient to focus on the embedding as a unit hyperboloid.

We can use various coordinate systems on the hyperboloid (2.14), which differ in terms of the resulting metric and the amount of AdS space they cover. The so-called global coordinates cover all of AdS_{d+1} and are defined by

$$X^0 = \cosh\mu \cos t \; , \tag{2.20}$$

$$X^i = \sinh\mu \, \omega^i \; , \quad i = 1, \ldots d \; , \tag{2.21}$$

$$X^{d+1} = \cosh\mu \sin t \; , \tag{2.22}$$

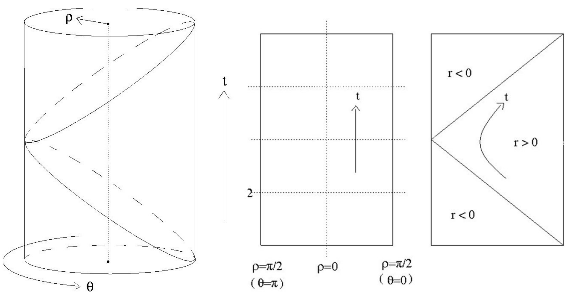


Figure 2.2: *AdS space can be represented as the interior of a cylinder (left). Top and bottom of this cylinder are identified and closed time-like curves arise. For the universal covering, an infinite number of cylinders is glued on top of each other. The boundary of the cylinder represents the boundary of AdS, where the dual field theory lives. The Penrose diagram of Anti-de Sitter space reveals the causal structure. Since it is a conformal projection, angles and thus the light-cones remain unchanged. A curious fact is that light rays can reach the boundary in finite global time, although it is infinitely far away. There are several useful coordinate systems such as global coordinates (middle) or Poincaré coordinates (right). One set of Poincare coordinates covers each of the diamond shaped regions, which are delimited in the figure on the left and which each are conformal to flat Minkowski space. Figure from [78]*

where $\omega^i = \sin\theta_i \ldots \sin\theta_{i-1}\cos\theta_i$ is a unit d-vector on the d-sphere and the range of the other coordinates is

$$0 \leq \mu \leq \infty \,, \tag{2.23}$$

$$0 \leq t \leq 2\pi \,. \tag{2.24}$$

The AdS$_{d+1}$ metric in global coordinates reads

$$\mathrm{d}s^2 = -\cosh^2\mu \mathrm{d}t^2 + \mathrm{d}\mu^2 + \sinh^2\mu \mathrm{d}\Omega_{d-1}^2 \,, \tag{2.25}$$

where $\mathrm{d}\Omega_{d-1}$ denotes the line element of a unit $(d-1)$-sphere. The topology of this space is $S^1 \times S^d$, since in the $d+1$ dimensional embedding, the time coordinate is required to be periodic. This implies the existence of unphysical closed time-like curves. Therefore, we need to unwrap the time circle and consider $t \in \mathbb{R}$. Then AdS space is the universal cover of the Lorentzian space defined by (2.14). The causal structure of AdS space is best understood by drawing its *Penrose diagram*, which is a finite conformal projection of the full space (see fig. 2.2). To obtain it, we introduce the tortoise radial coordinate

$$\sinh\mu = \tan\rho \,, \quad 0 \leq \rho \leq \frac{\pi}{2} \,. \tag{2.26}$$

With this substitution, the metric reads

$$\begin{aligned}\mathrm{d}s^2 &= -\sec^2\rho \mathrm{d}t^2 + \sec^2\rho \mathrm{d}\rho^2 + \tan^2\rho \mathrm{d}\Omega_{d-1}^2 & (2.27)\\ &= \tfrac{1}{\cos^2\rho}\left(-\mathrm{d}t^2 + \mathrm{d}\rho^2 + \sin^2\rho \mathrm{d}\Omega_{d-1}^2\right) \,. & (2.28)\end{aligned}$$

In these coordinates, the boundary lies at $\rho = \pi/2$, which corresponds to $r = \tan\rho = \infty$, and has the topology of a $S^{d-1} \times R$. When suppressing all but one angular coordinates, the Penrose diagram is a full cylinder, which can be projected onto an infinite stretch, when suppressing also the last angular coordinate. Surprisingly, null geodesics can reach the boundary in finite global time and return to where they started.

For our purposes, the relevant parametrization are *Poincaré coordinates*, which slice the hyperboloid with hyperplanes given by

$$X^0 - X^d = \frac{1}{z} , \quad z > 0 , \tag{2.29}$$

$$X^i = \frac{x^i}{z} , \quad i = 1, \ldots, d-1 , \tag{2.30}$$

$$X^{d+1} = \frac{x^0}{z} , \tag{2.31}$$

where we take $z = 1/r$. We obtain the metric

$$\mathrm{d}s^2 = \frac{\mathrm{d}z^2 + \eta_{ab}\mathrm{d}x^a\mathrm{d}x^b}{z^2} , \quad a, b = 0, \ldots, d-1 , \tag{2.32}$$

with η_{ab} the Minkowski metric. The position of the boundary is now at $z = 0$. These coordinates make the Poincaré symmetry manifest

$$x^a \to \Lambda^a{}_b x^b + b^a , \quad \Lambda \in \mathrm{SO}(1, d-1) . \tag{2.33}$$

The full group of $\mathrm{SO}(2, d)$ isometries is realized by the inversions

$$z \to \frac{z^2}{z^2 + \eta_{ab}x^a x^b} , \tag{2.34}$$

$$x^a \to \frac{x^a}{z^2 + \eta_{ab}x^a x^b} \tag{2.35}$$

and the dilations

$$z \to cz , \qquad x^a \to cx^a . \tag{2.36}$$

Poincaré coordinates make the connection to a field theory on the boundary Minkowski space more tangible. Via a conformal rescaling of the metric $g_{\mu\nu} \to z^2 g_{\mu\nu}$, such that

$$\mathrm{d}\tilde{s}^2 = \mathrm{d}z^2 + \eta_{ab}\mathrm{d}x^a\mathrm{d}x^b , \tag{2.37}$$

it can now also be seen that $z = 0$ is the *conformal boundary*. Note that the Poincaré patch is geodesically incomplete, because $X^0 - X^d > 0$, which means in global coordinates

$$\cosh\mu\cos t - \sinh\mu\sin\theta_1 \ldots \sin\theta_{d-1}\cos\theta_d > 0 . \tag{2.38}$$

Both null- and time-like geodesics can reach $z = \infty$ at a finite value of their affine parameter and escape from the Poincaré patch.

We are now going to investigate some general properties of this conformal boundary and which general properties we can derive for a field theory living on it.

2.2.2 The large-N limit of boundary quantum field theories

As predicted by the holographic principle, the quantum gravity in this AdS bulk should be described by a quantum field theory in one dimension less. I will now review a couple of general properties, that this dual theory must have on general grounds. Already in section 1.3, we have seen that the geometry of higher loop string scattering worldsheet diagrams resembles the topological properties of the $\frac{1}{N}$ expansion around the 't Hooft limit. It turns out that the boundary theory must be

1. conformal,
2. a large-N limit of a
3. gauge field theory,
4. which is dual in the sense that it relates strong to weak coupling.

I will reason on physical grounds why this is the expectation.

1. We have already established *conformality* of the boundary in the previous subsection 2.2.1. The isometries of the bulk act on the boundary as the group of conformal transformations in d dimensions

$$x^a \to \Lambda^a{}_b x^b + b^a \qquad \text{(Poincaré)}\ , \tag{2.39}$$

$$x^a \to c x^a \qquad \text{(dilatations)}\ , \tag{2.40}$$

$$x^a \to \frac{x^a}{x^2} \qquad \text{(inversions)}\ . \tag{2.41}$$

   These transformations leave the boundary invariant and henceforth, the field theory living thereon must be conformal.

2. The theory must allow for a *large-N limit*. This will be important to match the parameters of the gravitational theory to the ones of the gauge theory. At the core of the construction is the attempt to explain the black hole entropy as the entropy of the field theory in one dimension less. A Schwarzschild black hole in AdS space has the metric

$$\mathrm{d}s^2 = R_{\mathrm{AdS}}^2 \left(-f(r)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f(r)} + r^2 \mathrm{d}\Omega_{d-1}^2 \right) \tag{2.42}$$

$$f(r) = r^2 + 1 - \frac{2\widetilde{\mathrm{G}}_{\mathrm{N}} m}{r^{d-2}}\ , \tag{2.43}$$

   where

$$\widetilde{\mathrm{G}}_{\mathrm{N}} = \frac{{\mathrm{G}_{\mathrm{N}}}^{d+1}}{R_{\mathrm{AdS}}^{d-1}} \tag{2.44}$$

   is the (dimensionless) effective gravitational coupling at the scale of the AdS radius. The Schwarzschild radius of a large black hole is located at $r_s \sim \mathrm{G}_{\mathrm{N}} m$ so that its entropy scales with the surface

$$\mathcal{S} \sim \frac{r_s^{d-1}}{\widetilde{\mathrm{G}}_{\mathrm{N}}} \sim \frac{1}{\widetilde{\mathrm{G}}_{\mathrm{N}} \beta^{d-1}}\ , \tag{2.45}$$

where in the last part we have written the entropy in terms of the Hawking temperature $\beta \propto \frac{1}{r_s}$. This is the entropy expected for a (conformal) field theory in d dimensions with massless scalar fields at a temperature β

$$S \propto c\frac{V_{d-1}}{\beta^{d-1}} , \tag{2.46}$$

where c measures the effective number of fields. Comparing the gravitational entropy of a black hole with the entropy of the field theory, we fix this coefficient to be

$$c \sim \frac{1}{\widetilde{\mathrm{G}}_{\mathrm{N}}} = \frac{R_{\mathrm{AdS}}^{d-1}}{\mathrm{G_N}^{d+1}} . \tag{2.47}$$

This relation means that the effective number of fields is inversely proportional to the effective gravitational coupling at the AdS scale. Therefore, if the duality is to be defined (also) for a weakly coupled bulk, the corresponding field theory must allow for a large number of fields. In the very classical limit with $\widetilde{\mathrm{G}}_{\mathrm{N}} \approx 0$, the number of fields would be infinite.

3. The 't Hooft limit for quantum *gauge theories* provides a well defined prescription for how to achieve this. In an $\mathrm{SU}(N)$ gauge theory with the fields in the adjoint representation, the 't Hooft coupling
$$\lambda = g^2 N = \text{constant} \tag{2.48}$$
remains constant in the limit $N \to \infty$ and perturbative field theory remains valid also for large N, if $\lambda \ll 1$. The Hilbert space of such theories naturally has a Fock space structure, for which the energy of a multi-particle state is proportional to the sum of the energies of single-particle states up to small corrections. This is an important feature of weakly coupled theories, which the dual field theory needs to inherit.

   A gauge invariant local operator can be built by taking the trace over fundamental fields, like $\mathrm{tr}\, F_{\mu\nu}(x)F^{\mu\nu}(x)$. A product of two of such operators would be a multi-trace operator. The scaling dimension of a multi-trace operator is the sum of the dimensions of its constituent single trace operators up to $\frac{1}{N^2}$ corrections which are negligible in the large-N limit. Expanding the gauge theory in both λ and $\frac{1}{N}$ organizes the Feynman diagrams according to their genus which corresponds to a fixed order in $\frac{1}{N}$. The expansion of the field theory in the genus of a manifold resembles the world sheet of the loop expansion of a closed string with string coupling g_{s}. Adding an extra genus would correspond to an extra loop order of a string theory with coupling

$$g_{\mathrm{s}} \sim \frac{1}{N} . \tag{2.49}$$

   The classical limit of this string theory as a quantum theory corresponds to taking $g_{\mathrm{s}} = 0$ and therefore, having a classical limit for a holographic theory of quantum gravity suggest to use a large-N gauge theory.

4. Another condition for having classical gravity as a limit is that the graviton can be treated as a point-like particle. In a string theory, as suggested by the previous argument, the

graviton is the lowest oscillation mode of a string. Generically, a string theory will also have a tower of massive states with spin > 2, which we ignore in the classical limit. If the string is seen as an extended object, the graviton has the size of the string length l_s. The characteristic size of the AdS space needs to be large compared to the string length

$$\frac{R_\mathrm{AdS}}{l_\mathrm{s}} \gg 1 \ . \tag{2.50}$$

for the graviton to be point-like. In the dual gauge theory, higher spin operators have small scaling dimension at weak coupling. The bulk mass of the corresponding particles is then comparable to the inverse AdS radius R_AdS^{-1} and are therefore rather light in the point-like approximation and render classical gravity invalid. Therefore, the coupling of the dual gauge theory must be strong enough to render the masses of the higher spin states large in the limit where gravity is classically weakly coupled.

Having argued these properties on general grounds, it seems most natural to realize the holographic principle in terms of a string theory dual to the large-N limit of an $\mathrm{SU}(N)$ gauge theory. In [20–22], such a realization has been constructed, which I will describe in the next subsection.

2.2.3 The relation between the type IIB action and Super-Yang-Mills theory

The best understood instance of the gauge/gravity duality and the one I am going to use in this thesis is the correspondence between type IIB string theory on AdS$_5\times$ S^5 and $\mathcal{N} = 4$ superconformal Yang-Mills theory with an SU(N) gauge group in 4 dimensions. We have already argued the general properties that this gauge theory needs to have in the previous subsection. We are now giving a concrete example and explain how it fulfills the above conditions.

Quantum field theories on four-dimensional Minkowski space are usually not conformal. We can adapt a theory which is known in nature, quantum chromodynamics, which is an asymptotically free gauge theory with an SU(3) gauge group to our purposes. First, we will generalize the gauge group to SU(N) such that the theory has N colors with the gauge field in the adjoint representation. Also, we make the theory maximally, i.e. $\mathcal{N} = 4$, supersymmetric. This is maximal supersymmetry with four fermions χ_α and six scalars ϕ^I all in the adjoint representation [79, 80]. The Lagrangian is uniquely determined by super- and gauge symmetry to be

$$\mathcal{L}_\mathrm{SYM} = -\frac{1}{4g_\mathrm{YM}^2}\int \mathrm{d}^4x \ \mathrm{tr}\left[F^2 + 2(D_\mu\phi^I)^2 + \chi\Gamma^a D_a\chi + \chi\gamma^i[\phi_i,\chi] - \sum_{I,J}[\phi^I,\phi^J]\right] \tag{2.51}$$

$$+\frac{\theta}{8\pi^2}\int \mathrm{tr}\, F\wedge F \ . \tag{2.52}$$

The two free parameters are the coupling constant g_YM and the angle θ. The theory is conformal due to the non-renormalizations coming from supersymmetry. The 't Hooft coupling is defined as

$$\lambda = g_{\rm YM}^2 N \ . \tag{2.53}$$

It can be seen as the effective coupling of the theory. All the fields in the adjoint representation are $N \times N$ matrices with one fundamental and one anti-fundamental index. For two fields for which one of their indices are contracted, i.e. for which a color and an anti-color are entangled, there are N color degrees of freedom which can still be exchanged between them. This factor shows up as a closed index loop in the Feynman diagrams. The theory has a global SO(6) or SU(4) R-symmetry, which rotates the six scalars or the 4 fermions into each other. It does not commute with supersymmetry, since the fermions and bosons are in different representations of the R-symmetry group.

To find the corresponding theory of quantum gravity we now to match the parameters and symmetries of the field theory to a bulk theory. I have already argued in the previous section, that this is expected to be a string theory. Supersymmetric string theories are naturally living in ten dimensions. However, according to the holographic principle, we are looking for a five-dimensional theory, such that we have to compactify five dimensions.

A string theory, which contains only closed strings and reduces to a well defined supersymmetric theory of gravity at large distances is type IIB string theory with type IIB supergravity as its low energy effective theory [81]. Supersymmetry requires that this theory contains some massless fields besides the metric, in particular a five-form field strength F_5, which is completely anti-symmetric in all its indices and constrained to be self-dual $F_5 = \star F_5$ and a dilaton φ and the axion χ. The action of this theory is

$$\mathcal{S} = \frac{1}{(2\pi)^7 l_{\rm Planck}^8} \int {\rm d}^{10}x \sqrt{g}(R + F_5^2) + \dots \ , \tag{2.54}$$

where the Planck length is related to the string length and coupling

$$l_{\rm Planck} = g_s^{\frac{1}{4}} l_s \ . \tag{2.55}$$

The string coupling is related the vacuum expectation value of the dilaton

$$g_s = \langle {\rm e}^{\varphi} \rangle \ . \tag{2.56}$$

The equations of motion admit solutions of the form $\mathrm{AdS}_5 \times \mathrm{S}^5$, which provides the desired five-dimensional AdS factor. These solutions have a five-form flux along both directions with both electric and magnetic fields. Due to the Dirac quantization condition, the flux of F_5 over the sphere is quantized

$$\int_{S^5} F_5 \propto N \ . \tag{2.57}$$

The number of flux quanta N is the same as the number of colors in the gauge theory. The equations of motion give a relation between the rank of the gauge group and the radius of the AdS_5 and S^5 as

$$R_{\rm AdS}^4 = 4\pi N l_{\rm Planck}^4 = 4\pi g_s N l_s^4 \ . \tag{2.58}$$

Since we are interested in relating the five-dimensional AdS space to the holographic theory, we dimensionally reduce over the five-sphere. The dimensionally reduced action is

$$\mathcal{S}_{\text{IIB}} = \frac{2R_{\text{AdS}}^5 \text{Vol}(S^5)}{(2\pi)^7 l_{\text{Planck}}^8} \int \mathrm{d}^5x\sqrt{-g}\left(\frac{R^{(5)}}{2} + \frac{6}{R_{\text{AdS}}^2}\right) + \dots \tag{2.59}$$

$$= \frac{N^2}{4\pi^2} \int \mathrm{d}^5x\sqrt{-\overline{g}}\left(\frac{\overline{R}^{(5)}}{2} + 6\right) + \dots \,. \tag{2.60}$$

The boundary of this compactified space is four-dimensional Minkowski, indeed, since the conformal rescaling of the metric

$$\mathrm{d}\hat{s}^2 = \mathrm{d}z^2 + \eta_{ab}\mathrm{d}x^a\mathrm{d}x^b \tag{2.61}$$

shrinks the radius of the S^5 to zero at the boundary.

We now relate the dimensionless parameters of the field theory and the gravity theory. The axion is related to the angle in the field theory and the Yang-Mills coupling is related to the string coupling

$$g_{YM}^2 = 4\pi g_{\text{s}} \;, \qquad \theta = \langle\chi\rangle \;. \tag{2.62}$$

Also the global symmetries of both theories match as required, namely

- $\text{SO}(2,4)$ is the conformal group and the isometry group of AdS_5.
- There are 32 real supercharges in the field theory and $\text{AdS}_5 \times \text{S}^5$ is a maximally symmetric solution of $\mathcal{N}=2$ type IIB supergravity.
- The $\text{SO}(6)$ R-symmetry is the isometry group of the S^5.

Contemplating once more on the relation between the couplings in the two theories

$$\frac{R_{\text{AdS}}}{l_{\text{s}}} = g_{\text{YM}}^2 N = g_{\text{s}} = \lambda_{\text{'t Hooft}} \;, \tag{2.63}$$

we see that this correspondence, indeed, satisfies the general principles explained in section 2.2.2. If we take $N \to \infty$ at $\lambda_{\text{'t Hooft}} \ll 1$ fixed, the string coupling vanishes, $g_{\text{s}} \to 0$. In the field theory's perturbative expansion only planar diagrams survive, whereas in the string theory higher genus contributions to the string scattering amplitude vanish. For a weakly coupled bulk theory, we need a large rank of the gauge group $N \gg 1$. For Einstein gravity to be trustworthy, the effective coupling must also be large and we have two regimes

- $\lambda_{\text{'t Hooft}} \gg 1$: classical gravity is valid, field theory is strongly coupled;
- $\lambda_{\text{'t Hooft}} \ll 1$: classical gravity is invalid, field theory is weakly coupled;

Thus, for each regime, there is a predestinated theory, in which to do the calculations. Either, a well-established theory of gravity describes the non-perturbative regime of a quantum field theory or a well-understood perturbative quantum field theory sheds light on gravity beyond the classical approximation. The latter case in an application cosmology is of interest in this thesis.

2.2.4 THE IMPORTANCE OF BULK BOUNDARY CONDITIONS

What makes the AdS/CFT correspondence so powerful is that the equivalence holds at the full quantum level. This equivalence is encoded in the equivalence of the partition functions of the bulk and boundary theories

$$\mathcal{Z}_{\text{gravity}}[\phi_0(x)] = \mathcal{Z}_{\text{field theory}}[\phi_0(x)] \tag{2.64}$$

which is a central statement to the duality. From it, the AdS/CFT dictionary can be derived. It states that to each field in the bulk corresponds a primary operator in the field theory.

The boundary of the space-time is where the field theory lives. The boundary data of the bulk fields are sources of gauge invariant operators of the field theory. Therefore, the boundary conditions, which we impose on the bulk (scalar) fields, determine the solution in the bulk, which corresponds to picking specific operators in the field theory. Hence, for each field, we impose a boundary condition

$$\phi(0, x^a) \sim \phi_0(x^a) \ . \tag{2.65}$$

Then the partition function of the bulk theory, subject to the boundary conditions, is identical to the generating functional of the field theory with the boundary values of the fields as sources

$$\mathcal{Z}_{\text{bulk}}[\phi_0] = \langle \mathrm{e}^{\int \mathrm{d}^4 x \phi_0(x) \mathcal{O}(x)} \rangle_{\text{boundary}} \ . \tag{2.66}$$

This relation can be employed to concretely obtain correlators on either side from the other by functional differentiation. It becomes particularly predictive and useful in the 't Hooft limit, where $g_S \to 0$ and the bulk is essentially classical. The path integral determining the partition function $\mathcal{Z}_{\text{bulk}}$ is then largely dominated by the solution to the classical field equations.

To make the effect of boundary conditions concrete, let us consider the Euclidean action of a free bulk scalar field of mass m

$$I = \int \mathrm{d}z \mathrm{d}^4 x \sqrt{g} \left[\frac{1}{2}(\partial\phi)^2 + \frac{1}{2} m^2 \phi^2 \right] \ . \tag{2.67}$$

Connected correlation functions in the boundary theory can be computed by functional differentiation of the bulk partition function

$$\mathcal{Z}_{\text{bulk}}[\phi_0] = \mathrm{e}^{W[\phi_0]} \tag{2.68}$$

$$\langle \mathcal{O}(x_1) \dots \mathcal{O}(x_n) \rangle_{\text{connected}}[\phi_0] = \left. \frac{\delta^n W}{\delta\phi_0(x_1) \dots \delta\phi_0(x_n)} \right|_{\phi_0} \ , \tag{2.69}$$

where

$$W_{\text{bulk}}[\phi_0] = I_{\text{on-shell}}[\phi_0] + \text{quantum corrections} \ . \tag{2.70}$$

where $I_{\text{on-shell}}[\phi_0]$ is the Euclidean action of the classical solution satisfying the boundary conditions set by ϕ_0. Imposing a finite boundary condition corresponds to adding a source term to the boundary action

$$I_{\text{CFT}} \to I_{\text{CFT}} - \int \mathrm{d}^4 x \phi_0(x) \mathcal{O}(x) \ . \tag{2.71}$$

The linearized field equations in Poincaré coordinates show that the general solution behaves near the boundary as

$$\phi|_{z\to 0} \sim \left(\alpha(x)z^{4-\Delta} + \beta(x)z^{\Delta}\right)\left(1 + \mathcal{O}(z^2)\right) , \tag{2.72}$$

$$\Delta = 2 + \nu , \qquad \nu = 4 + m^2 . \tag{2.73}$$

Here, α and β are the mode functions defined on a Ponicaré slice $z = \text{constant}$. We set the boundary conditions for the mode α, which diverges near the boundary

$$\alpha(x) = \phi_0(x) . \tag{2.74}$$

The solution to the free field equation respecting this condition is

$$\phi(z,x) = \frac{(\Delta-1)(\Delta-2)}{\pi^2}\int \mathrm{d}^4y \left(\frac{z}{z^2+(x-y)^2}\right)^{\Delta} \phi_0(y) . \tag{2.75}$$

From that expression, the value of the decaying mode β can be extracted

$$\beta(x) = \frac{(\Delta-1)(\Delta-2)}{\pi^2}\int \mathrm{d}^4y \frac{\phi_0(y)}{(x-y)^{2\Delta}} . \tag{2.76}$$

The on-shell action of the boundary theory is defined by an integral on the AdS boundary. It is divergent and needs to be renormalized by introducing suitable boundary counterterms, which do not affect the bulk equations of motion. The finite part of the action is

$$I_{\text{on-shell}} = -\frac{(\Delta-1)(\Delta-2)}{\pi^2}\int \mathrm{d}^4x\mathrm{d}^4y \frac{\phi_0(x)\phi_0(y)}{(x-y)^{2\Delta}} . \tag{2.77}$$

Let us now examine the one- and two-point functions of the boundary theory. The vacuum expectation value is calculated to be

$$\langle\mathcal{O}(x)\rangle = \frac{2(\Delta-1)(\Delta-2)^2}{\pi^2}\int \mathrm{d}^4y \frac{\phi_0(y)}{(x-y)^{2\Delta}} = 2\nu\beta(x) . \tag{2.78}$$

We see that the normalizable mode corresponds to the VEV of the dual field theory operator. This must vanish in the absence of sources not to break conformal invariance. The two-point function

$$\langle\mathcal{O}(x)\mathcal{O}(y)\rangle_{\text{connected}} = \frac{2(\Delta-1)(\Delta-2)^2}{\pi^2}\frac{1}{(x-y)^{2\Delta}} \tag{2.79}$$

has the form expected for a two-point function of an operator with dimension Δ in a conformal field theory. Hence, the mass of the bulk (scalar) field determines the dimension of the operator.

We can even generalize the boundary conditions, which we impose by deforming the field theory action [82, 83] and re-write (2.71) as

$$I_{\text{CFT}} = I_{\text{CFT}} + 2\nu W[\hat{\beta}(x)] , \tag{2.80}$$

$$\hat{\beta}(x) = \frac{\mathcal{O}(x)}{2\nu} , \tag{2.81}$$

where

$$W[\hat{\beta}(x)] = \int \mathrm{d}^4x \ \phi_0(x)\hat{\beta}(x) \ . \tag{2.82}$$

The bulk boundary condition are then re-written as

$$\alpha(x) = -\frac{\delta W}{\delta\beta(x)} \ . \tag{2.83}$$

This is a way of imposing general, non-linear boundary conditions.

Such generalized bulk boundary conditions will generically correspond to multi-trace deformations of the field theory. They can also be imposed as field equations, if suitable boundary terms are added to the bulk action. The field equations of a scalar field in $d+1$ dimensional AdS space

$$\Box_x\phi - \frac{d-1}{z}\partial_z\phi + \partial_z^2\phi = \frac{m^2}{z^2}\phi \tag{2.84}$$

lead to the general asymptotic behavior of the field

$$\phi|_{z\to 0} \sim \left(\alpha(x)z^{d-\Delta} + \beta(x)z^{\Delta}\right)\left(1+\mathcal{O}(z^2)\right) \ , \tag{2.85}$$

$$\Delta = \frac{d}{2} + \nu \ , \qquad \nu = \sqrt{\frac{d^2}{4} + m^2} \ , \tag{2.86}$$

where we choose the larger root for ν. We see that Δ is real even for fields with a negative mass squared, as long as the inequality $m^2 \geq m^2_{\mathrm{BF}} = -\frac{d^2}{4}$ is obeyed, which is called the Breitenlohner-Friedman bound [84]. Such fields do not destabilize AdS space. For a scalar field which saturates the bound $m^2 = m^2_{\mathrm{BF}}$, the two roots of ν are the same and the asymptotic behavior of the field is

$$\phi|_{z\to 0} \sim \left(-\alpha(x)z^{\frac{d}{2}}\log z + \beta(x)z^{\frac{d}{2}}\right)\left(1+\mathcal{O}(z^2)\right) \ . \tag{2.87}$$

The two variables α and β are canonically conjugate and have the interpretation as the source and expectation value (2.78) of the scalar field.

For the example at hand, we can focus on this latter case where $m^2 = m^2_{\mathrm{BF}}$, which is a slight variation on the general case with $m^2 > m^2_{\mathrm{BF}}$. To impose the boundary condition $\alpha(x) = \phi_0(x)$, we have to add a boundary term to the bulk action

$$I_b[\phi_0] = \frac{\pi^3 R^8_{\mathrm{AdS}}}{2\kappa^2}\int_{z=\epsilon} \mathrm{d}^dx z^{-d}\left(\frac{\Delta}{2}\phi^2 - z^{d-\Delta}\phi_0\phi\right) \ , \tag{2.88}$$

and we know from (2.58) that $R^8_{\mathrm{AdS}} \propto N^2$ in the context of the $\mathrm{AdS}_5/\mathrm{CFT}_4$ correspondence [22]. This leads to the general solution

$$\phi(z,x) = \frac{\Gamma(\Delta+1)}{d\pi^{\frac{d}{2}}}\int \mathrm{d}^dy \left(\frac{z}{z^2+(x-y)^2}\right)^{\Delta}\phi_0(y) \ . \tag{2.89}$$

For the boundary term to vanish in general, the classical solution must satisfy

$$z^{-\Delta}(z\partial_z - \Delta)\phi\Big|_{z=\epsilon} = -\phi_0 \ , \tag{2.90}$$

which reduces to $\alpha = \phi_0$ when taking the cut-off $\epsilon \to 0$. In the large-N limit, the equivalence of the partition functions can be re-stated as

$$\langle \mathrm{e}^{-I_b[\phi_0]} \rangle_{\text{bulk}} = \langle \mathrm{e}^{\int \phi_0(x)\mathcal{O}(x)} \rangle_{\text{boundary}} \ . \tag{2.91}$$

The complete action can be written such that only the boundary term is non-vanishing on-shell

$$I + I_b[\phi_0] = -\frac{N^2}{2}\int_\epsilon^\infty \mathrm{d}z\mathrm{d}^d x z^{-d+1}\phi\left(\Box_x - \frac{d-1}{z}\partial_z + \partial_z^2 - \frac{m^2}{z^2}\right)\phi \tag{2.92}$$

$$+ N^2 \int_{z=\epsilon} \mathrm{d}^d x z^{-d}\phi\left(-\frac{1}{2}z\partial_z\phi + \frac{\Delta}{2}\phi - z^{d-\Delta}\phi_0\right) \ . \tag{2.93}$$

Using the boundary condition (2.90), the on-shell action is

$$(I + I_b[\phi_0])_{\text{on-shell}} = -\int_{z=\epsilon} \mathrm{d}^d x z^{-\Delta}\phi_0\phi \ . \tag{2.94}$$

We then use the general solution (2.89) and remove the regulator $\epsilon \to 0$

$$(I + I_b[\phi_0])_{\text{on-shell}} = -N^2\frac{\Gamma(\Delta+1)}{2d\pi^{\frac{d}{2}}}\int \mathrm{d}^d x \mathrm{d}^d y \frac{\phi_0(x)\phi_0(y)}{(x-y)^{2\Delta}} \ . \tag{2.95}$$

From that, we obtain the expectation value

$$\langle \mathcal{O}(x) \rangle = N^2\frac{\Gamma(\Delta+1)}{d\pi^{\frac{d}{2}}}\int \mathrm{d}^d x \frac{\phi_0(x)}{(x-y)^{2\Delta}} = N^2\beta(x) \ . \tag{2.96}$$

For the field theory, we consider a deformation of the Euclidean action by a generic functional of $\mathcal{O}$

$$I_{\text{CFT}} \to I_{\text{CFT}} + N^2 W[\hat{\beta}] \ , \qquad \hat{\beta} = \frac{\mathcal{O}(x)}{N^2} \ . \tag{2.97}$$

For the partition functions, this deformation leads to

$$\langle \mathrm{e}^{-N^2 W[\hat{\beta}]}\mathrm{e}^{\int \phi_0(x)\mathcal{O}(x)} \rangle_{\text{boundary}} = \tag{2.98}$$

$$\mathrm{e}^{-N^2 W\left[N^2\frac{\delta}{\delta\phi_0(x)}\right]}\langle \mathrm{e}^{\int \phi_0(x)\mathcal{O}(x)} \rangle_{\text{boundary}} = \mathrm{e}^{-N^2 W\left[N^2\frac{\delta}{\delta\phi_0(x)}\right]}\langle \mathrm{e}^{-I_b} \rangle_{\text{bulk}} \tag{2.99}$$

$$= \langle \mathrm{e}^{-N^2 W\left[z^{-\Delta}\phi(\epsilon,x)\right] - I_b} \rangle_{\text{bulk}} \ . \tag{2.100}$$

That means that the deformation of the field theory imposes a change in the bulk boundary term

$$I_b^W = I_b + N^2 W[\epsilon^{-\Delta}\phi(\epsilon,x)] \ , \tag{2.101}$$

which changes the on-shell constraint to

$$z^{-\Delta}(z\partial_z - \Delta)\phi\Big|_{z=\epsilon} = \phi_0 + \frac{\delta W}{\delta\hat{\beta}(x)}\left[\epsilon^{-\Delta}\phi(\epsilon,x)\right] \ . \tag{2.102}$$

Since the argument of the functional derivative of W is divergent for $\epsilon \to 0$

$$\epsilon^{-\Delta}\phi(\epsilon,x) \sim z^{d-2\Delta}\alpha(x) + \beta(x) \ . \tag{2.103}$$

That means, that the deformation of the field theory leads to ill-defined boundary conditions, which is generally the case, if W is itself independent of the cut-off.

In general, this means that we need to renormalize the deformed field theory. This is not a surprise, since with the deformation, we have broken conformal invariance of the theory, which now must be renormalized (see e.g. [85]). We assume that this can always be done consistently and that the finite part of (2.103) corresponds to the bare deformation of the dual field theory. Then, using the prescription of [82] imposing the boundary condition

$$\alpha(x) = -\frac{\delta W}{\delta \beta(x)} \tag{2.104}$$

corresponds to a deformation of the dual CFT's Euclidean action by

$$I_{\mathrm{CFT}} \to I_{\mathrm{CFT}} + N^2 W[\hat{\beta}] \,, \qquad \langle \hat{\beta} \rangle = \beta \,. \tag{2.105}$$

We will use this way of incorporating boundary conditions for a double trace deformation of the field theory in section 2.3.3. It will break both supersymmetry and conformal invariance. In the bulk, this means that the asymptotic AdS invariance is broken by back-reaction of the scalar field.

In the following, I am going to present a way to use such a deformation to construct an AdS space with a cosmological singularity and its dual field theory.

2.3 A singularity toy model

The AdS/CFT correspondence carries the exciting prospect of being able to address some long-standing issues in (Anti-de Sitter) quantum gravity in terms of well-defined and usually better understood quantum field theories. Of particular interest are "big bang" and "big crunch" singularities in the supergravity theory which in principle should have a holographically dual description in terms of a conformal field theory. In this section, I will describe a setup, in which an unstable bulk theory, which exhibits a singularity, is related to the deformation of a conformal field theory, using the prescription of generalized boundary conditions.

The idea to describe spacelike singularities in a dual theory reaches back to matrix models in two space-time dimensions [86, 87]. Light-like singularities have been investigated with matrix theory in higher space-time dimensions [88] and a non-commuting matrices have been suggested as a model of space-time near a singularity [89–92]. Insights can be gained from the study of singularities inside black holes [93–96] but the horizon concealing it protects the CFT from ever seeing the singularity. There are numerous other models of cosmological singularities in AdS/CFT [97–103].

The results of a first attempt to consider such an application of the AdS/CFT correspondence to four-dimensional space-times [104, 105] were somewhat inconclusive. In AdS_5/SYM$_4$ [106], the cosmological setup is more precise and better understood. The generalized scalar field boundary conditions correspond to a double trace deformation of the field theory [82] and to leading order in $1/N$ the singular nature of the evolution in the bulk is reflected by an unbounded double trace potential in the field theory. In a more detailed string theoretical construction of the holographic set-up, D3-branes induce an instability, which leads to the crunch [85].

A negative and unbounded potential for the scalar fields leaves the field theory without a proper vacuum and is in obvious conflict with unitarity. In the large N limit, loop corrections cannot improve on this situation since the double trace deformed theory is 1-loop exact and asymptotically free. One particular proposal to deal with this problem is to impose special self-adjoint boundary conditions for the field theory at infinity, which basically means to reduce the Hilbert space ad hoc such that it contains only symmetric states which have a wave packet coming in from infinity for every one vanishing there. When the field rolls up the potential, the singularity retracts from the boundary and this procedure results in a bouncing cosmology for the bulk [106, 107]. This is precisely whereon the aforementioned ekpyrotic model of cosmology is based, in which the initial conditions for the big bang are created by previous cosmological cycles. This idea has its problems, most notably having to do with particle creation and induced back-reaction, and relies on the assumption that the unboundedness of the potential remains an unavoidable consequence for all values of the parameters in the field theory. Strictly speaking, however, the unboundedness of the potential has only been checked in the limit $N \to \infty$. We have seen that this limit corresponds to effectively turning off quantum gravity in the bulk. A plausible alternative is that at finite N, a full (perturbative) analysis results in an effective potential with a new stable minimum in the far UV, which would drastically change the qualitative behavior of the field theory. In the bulk this picture would suggest a resolution of the big crunch singularity by higher order string corrections that become more and more important as the bulk scalar field flows down the unbounded supergravity potential.

Another important reason to suspect that the dynamics is more subtle than so far presented is that the double trace deformation breaks the conformal symmetry. I expect that this gives rise to the running of not just the double trace coupling, but also a scale-dependence of the gauge coupling at higher non-planar order. This implies a coupled set of (nonlinear) flow equations that should be studied carefully to determine the UV behavior. In particular this might result in the double trace deformation becoming marginally irrelevant, instead of asymptotically free, requiring the introduction of a UV cut-off in the theory. In the bulk gravitational description this should be related to the appearance of an additional dilatonic scalar degree of freedom, which might have important consequences on the solutions in the bulk and the validity of the supergravity limit near the crunch singularity. The appearance of a new stable minimum in the far UV stabilizing the dynamics was also discussed in recent work on AdS instanton solutions in the bulk and their dual interpretation in conformal field theory [108], although their proposal depends crucially on a positive conformal single trace contribution to the quartic potential that was added by hand. In our case we will only be interested in the effects of corrections suppressed by $\frac{1}{N}$ at higher loop order in the gauge theory deformed by a double trace interaction. If they are sufficient to stabilize the potential, regularization by hand as in e.g. [109] is not necessary.

In the remainder of this section I will review of the specific AdS bulk and then relate it to the deformed field theory. I will comment on the the 1-loop, large N, effective potential in the presence of the double trace deformation.

2.3.1 Cosmology in Anti-de Sitter space

To setup a cosmology which has a field theory dual I look at five-dimensional supergravity in AdS space. I use precisely the same model as [106, 107]. Type IIB string theory is defined in ten dimensions, five of which need to be compactified on a five-sphere. The low energy effective theory is $\mathcal{N}=8$ gauged supergravity [110–112], which is a consistent truncation of ten-dimensional type IIB supergravity dimensionally reduced on an S^5. In total, this compactification produces 42 scalars. For our purpose, we focus on the subset of the five scalars thereof, $\alpha_i, i = 1\dots5$, which describe the different quadrupole distortions of S^5.

Their action is [113]

$$\mathcal{S} = \int \sqrt{-g}\left[\frac{R}{2} - \sum_{i=1}^{5}\frac{1}{2}(\nabla\alpha_i)^2 - V(\alpha_i)\right] , \tag{2.106}$$

where units are such that the five dimensional Planck mass is unity. Supergravity is defined by its superpotential W, in terms of which the F-term potential for the scalar fields is given by

$$V(\alpha_i) = \frac{1}{R^2_{\rm AdS}}\sum_{i=1}^{5}\left(\frac{\partial W}{\partial\alpha_i}\right)^2 - \frac{4}{3R^2_{\rm AdS}}W^2 . \tag{2.107}$$

The easiest, most symmetric way to write the superpotential is in terms of new fields β_i, which are defined such that $\sum_i\beta_i = 0$. They are related to the five original scalars by

$$\begin{pmatrix}\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 0 & 1/2\sqrt{3}\\ 1/2 & -1/2 & -1/2 & 0 & 1/2\sqrt{3}\\ -1/2 & -1/2 & 1/2 & 0 & 1/2\sqrt{3}\\ -1/2 & 1/2 & -1/2 & 0 & 1/2\sqrt{3}\\ 0 & 0 & 0 & 1/\sqrt{2} & -1/\sqrt{3}\\ 0 & 0 & 0 & -1/\sqrt{2} & -1/\sqrt{3}\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\\\alpha_5\end{pmatrix} . \tag{2.108}$$

In terms of these new fields, the superpotential is given by

$$W = -\frac{1}{2\sqrt{2}}\sum_{i=1}^{6} e^{2\beta_i} . \tag{2.109}$$

If all the fields vanish, $\alpha_i = 0\ \forall i$, the compact S^5 is unperturbed. There, the scalar potential has a local maximum, which is the maximally supersymmetric AdS state. Around this state, each scalar obeys a free wave-equation with a mass that saturates the Breitenlohner-Freedman

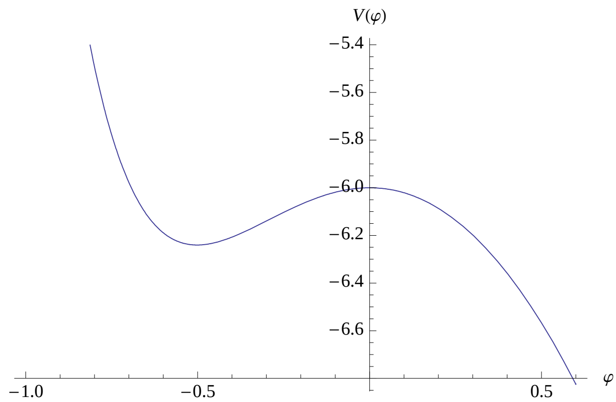


Figure 2.3: *The truncated* $5d$ *supergravity potential for an* $SO(5)$ *invariant scalar field, which is unbounded from below for large scalar field values.*

bound $m^2_{\rm BF} = -\frac{4}{R^2_{\rm AdS}}$. This means that the background considered here is perturbatively stable and it appears as if at least the gravity theory is not fundamentally flawed. Therefore, we expect that quantum effects should play an important rôle in resolving the cosmological singularity. One possibility to truncate the theory further to a single scalar is

$$\beta_i = \frac{\varphi}{\sqrt{30}}\ ,\ ,i = 1, \ldots 5, \qquad \beta_6 = -\frac{5\varphi}{\sqrt{30}}\ . \tag{2.110}$$

This theory is an SO(5) invariant scalar coupled to gravity [111] with a potential

$$V(\varphi) = -\frac{1}{4R^2_{\rm AdS}}\left(15e^{2\gamma\varphi} + 10e^{-4\gamma\varphi} - e^{-10\gamma\varphi}\right)\ , \tag{2.111}$$

with $\gamma = \sqrt{\frac{2}{15}}$, which is shown in figure (2.3).

Note that the supergravity potential is unbounded from below for positive values of the scalar field, a feature that is replicated in the dual gauge theory description. In global coordinates the AdS_5 metric reads

$$\mathrm{d}s^2 = R^2_{AdS}\left(-(1+r^2)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1+r^2} + r^2 d\Omega_3\right)\ . \tag{2.112}$$

As a reminder, scalar field perturbations behave as follows near the boundary as $r \to \infty$

$$\phi(r) = \frac{\alpha \ln r}{r^2} + \frac{\beta}{r^2}\ , \tag{2.113}$$

where the coefficients α and β depend on the other coordinates (t, x) and are related to each other in some specific way, in terms of a specified boundary condition, for the dynamics of the theory to be well-defined. For cosmological, time-dependent, behavior of the background to occur one adopts boundary conditions of the form

$$\alpha = -\frac{\partial W}{\partial \beta}\ , \tag{2.114}$$

where the function $W(\beta)$ is a priori an arbitrary function of the remaining coordinates, which will appear as an additional potential term in the dual field theory. The case of of interest here is when $\alpha = f\beta$, corresponding to a double trace deformation of the SYM theory. By allowing $\alpha \neq 0$ the scalar field falls off more slowly than in the standard (empty) AdS case, where $\alpha = 0$, and as a consequence the full AdS isometry group is partly broken.

Before moving on to briefly discuss the cosmological nature of the bulk solutions, let us remind the reader of the supergravity limit and the corresponding parameter map to the dual field theory. Taking supergravity as the low energy limit of string theory first of all requires that the string tension α' becomes large and strings shrink to point particles in comparison to the AdS radius R_{AdS}, i.e. $\frac{R_{\text{AdS}}}{\alpha'} \gg 1$. In the dual field theory this corresponds to the non-perturbative regime of large 't Hooft coupling, to be precise

$$\lambda \equiv g_{\text{YM}}^2 N = \left(\frac{R_{\text{AdS}}}{\sqrt{\alpha'}} \right)^4 \gg 1 \, . \tag{2.115}$$

The other requirement of a valid supergravity limit is that of small string coupling $g_s \ll 1$, such that string loop effects can be neglected. The AdS/CFT dictionary dictates that $2\pi g_s = g_{\text{YM}}^2$ and as a consequence a fixed 't Hooft coupling $\lambda = g_{\text{YM}}^2 N$ necessarily implies a large N planar limit. The effects of quantum string corrections are mapped to $g_s \propto \frac{1}{N}$ non-planar corrections in the dual field theory, whereas the free planar field theory limit ($\lambda \to 0$) should describe the (free) string theory to all orders in α', for which a bulk description in terms of supergravity breaks down completely. We would like to stress that the strict $N \to \infty$ limit corresponds to a free, classical, AdS string theory, in which the effects of quantum gravity are effectively turned off. As a consequence one should be careful to extend results derived in the (classical) $N \to \infty$ limit to the large, but finite, N case with gravity turned on. When confronted with singularities in the AdS bulk one would naively expect that an ever increasing strength of gravitational interactions should play an important, if not crucial, role in any mechanism to resolve the singularity and therefore a strict $N \to \infty$ limit could give rise to misleading results. As a corollary, non-planar contributions might result in drastically different conclusions regarding the effective potential and the corresponding behavior in the gravitational bulk. Looking at this from a pure bulk perspective this might be related to a non-perturbative inconsistency of the single $SO(5)$ invariant scalar field supergravity truncation. Since the double trace deformation is marginally relevant, breaking the super-conformal symmetries, higher order running of the gauge coupling should be expected and correspondingly the dilaton in the bulk should become dynamical, which is not described by the truncated supergravity Lagrangian and the corresponding instanton solutions.

2.3.2 Singular cosmology as instanton solution

I now explain in more detail, how the authors of [106] construct a solution to (2.111), which satisfies the generalized boundary conditions $\alpha = f\beta$ and develops a space-like singularity.

Those data are constructed as a slice of an O(5) invariant Euclidean instanton solution with metric

$$\mathrm{d}s^2 = R_{\mathrm{AdS}}^2 \left(\frac{\mathrm{d}\rho^2}{b^2(\rho)} + \rho^2 \mathrm{d}\Omega_4 \right) . \tag{2.116}$$

The function b can be determined in terms of ϕ from the field equations, which asymptotically read

$$b^2 = \rho^2 + 1 + \frac{1\alpha^2(\ln\rho)^2}{3\rho^2} + \frac{\alpha(4\beta-\alpha)\ln\rho}{3\rho^2} + \frac{8\beta^2 - 4\alpha\beta + \alpha^2}{12\rho^2} , \tag{2.117}$$

whereas the scalar field ϕ is subject to

$$b^2\phi'' + \left(\frac{4b^2}{\rho} + bb' \right) \phi' - R_{\mathrm{AdS}}^2 V_{,\phi} = 0 , \tag{2.118}$$

where $' = \partial_\rho$. Furthermore, it is required that the solution is regular at the origin, $\phi'(0) = 0$, such that the instanton solutions can be labeled by $\phi_0 = \phi(0)$. An instanton can be constructed by integrating (2.118) for a given boundary condition ϕ_0. One finds, indeed, that asymptotically

$$\phi(\rho) = \frac{\alpha \ln\rho}{\rho^2} + \frac{\beta}{\rho^2} , \tag{2.119}$$

where α, β are now constants.

This is used to construct time symmetric initial data for the Lorentzian solution by restricting to the equator of the S^4. The Euclidean radial coordinate ρ thus becomes the radial distance r on the initial data slice. For a given boundary condition $\alpha(\beta)$, one selects that point, which is also an instanton solution. For $f > 0$, there is precisely one such configuration. This solution is then analytically continued to a Lorentzian solution, which describes the evolution of such initial data under AdS-invariant boundary condition. Note that those slightly differ from $\alpha = f\beta$ and are expressed as

$$\alpha \left(1 - \frac{f}{2} \ln\alpha \right) = f\beta . \tag{2.120}$$

For small f, this difference is negligible. The analytic continuation transforms the origin of the Euclidean instanton to the lightcone of the Lorentzian solution and the $O(5)$ symmetry to $SO(4,1)$. This ensures, that inside the lightcone, the solution must behave like an open FRW universe

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\mathrm{d}H_4 , \tag{2.121}$$

where $\mathrm{d}H_4$ is the metric on the four-dimensional unit hyperboloid. As the field ϕ rolls down the negative, unbounded potential, the scale factor shrinks and vanishes in finite time. The associated degeneracy of the metric is the big crunch singularity. Outside the lightcone, the scalar field remains bounded and the solution is given by (2.116) with the sphere $\mathrm{d}\Omega_4$ replaced by four-dimensional de Sitter space.

For initial scalar field profiles satisfying the generalized $\alpha = f\beta$ boundary conditions one can argue on general grounds that a big crunch singularity will develop and spread to the boundary in finite global time [104, 105]. Approximate solutions can be found by analytically continuing

$SO(5)$ invariant Euclidean instanton solutions, describing the decay of the maximally supersymmetric AdS vacuum along the direction of the potential that is unbounded from below. Inside the light-cone that spreads from the origin the solution is described by a crunching FRLW cosmology, which hits the boundary in finite global time. Because the scalar field ends up rolling down an unbounded exponential potential the appearance of a big crunch singularity should not come as a surprise. This case should probably be considered much more severe than the relatively mild singularities appearing in Coleman-De Luccia instantons describing the decay of a false AdS vacuum into another stable AdS minimum, which have recently received renewed attention because of their potential holographic description in terms of a cut-off field theory at the spherical domain wall separating the two AdS vacua [114, 115]. It would certainly be of interest to see how any of these ideas apply in this, more extreme, case. Having briefly summarized the bulk story, we would now like to move on to the holographically dual gauge theory description.

2.3.3 The effective potential in $\mathcal{N} = 4$ SYM with a double trace deformation

In the maximally supersymmetric AdS vacuum the holographic dual is of course the $\mathcal{N} = 4$ Super-Yang-Mills theory with $SU(N)$ gauge group, whose action is [106]

$$S_0 = \int \mathrm{d}^4x \,\mathrm{tr}\, \left\{ -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}D_\mu\Phi^i D^\mu\Phi^i + \frac{1}{4}g^2[\Phi^i,\Phi^j]\,[\Phi_i,\Phi_j] + \text{ fermions} \right\} , \quad (2.122)$$

with field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu]$ and covariant derivative $D_\mu\Phi^i = \partial_\mu\Phi^i + ig[A_\mu,\Phi^i]$.[1] We deform this action by adding a double trace potential [82, 106]

$$V_{\mathrm{Tr}^2} = -\frac{f}{2}\int d^4x\, \mathcal{O}^2 ,\ f > 0 \,, \quad (2.123)$$

where the operator $\mathcal{O}$ is chosen to be the half-BPS operator of dimension two, holographically dual to the $SO(5)$ invariant bulk scalar field φ,

$$\mathcal{O} = \frac{1}{N}\,\mathrm{Tr}\,\left[\Phi_1^2 - \frac{1}{5}\sum_{i=2}^{6}\Phi_i^2\right] \,, \quad (2.124)$$

where $\Phi_1 \ldots \Phi_6$ are the six scalars of the theory.

Fixing the $\frac{1}{N}$ counting

Since we want to calculate $\frac{1}{N}$ corrections to the β function, we need to fix the counting of N consistently. Observe that in our theory (2.122) a single trace 4-vertex comes with g^2 and a

[1] Actually, the coupling $g = g_{\mathrm{YM}}$. Since we are only interested in the scalar sector of the theory for the moment, we keep it as g, though.

double trace 4-vertex with $\frac{f}{N^2}$. In the following, n denotes the number of loops of a given diagram.

Let us first fix the overall scaling by looking at diagrams with only single trace vertices. A general single vertex diagram scales as

$$\sim g^{2(n+1)}N^n = (g^2N)^{n+1}\frac{1}{N} \; . \tag{2.125}$$

For a 't Hooft coupling for the single trace vertex

$$\lambda = g^2N \overset{N\to\infty}{=} \text{const.} \tag{2.126}$$

we find that diagrams scale at the same order in N for all loops as the tree-level diagram. Hence all diagrams scale at $\frac{1}{N}$ in the large N limit and no diagram is allowed to scale at a power higher in N.

We now fix the double trace 't Hooft coupling. Because there are different ways to impose this group structure, the number of closed index loops can vary according to the specific diagram and hence the order in N. We focus on the index structure which produces the highest order in N. Such n-looop diagrams scale at

$$\sim \frac{f^{n+1}}{N^2} = (fN^i)^{n+1}N^{-2-(n+1)i} \; . \tag{2.127}$$

Requiring that each diagram scales at the same order in N as the tree level diagram

$$\sim \frac{f}{N^2} = (fN^i)N^{-(2+i)} \tag{2.128}$$

yields the condition

$$-(2+i) = -2-(n+1)i \Leftrightarrow i = 0 \; . \tag{2.129}$$

Therefore the consistent 't Hooft coupling for the double trace interaction is

$$f \overset{N\to\infty}{=} \text{const.} \tag{2.130}$$

and every diagram with only double trace vertices scales at order $\frac{1}{N^2}$, which is one order lower than the single trace diagrams above and hence consistent with the maximal scaling requirement.

We now check that this definition of 't Hooft couplings is still consistent for diagrams with different kind of vertices. For such diagrams, we also assume the group structure that produces the highest possible order in N. Such n-loop diagrams with j double trace and $n+1-j$ single trace vertices have one index loop per boson loop and one additional index loop for a boson loop between two double trace couplings, hence for all up to one. The diagram scales as

$$\sim \left(\frac{f}{N^2}\right)^j \left(g^2\right)^{n+j-1} = f^j \left(gN^2\right)^{n+1-j} N^{-2} \; , \tag{2.131}$$

which is also one order less than single trace diagrams and consistent with the scaling requirement.

Renormalization in the large-N limit

The squaring of the trace results in a contraction of the gauge degrees of freedom differently than in the single trace case. The scalar field Φ_1 is identified as the steepest negative direction in the effective potential (2.124) and we focus on the dynamics of Φ_1 rolling down a fixed direction $\Phi_1(x) = \phi(x)U$ [2]. We should add that strictly speaking the supersymmetric gauge theory is defined on a 3-sphere, in which case a mass term appears for the Φ_1 scalar due to the conformal coupling to the curvature of the S^3. This quadratic mass term can be neglected in the UV, or equivalently for large enough field values.

The bare potential of this theory is negative and unbounded from below. The double trace coupling gets renormalized at one-loop level. The one-loop effective potential reads

$$V(\mathcal{O}) = -\frac{f}{2}\mathcal{O}^2\left[1 - \frac{f}{2}\ln\left(\frac{\mathcal{O}}{\mu^2}\right)\right] , \tag{2.132}$$

where μ is a UV cut-off scale. Following the Coleman-Weinberg prescription [116] we define the renormalized coupling $f_{\text{ren.}}(\mu)$ by the renormalization condition

$$V(\mu) = V(\mathcal{O})|_{\mathcal{O}=\mu^2} = -\frac{f_{\text{ren.}}(\mu)}{4}\mu^4 . \tag{2.133}$$

This corresponds to having the sliding scale μ to be set by the (homogeneous) expectation value of the operator $\mathcal{O}$, rather than by an external momentum[3], implementing dimensional transmutation. The beta function is most readily obtained by demanding that the effective potential (2.132) is independent of the scale μ, leading to

$$\mu\frac{\partial f}{\partial \mu} = -f^2 . \tag{2.136}$$

After identifying $\phi = \sqrt{\mathcal{O}}$, this gives the following result for the renormalized coupling

$$f_{\text{ren.}}(\phi) = \frac{f(M)}{1 + f(M)\,\ln(\phi^2/M^2)} , \tag{2.137}$$

with an arbitrary scale M acting as the scale at which the perturbative theory is defined. On physical (continuity) grounds it is natural to suppose that this infrared scale is close to the

[2] U is a constant Hermitian matrix satisfying $\text{Tr}U^2 = 1$, so that ϕ is a canonically normalized scalar field. We focus on the dynamics of Φ_1, only.

[3] Note that this means that the renormalization scale of the theory changes as $\phi \sim \sqrt{\mathcal{O}}$ rolls down the potential. This can only be done adiabatically and therefore we have to obey the "slow-roll condition" that the time-scale at which the system is probed is small compared to the time-scale on which the system changes

$$\frac{1}{|\mu|} \ll \frac{1}{|\dot{\phi}|} , \tag{2.134}$$

which yields with $\mu = \phi$

$$\frac{|\dot{\phi}|}{|\phi|} \gg 1 . \tag{2.135}$$

However, the larger ϕ becomes, the larger becomes its slope and perturbation theory might break down.

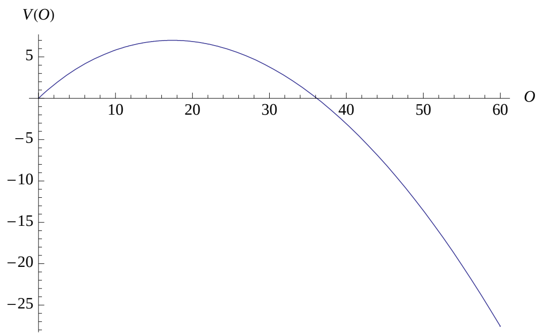


Figure 2.4: *The one-loop effective potential $V(O)$ with $\epsilon = 0.1$ and $M = 1$ (for large values of O), regularized for small values of O by a conformal mass term.*

scale where the conformal masses will start dominating the effective potential. The Coleman-Weinberg potential at one-loop level now reads, for $\mu = \phi \gg M$,

$$V = -\frac{1}{4}\frac{\phi^4}{\ln(\phi^2/M^2)} \,. \tag{2.138}$$

This 1-loop result given in [106] is exact in the large-N limit for $N_c = 4$ SYM theory. This means, that there are no higher order corrections to the coupling (in the large-N limit) and as a consequence the theory is asymptotically free. Therefore, the larger the field value the smaller the coupling and one concludes that perturbation theory should become an ever more accurate description for larger field values, i.e. at larger energies. On the other hand, the absence of an (approximately) defined vacuum state for large field values is obviously a source of concern. The (renormalized) field theory potential, regularized in the IR, is depicted in fig. 2.4.

In the absence of a ground state, the wave-function of the scalar field will spread to infinity in finite time. This means that unitarity will be lost. As a solution to this dilemma, it has been proposed [106] to employ specific boundary conditions at infinity, which reflect each mode back as it rolls down the potential. Such boundary conditions are known as a *self-adjoint extension* [117,118]. Rather than an extension, it is a restriction of the Hilbert space to contain only such modes for which the Hamiltonian

$$\widehat{H} = -\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{1}{4}\lambda x^p \tag{2.139}$$

is self-adjoint. For large x, the WKB approximation is increasingly accurate and the WKB wavefunctions

$$\chi_E^{\pm}(x) = \left[2\left(E + \frac{\lambda x^p}{4}\right)\right]^{-\frac{1}{4}} \exp\left(\pm\mathrm{i}\int^x \sqrt{2\left(E + \frac{\lambda y^p}{4}\right)}\mathrm{d}y\right) \tag{2.140}$$

for a given energy E can be used as an ansatz to study the generic behavior of energy eigenfunctions. It can be seen that for a linear combination of these eigenfunctions

$$\psi_E^{\alpha}(x) \sim x^{-p/4} \cos\left(\frac{\sqrt{2\lambda}x^{p/2+1}}{p+2} + \alpha\right) \,, \tag{2.141}$$

the Hamiltonian is, indeed, self-adjoint. Such a procedure works in quantum mechanics, but near the singularity, the evolution of the field is ultra-local, which means that spatial gradients become unimportant for the field evolution. That implies that the quantum field theory can be seen as a collection of independent quantum mechanical oscillators at each point in space and it can be attempted to impose these conditions at every spatial point.

There are, however, serious doubts about the validity and motivation of such an approach. First of all, such a selection of a sub-Hilbert space is ad hoc and not justified by any symmetries or other properties of the Hamiltonian. They are rather imposed ex post to justify the theory. Secondly, although this procedure removes the unitarity violation, it still does not give the theory a ground state. Hence, it remains unclear on how to build a Fock space from the vacuum. In essence, the prescription applies to quantum mechanics only and is, here, extended to a limit of quantum field theory. In fact, it has been observed numerically [109], that very quickly, the energy of the initial configuration gets converted into gradient energy, which eventually diverges. The initial wave package evolves non-adiabatically and particles are produced. This energy is not converted back into a homogeneous mode. In particular, there is no transition from a Big Crunch to a Big Bang. On these grounds, I am not satisfied with the self-adjoint extension as a solution to the problem of having a dual quantum field theory without a ground state. Rather than imposing a solution, I want to examine if the theory itself regularizes its potential by taking into account all the quantum effects, in particular those suppressed by $1/N$.

We expect that this potential gets turned around by $\frac{1}{N}$ corrections, i.e. that including finite-N diagrams in the calculation of the effective potential will render it finite and create a true vacuum. The minimal change of the coupling, which would achieve such a behavior, is

$$f(\phi) = \frac{\epsilon}{\ln\phi^2 + \alpha\phi^A} \ , \tag{2.142}$$

where A is a number of function determined by renormalization below, that needs to reach a value $A.4$ for some large value of ϕ in order to cancel the ϕ^4 in the numerator. Note that the scalar field occurs here, because of the renormalization procedure, in which $\mu = \phi$ and thus the exponent A doesn't need to respect the invariance of the theory under $\phi \to -\phi$. In fact, since A is determined by renormalization theory, it doesn't even need to be integer and will also change it's value, as the renormalization scale ϕ increases. The behavior of the potential indeed changes as desired as seen in fig. 2.5. Note that a singularity at $\phi_0 < 1$ remains and that the potential is unbounded there. This is a region in which perturbation theory is not valid. We remark that the sign of the added term should be $\alpha > 0$, because otherwise we do not create a minimum but a new maximum.

A few comments are probably in order. The deformed theory has a UV conformal fixed corresponding to the standard super Yang-Mills theory which is suggestive of a consistent and complete holographically dual description, i.e. no new degrees of freedom have to be introduced at or above some UV cut-off scale, which is clearly important if our ambition is to understand (or resolve) the appearance of space-like crunch singularities in the bulk. On the other hand, the bad news is that the theory does not have a well-defined vacuum state. One

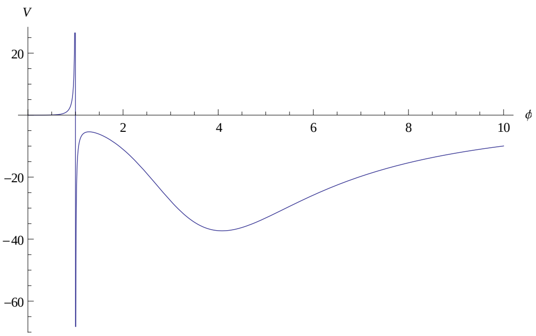


Figure 2.5: *The effective potential* (2.142) *which now features a vacuum.*

could imagine "fixing" this by regulating the potential such that instead of being unbounded from below it features a new globally stable minimum at some large field value. One obvious way to do this would be to add higher-dimensional irrelevant corrections to the potential. This has one obvious drawback, namely that it requires the introduction of irrelevant operators that turn the theory incomplete beyond some UV cut-off. Instead of adding such operators by hand, one could even imagine that higher-order, non-planar, quantum corrections could change the anomalous dimension of the coupling f to become irrelevant at some high energy scale, disturbing the asymptotically free nature of the double trace deformation. Even though this might produce a stable vacuum, it would be disastrous from the point of view of having to rely on a theory with a (perturbative) UV fixed point. To avoid this one could consider stabilizing the potential with a positive and exactly marginal contribution, like the single trace quartic operator [108]. This is not what we are after in this work. Instead, the modest but difficult goal we have set is to investigate the behavior of the beta-functions for the double trace deformation at the non-planar two-loop order. The aim is to explicitly check whether the double trace deformed theory remains asymptotically free and if the effective potential remains unbounded from below. Because two-loop non-planar corrections will involve mixing between the gauge and the double trace coupling we will be forced to also study the running of the gauge coupling at higher order. Analysis of the coupled system of RG-flow equations can then reveal the behavior of the effective double trace coupling and potential. The necessary inclusion of a second, coupled, degree of freedom in the form of the gauge coupling will turn out to have important consequences for the UV behavior of the deformed theory.

After this introduction, let us now move on to an analysis of two-loop nonplanar corrections.

2.4 Limitations of the large-N limit

In the previous subsection, I have mentioned that the effective potential was one-loop exact and that the Yang-Mills coupling does not contribute to the RG-flow of the double trace coupling. In this subsection I want to explain these statements and justify why I expect them to change

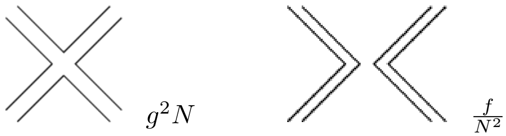


Figure 2.6: *The singe (left) and double (right) trace vertices depicted in double line notation. Note that in the large-N limit, the single trace vertex scales as g^2N, whereas the double trace vertex scales with $\frac{f}{N^2}$.*

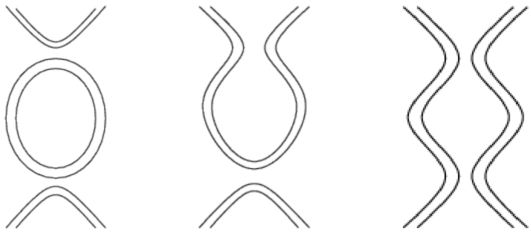


Figure 2.7: *Different diagrams of order f^2 at the one-loop level. Different contractions of the double-trace vertex yield different orders in $1/N$. Hence, only the very left diagram survives the large N limit, which dominates the other two at order N^2.*

qualitatively, when we take into account finite-N effects.

2.4.1 Effects of the large-N limit

When dealing with a double-trace vertex, it is important to realize, that as opposed to the single trace vertex, it can be contracted in two different ways. Depending on the contraction, the number of closed index loops changes. This means, that with a double trace vertex in a diagram, not only the number of vertices but also the specific contraction at each vertex determines the order in $1/N$ of a given diagram. Looking at the one-loop order first, depicted in figure 2.7, we see that only one of the possible three diagrams will contribute in the large N limit. Each vertex contributes a factor f/N^2, whereas each closed index loop contributes a factor of N. Combined, we see that only the contraction which yields two closed index loops has the same order in $1/N$ as the vertex (cf. figure 2.6). We see that the large N limit significantly reduces the number of diagrams, we have to take into account. I treat the contributions of the diagrams sub-leading in $1/N$ to the one-loop renormalization in detail in appendix B.2.

Let us now turn our attention to the two-loop level. There are two prototypical shapes of two-loop diagrams, which are represented in figure 2.8. To see the group structure, we have to thicken all the propagators to be double lines and replace the vertices by single trace and double trace contractions. Since there are quite a few of these diagrams, this is reserved for appendix B.3. It is, however, clear by inspection, that the diagrams with the two loops intertwined can at most have two closed index loops as opposed to the left diagram, for which the maximal number of index loops is four as depicted in figure 2.10. We see that for this contraction,

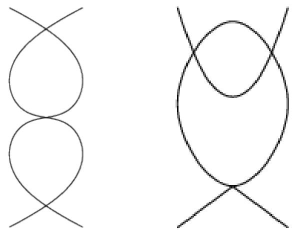


Figure 2.8: *The momentum structure of two loop diagrams correcting the four-vertex. For our purposes, the propagators have to be doubled and the vertices have to be replaced by single and double trace contractions. If we draw all the diagrams and do the counting, we observe that only diagrams of the left shape survive in the large N limit.*

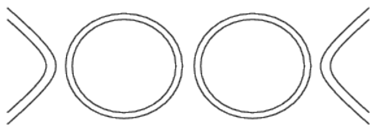


Figure 2.9: *The diagram with the most closed index loops at the two-loop level, which is leading in the large N limit. Note, that it has the same order in $1/N$ as the vertex. It factorizes into a square of the leading one-loop diagram in figure 2.7.*

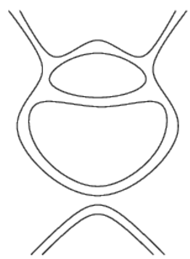


Figure 2.10: *The non-factorizable two-loop diagram in double-line notation. The maximum number of index loops is 2 such that the diagrams of this type are always subleasing in $\frac{1}{N}$ as compared to the diagram in figure 2.9.*

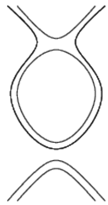


Figure 2.11: *The one-loop diagram containing a single and a double trace vertex. The contractions are such that this diagram renormalizes the double trace coupling, which can be seen when following the flow of the gauge degrees of freedom. This diagram scales as $\frac{f}{N^2}(g^2N)$, which is, keeping the 't Hooft coupling fixed, again of the same order as the double trace vertex.*

a new vertex, which counts f/N^2 comes with two new index loops contributing N^2 so that these diagrams are the leading ones in the large N limit. Besides, for any loop, the incoming momentum is the same, which means that the diagram can be factorized into a square of the leading one-loop diagram in figure 2.7.

The same argument can be repeated for any order of loops: The chain-type diagrams are leading in the large N limit and are factorizable. This implies that, in the large N limit, higher loops do not contribute any new divergences to the Green's function and hence, there is no need to introduce new counterterms, which means that the β-function does not change. In other words, it is *one-loop exact* (cf. [82, 106, 107]). We see that the large N limit leads to an amazing simplification of this theory.

The one-loop exactness of the theory has another astonishing consequence. So far, I have only talked about diagrams, that contain only one type of vertex, namely the double trace one. It is, however, conceivable, that a diagram contains single and double trace vertices at the same time. In figure 2.11, I depict such a one-loop diagram.

It can be shown that, actually, such diagrams do not contribute to the effective potential [106]. To see this, we expand the (real) scalar fields around a constant background $\Phi^1(x) = \phi(x)U$ for the steep direction, such that $\operatorname{tr} U^2 = 1$.[4] Observe that U can be chosen to be diagonal to simplify the calculations. We compute the masses of the various modes in this background

$$\Phi^1 = \phi U + R^1 \qquad (2.143)$$

$$\Phi^i = 0 + R^i\,, \quad i = 2\dots 6\,. \qquad (2.144)$$

Taking the terms quadratic in R, we see that the masses of the scalars are

$$(M^1_{ab})^2 = g^2\phi^2(U_{aa} - U_{bb})^2 + \frac{2fa^2\phi^2}{N^2}(1 + \delta_{ab}U^2_{aa}) \qquad (2.145)$$

$$(M^i_{ab})^2 = g^2\phi^2(U_{aa} - U_{bb})^2 - \frac{2fa^2\phi^2}{5N^2}\,. \qquad (2.146)$$

[4]See section 2.5 for details on the background field method and its extension to two loops.

Observe that those are already diagonal in field space. Each bosonic mode contributes

$$\frac{1}{32\pi^2}\frac{1}{2}M^4\ln\left(\frac{M^2}{\Lambda^2}\right) \tag{2.147}$$

to the effective potential. Since we are looking for a term proportional to f, we focus on the cross-terms of the two couplings. Those need to come from off-diagonal entries, since otherwise the single trace term vanishes. We find for Φ^1

$$2g^2\phi^2(U_{aa}-U_{bb})^2\frac{2fa^2\phi^2}{N^2}\ln\left(g^2\phi^2(U_{aa}-U_{bb})^2+\frac{2fa^2\phi^2}{N^2}\right) \tag{2.148}$$

and for $\Phi^2\ldots\Phi^6$

$$5\cdot 2g^2\phi^2(U_{aa}-U_{bb})^2\frac{2fa^2\phi^2}{5N^2}\ln\left(g^2\phi^2(U_{aa}-U_{bb})^2+\frac{2fa^2\phi^2}{5N^2}\right)\,. \tag{2.149}$$

To simplify this, we need to expand the logarithm around small f, which is possible, because we perceive the double trace interaction as a perturbation of SYM. Then the contributions of the 6 scalar fields to the mixed term precisely cancel. Thus, this is a feature of the scalar sector and of the exact form of the scalar potential rather than a property due to supersymmetry. The commutator squared part for the single trace interaction is typical for SYM theory, whereas the form of the double trace operator, which preserves an $SO(5)$ sub-symmetry is due to the specific choice of truncation in the dual supergravity theory.

2.4.2 Qualitative changes beyond large-N

As I have been mentioning before, the one-loop renormalized theory, which is exact in the large-N limit, does not have a ground state. If a theory turns pathological in a certain approximation, this can be taken as a hint that the approximation was not a valid one to do. In this case, there is another indication that $1/N$ effects might be important. They correspond to quantum gravity or g_s-effects in the bulk, which certainly play a rôle around around singularities, where gravity is strong. For a field theory that is dual to a bulk with a cosmic singularity, it is therefore expected that finite-N effects will qualitatively change the behavior of the theory.

The reason why new features are conceivable, is because at sub-leading order in $1/N$, non-factorizable diagrams such as the one depicted on the right of figure 2.8 come into play at the two-loop level. With them, new divergences arise, which contribute new terms to the β-function. Details on the full two-loop renormalization of ϕ^4 theory can be found in appendix B.3. In particular, it is possible that the RG-flows of the Yang-Mills coupling g and the double-trace coupling f mix. Recall that the cancellation of such diagrams was due to a conspiracy of supersymmetry and the resident SO(5) R-symmetry. Whereas the one-loop effective potential only depends on the masses of the scalars, the two-loop contributions also depend on the couplings directly in a non-trivial way.

The most drastic qualitative change of the theory would be that quantum corrections regularize the effective potential. Such an effect can happen, if there is a contribution to the β-function, which is linear in the double trace-coupling f, as I will show below. Recall that the argument, why such corrections where absent at the one-loop order is independent of the order in $1/N$. Therefore, we expect such contributions to arise at the two loop level, where they would be encoded in diagrams with two sinlge and one double trace vertices.

2.4.3 POTENTIAL SELF-CORRECTION FROM AN RG POINT OF VIEW

As mentioned above, the theory is one-loop exact in the large-N limit, such that higher-loop effects alone wouldn't change the behavior of the potential. Hence, we expect that a "turnaround" as described could only be created by including finite-N effects in perturbation theory at 2-loop order. If quantum corrections were to take care of the unboundedness of the potential and would regularize it, this would show up as a higher order correction to the coupling constant $f(\phi)$. In the following, I determine a necessary condition for a turnaround by presuming a favorable form of the coupling. I integrate its β-function to determine its dependence on the cutoff.

The β-function is the derivative of the coupling with respect to the field (multiplied by the field in order to render it dimensionless). Thus for (2.142)

$$\beta(f) = \phi\frac{\partial f}{\partial\phi} = -\frac{\epsilon}{(\ln\phi^2+\alpha\phi^A)^2}\left(\frac{1}{\phi^2}2\phi+\alpha A\phi^{A-1}\right)\phi \tag{2.150}$$

$$= -\frac{f^2}{\epsilon}\left(2+\alpha A\phi^A\right) \tag{2.151}$$

$$= -\frac{f^2}{\epsilon}\left(2+\frac{\epsilon A}{f}+\dots\right) \tag{2.152}$$

$$= -\frac{2}{\epsilon}f^2 - Af\ , \tag{2.153}$$

where in the second last line we used that $f\to\frac{\epsilon}{\alpha\phi^A}$ for large values of ϕ. Note that the prefactor of the linear term determines the scaling and needs to be >4 in order to produce a turnaround. From the form of the β-function in (2.153) we see that the desired correction needs to be caused by diagrams

- involving only one double trace vertex (and thus two single trace vertices)
- contributing with the same sign as the 1-loop correction

We now integrate $\beta(f)$ to get the RG-flow of f with respect to the cutoff Λ, which is set up to equal the field ϕ in [106].

$$\frac{\partial f}{\partial\ln\Lambda} = -\frac{2}{\epsilon}f^2-6f \Longleftrightarrow -\frac{\partial f}{\frac{2}{\epsilon}f^2+6f} = \partial\ln\Lambda\ , \tag{2.154}$$

integrating which on both sides yields

$$-\frac{1}{6}\left(\ln f+\ln(3+\epsilon f)\right) = \ln\Lambda\ , \tag{2.155}$$

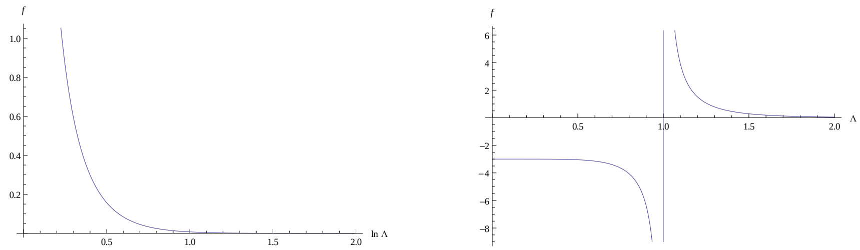


Figure 2.12: *The RG-flow of f with the expected two-loop correction with $\epsilon = 1$. The theory remains asymptotically free.*

which we can solve for f to get

$$f = \frac{3}{-\epsilon + e^{6 \ln \Lambda}} = \frac{3}{-\epsilon + \Lambda^6} \; . \tag{2.156}$$

This is still an asymptotically free coupling as can be seen in fig. 2.12, where we have plotted f with respect to $\ln \Lambda$ and Λ, respectively.

2.4.4 Coupled RG-flow of two scalar couplings

As explained above we are looking for corrections to the β-function which are sub-leading in $1/N$. Such corrections will show up at two-loop order. In the following we are laying out, how to extract the β-function from the two-loop counterterms.

It is important to notice that at the two-loop level, mass and field renormalizations kick in in all the theories under consideration. Thus the full renormalization of all parameters needs to be done and subsequently the Callan-Symanzik-equation needs to be solved. Note also, that the β-function for g and f are coupled. Furthermore, in a massive theory, the mass term is reparametrised to $m^2\phi^2 = a\mu^2\phi^2$ for the sake of dimensional regularization. The Callan-Symanzik-equation for an n-point Green's function $G^{(n)}$ in a theory with only one field then reads

$$\left(\mu \frac{\partial}{\partial \mu} + \beta_g \frac{\partial}{\partial g} + \beta_f \frac{\partial}{\partial f} + \beta_m \frac{\partial}{\partial a} + \sum_{k \in \text{fields}} n_k \gamma_k \right) G^{(n)}(\{x_i\}, \mu, g, f, a) = 0 \; . \tag{2.157}$$

where $\beta_m = (d - 6 + \gamma_m)a$. Here, d is the (regularized) dimension of the field theory and γ_m takes the scaling of the mass operator in the Green's function into account. Later on we are working in a massless theory and hence, this part of the Callan-symanzik-equation drops out.

The correction of the β-function which regularizes the effective potential finite after including finite-N corrections has been expected in (2.153) and the comments thereunder to come from a diagram involving one double trace and two single trace interactions as depicted in fig. 2.13.

However, there are more diagrams that could renormalize the double trace interaction coming from couplings to fermions and vectors as well as from mass renormalizations of bosons and

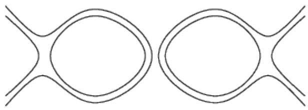


Figure 2.13: *The diagram we expect to give us the desired correction to the β-function at the two-loop level involves two single trace and one double trace vertex.*

fermions. We neglect those effects, because they would not be different from standard, undeformed $\mathcal{N}=4$ super Yang Mills theory and cancel due to supersymmetry to yield a conformal theory. The deformation, which breaks supersymmetry, and hence, on a quantum level, also conformal invariance, applies to the scalar sector, only.

At the two-loop level, the β-functions of the two remaining couplings have the following terms

$$\frac{\partial g}{\partial \ln \mu} = \eta^{(g)}_{4,0} g^4 + \eta^{(g)}_{2,1} g^2 f + \eta^{(g)}_{0,2} f^2 + \eta^{(g)}_{6,0} g^6 + \eta^{(g)}_{4,1} g^4 f + \eta^{(g)}_{2,2} g^2 f^2 + \eta^{(g)}_{0,3} f^3 \quad (2.158)$$

$$\frac{\partial f}{\partial \ln \mu} = \underbrace{\eta^{(f)}_{4,0} g^4 + \eta^{(f)}_{2,1} g^2 f - \eta^{(f)}_{0,2} f^2}_{1-\text{loop}} + \underbrace{\eta^{(f)}_{6,0} g^6 + \eta^{(f)}_{4,1} g^4 f + \eta^{(f)}_{2,2} g^2 f^2 + \eta^{(f)}_{0,3} f^3}_{2-\text{loop}} . \quad (2.159)$$

In principle, also terms $\sim \mathrm{e}^{-f/g}$ and $\sim \ln f/g$ could occur in the β-function but because they would correspond to non-perturbative contributions we will neglect them here. These two β-functions form a system of coupled differential equations we would like to study more carefully. In particular, the actual presence and signs of the different terms will of course play a crucial role in determining the UV behavior of the couplings.

Obviously some of the terms appearing in (2.158) will have to vanish or are simply of no interest to us. First of all, in the absence of the double trace deformation the gauge theory is super-conformal and therefore all terms independent of f should vanish in the beta-function for g^2. So we immediately conclude that

$$\eta^{(g)}_{4,0} = \eta^{(g)}_{6,0} = 0 . \quad (2.160)$$

Furthermore, for $N=4$ SYM theory, a non-renormalization argument excludes the diagram with one single- and one double-trace vertex at the one-loop level

$$\eta^{(f)}_{2,1} = 0 . \quad (2.161)$$

Under the assumption that $f \ll g \ll 1$ we can also neglect terms of higher order in f, implying that $\eta^{(g)}_{2,2} = \eta^{(f)}_{2,2} = 0$. Finally, and importantly, we are not interested in the term proportional to f^3 because it cannot affect the UV behavior of the double trace coupling that we are interested in. Basically, the only terms that can have interesting effects on the couplings in the far UV are the leading contribution to the scale dependence of the gauge coupling and the next to leading, gauge coupling dependent, contribution to the running of the double trace coupling. Hence, under this assumption and using that we have already shown that $\eta^{(f)}_{0,2} = 1$, we are left with

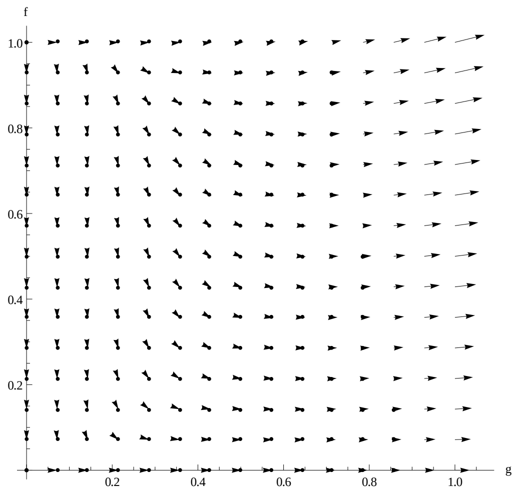


Figure 2.14: *Vector plot of the coupled system of β-functions for g and f assuming all coefficients to be 1. the point $(g = 0, f = 0)$ is an unstable fixed point. Including corrections to the 't Hooft limit drive the theory away from this point.*

the following set of coupled differential equations,

$$\begin{aligned} \frac{\partial g^2}{\partial \ln \mu} &= \eta^{(g)}_{2,1} g^2 f + \eta^{(g)}_{4,1} g^4 f \\ \frac{\partial f}{\partial \ln \mu} &= -f^2 + \eta^{(f)}_{4,1} g^4 f \ . \end{aligned} \tag{2.162}$$

Note that for now we kept two terms in the beta function for g^2. In the final analysis we will only be interested in the leading contribution (which is the $g^2 f$ term, unless the coefficient vanishes). Determining the coefficients of these terms requires either a standard perturbative Feynman diagram analysis. The results for all the coefficients are given in appendix B.4.

The structure of these flow equations is such that a number of different things could happen, depending on the signs of the coefficients. Due to the mixing with the running gauge coupling one possibility is that both the double trace and gauge coupling increase towards the UV, implying the double trace deformed theory is actually ill-defined, contrary to the result at one-loop. Another option would keep the double trace coupling behavior relevant, but depending on the behavior of the gauge coupling in the UV limit, the effective potential could be turned around featuring a stable vacuum state.

For this reduced coupled system of β-functions we give the RG-flow in a vector plot in fig. 2.14. Here, we see that the point of the free theory for $f = g = 0$ is an unstable fixed point, from which the theory is driven away, if $g \neq 0$, hence, if we include corrections to the large N limit in which effectively $g = 0$. As commented earlier, the term which has a chance to cause the turnaround is $\eta^{(f)}_{4,1} g^4 f$. If its sign changes, the result might be different. However, the picture doesn't change qualitatively as we see in fig. 2.15. It is worth to contemplate about the bulk

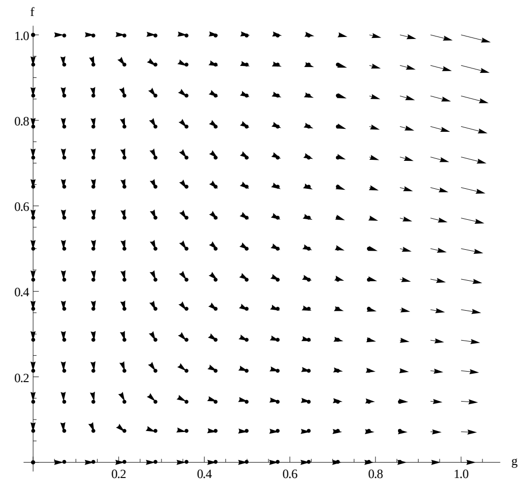


Figure 2.15: *As in fig. 2.14 but with $\eta_{4,1}^{(f)} < 0$. The point $(g = 0, f = 0)$ remains unstable.*

dual of this term. Since it is suppressed by $1/N$ and contains two 't Hooft couplings,

$$\eta_{4,1}^{(f)} g^4 f \leftrightarrow \eta_{4,1}^{(f)} g_s \left(\frac{R}{\alpha'}\right)^8 . \tag{2.163}$$

This should be matched with a corresponding term in the supergravity action. Note also that the term $\eta_{4,0}^{(f)} g^4$ will lead to a logarithmic scaling of f.

2.5 Results

A convenient way to calculate the effective potential including all kinds of fields, such as scalars, fermions, vectors and ghosts is the background field method [119–122]. Here, one can fix a gauge and compute quantum corrections without loosing explicit gauge invariance. Although the effective potential is gauge dependent, its physical properties are not [123, 124]. In particular, we employ the setup proposed in [125], which performs the calculation in Landau gauge. As opposed to calculating the 4-vertex Green's function, this has the advantage, that the effective potential is calculated by summing only vacuum graphs without external momenta, which simplifies the calculations enormously.

In this formalism, all the fields are separated into a classical background and its quantum fluctuations about it. For instance, a (real) scalar field is represented as $\phi + R$ with its background ϕ and the perturbations R around it.[5] The effective potential is the tree-level potential in the classical background plus the sum of all connected one-particle-irreducible vacuum graphs

$$V_{\text{eff.}} = V^{(0)} + \frac{1}{16\pi^2} V^{(1)} + \frac{1}{(16\pi^2)^2} V^{(2)} + \ldots , \tag{2.164}$$

[5] A complex scalar field would be represented as a background with two real fluctuations.

where $V^{(n)}$ denotes the n-loop correction. When calculating those with Feynman rules, the couplings and masses acquire a dependence on the background.

Since the one-loop diagram does not contain a vertex, the one-loop correction only depends on the masses, induced by the background, and reads

$$V^{(1)} = \frac{1}{4} \sum_{\text{fields } i} (-1)^{2s_i}(2s_i+1)(m_i^2)^2 \left(\ln \frac{m_i^2}{Q^2} - k_i \right) , \tag{2.165}$$

where the index i runs over the real scalars, two-component fermions and vector degrees of freedom in the theory. The renormalization scale is denoted by Q, $s_i = 0, 1/2, 1$ for scalars, fermions and vectors, respectively, and the constants k_i depend on the renormalization scheme. The two-loop contributions, which we are particularly interested in, here, are of the schematic form

$$V^{(2)} = \sum_{i,j} g^{iijj} f_{ij}(m_i^2, m_j^2, Q) + \sum_{i,j,k} |g^{ijk}|^2 f_{ijk}(m_i^2, m_j^2, m_k^2, Q) , \tag{2.166}$$

where g^{ijkl} and g^{ijk} are field dependent four- and three-particle couplings and the functions $f_{ij}(x, y, Q)$ and $f_{ijk}(x, y, z, Q)$ are the results of the two-loop integrals, which depend on the renormalization scale.

Turning on a background effectively produces masses for both the scalar and the gauge fields. It also leads to cross-terms between gauge and scalar fields., which need to be eliminated by adding a gauge fixing term to the Lagrangian

$$\mathcal{L}_{\text{gf}} = \frac{g^2}{2} \left[R^1, \phi \right] \left[\phi, R^1 \right] , \tag{2.167}$$

which gives rise to a mass also for the scalar field along the steep direction [126]. Having masses for the fields means, that we first have to transform them to square-mass eigenstates. Since our theory comprises a deformation of the scalar sector, I will focus on the scalar part of the Lagrangian in the following. Its kinetic part contains terms like

$$-\mathcal{L} = \frac{1}{2} m_{ij}^2 R_i' R_j' , \tag{2.168}$$

where i, j run over all the (real) scalar fields and m_{ij}^2 are real symmetric matrices, which depend on the classical background field. The primes denote, that the scalars are not yet in a squared-mass eigenstate. They are rotated like

$$R_i' = N_{ji}^{(S)} R_j . \tag{2.169}$$

by an orthogonal matrix $N^{(S)}$ defined by

$$N_{ik}^{(S)} m_{kl}^2 N_{jl}^{(S)} = \delta_{ij} m_i^2 , \tag{2.170}$$

where m_i^2 are the scalar squared-mass eigenvalues. This basis rotation also has an effect on the interaction terms of the theory, which we denote, again for the scalar sector, as

$$\mathcal{L}_S = -\frac{1}{6} \lambda^{ijk} R_i R_j R_k - \frac{1}{24} \lambda^{ijkl} R_i R_j R_k R_l . \tag{2.171}$$

For the calculation of the effective potential, the couplings have to be used in this basis as defined here. The interactions λ^{ijk} and λ^{ijkl} are symmetric in all their indices and real. They generically depend on the classical background field as they depend on the rotation matrix $N^{(S)}$. Note that we are suppressing interactions of the scalars with other fields, since those do not differ from the standard $\mathcal{N}=4$ SYM theory.

In our case, following [106], we expand the (real) scalar fields around a constant background $\Phi^1(x) = \phi U$ for the steep direction, such that $\mathrm{tr}\, U^2 = 1$. Observe that U can be chosen to be diagonal to simplify the calculations. This background introduces effective masses which can be read off from the interaction terms. We have already mentioned the expressions for the masses of the various modes in this background

$$\Phi^1 = \phi U + R^1 \tag{2.172}$$

$$\Phi^i = 0 + R^i \ , \quad i = 2\ldots 6 \ . \tag{2.173}$$

We should carefully check that with the double trace interaction the fields are still in a square mass eigenstate for the chosen basis. For the off-diagonal modes, the contribution to the squared masses from the double trace interaction are

$$(M^1_{ab})^2_f \sim \frac{fa^2}{2N^2}\left[\mathrm{tr}\left(\phi^2 U^2 + (R^1)^2 + 2\phi U R^1 - \frac{1}{5}(R^i)^2\right)\right] \tag{2.174}$$

from which the masses only comprise the terms quadratic in the fields R, namely

$$(M^1_{ab})^2_f = \frac{fa^2}{2N^2}\left(\phi^2 \underbrace{U_{aa}U_{aa}}_{\mathrm{tr}\,U^2=1} + R^1_{ab}R^1_{ba} + 2\phi U_{aa}R^1_{aa}\right)^2 \tag{2.175}$$

$$= \frac{fa^2}{2N^2}\left(2\phi^2 R^1_{ab}R^1_{ab} + 4\phi^2 U_{aa}R^1_{aa}U_{cc}R^1_{cc}\right) \ , \tag{2.176}$$

where a and c are different indices to be summed over. In the last term in the last line we see that there are cross-terms between different $SU(N)$ degrees of freedom. Hence, we still need to diagonalize these masses.

The way I propose to do this was by rotating the fields R to a basis, with one element parallel and all others perpendicular to the background U

$$R^i_{\parallel} = R^{i\alpha}U^{\alpha}U \tag{2.177}$$

$$R^i_{\perp} = R^i - R^{i\alpha}U^{\alpha}U \ , \tag{2.178}$$

where we have used the scalar product on $SU(N)$, $X\cdot Y = \mathrm{tr}\,(XY)$ and the normalization of the background $\mathrm{tr}\, U^2 = 1$. In this basis, the mass squared term looks like

$$(M^1_{ab})^2_f = \frac{fa^2}{2N^2}\left(\underbrace{6\phi^2}_{\frac{1}{2}m^2_{\parallel}}(R^1_{\parallel})^2 + \underbrace{2\phi^2}_{\frac{1}{2}(m^2_{\perp})^a} R^{1a}_{\perp}R^1_{\perp a}\right) \ , \tag{2.179}$$

where the index a runs over the $N^2 - 1$ orthogonal components of R.

With all these data at hand, it is easy to verify the one-loop result for the effective potential. It is most concisely expressed in the $\overline{\mathrm{DR}}'$ renormalization scheme as

$$V^{(1)} = \sum_i (-1)^{(2s_i)}(2s_i + 1)h(m_i^2) \ , \tag{2.180}$$

with

$$h(x) = \frac{x^2}{4}\left(\ln\frac{x}{Q^2} - \frac{3}{2}\right) \ . \tag{2.181}$$

I have already commented above in subsection 2.4.1 that summing over all the scalar fields will produce the curious effect that any contributions to the effective potential with both single and double trace interactions vanish.

This is why I turn my attention to the two-loop level, again restricting myself to the sector with scalar interactions, only. The contributions to the effective potential are

$$V^{(2)}_{SSS} = \sum_{i,j,k} \frac{1}{12}(\lambda^{ijk})^2 f_{SSS}(m_i^2, m_j^2, m_k^2) \ , \tag{2.182}$$

$$V^{(2)}_{SS} = \sum_{i,j} \frac{1}{8}\lambda^{iijj} f_{SS}(m_i^2, m_j^2) \ . \tag{2.183}$$

The loop-integral functions are given in terms of the standard functions

$$f_{SSS}(x, y, z, Q) = -I(x, y, z, Q) \ , \tag{2.184}$$

$$f_{SS}(x, y, Q) = J(x, y, Q) \ . \tag{2.185}$$

which were introduced in [127] as

$$J(x, y, Q) = xy(\ln\frac{x}{Q^2} - 1)(\ln\frac{y}{Q^2} - 1) \ , \tag{2.186}$$

$$I(x, y, z, Q) = \frac{1}{2}(x - y - z)\ln\frac{y}{Q^2}\ln\frac{z}{Q^2} + \frac{1}{2}(y - x - z)\ln\frac{x}{Q^2}\ln\frac{z}{Q^2} \tag{2.187}$$

$$+ \frac{1}{2}(z - x - y)\ln\frac{x}{Q^2}\ln\frac{y}{Q^2} \tag{2.188}$$

$$+ 2x\ln\frac{x}{Q^2} + 2y\ln\frac{y}{Q^2} + 2z\ln\frac{z}{Q^2} - \frac{5}{2}(x + y + z) - \frac{1}{2}\xi(x, y, z) \ , \tag{2.189}$$

where the function $\xi(x, y, z)$ is expressed in terms of dilogarithms

$$\frac{\xi(x, y, z)}{R} = 2\ln\frac{z + x - y - R}{2z}\ln\frac{z + y - x - R}{2z} - \ln\frac{x}{z}\ln\frac{y}{z} \tag{2.190}$$

$$- 2\mathrm{Li}_2\frac{z + x - y - R}{2z} - 2\mathrm{Li}_2\frac{z + y - x - R}{2z} + \frac{\pi^2}{3} \tag{2.191}$$

with

$$R = \sqrt{x^2 + y^2 + z^2 - 2xy - 2xz - 2yz} \ . \tag{2.192}$$

This information specifies the two-loop potential of the scalar sector completely, which can now be calculated systematically on a computer.

I report on the calculation performed with the Mathematica Software. To simplify the computational load, I have considered a model with $N_f = 2$ scalar fields and rank $N = 2$ of the gauge group. The qualitative result of this calculation can then be extended to large N by performing a well-defined 't Hooft limit. To account for the fact that we only consider two scalar fields in this example, the double trace operator is adjusted to be $\mathcal{O} = \mathrm{tr}\,\left[(\Phi_1)^2 - (\Phi_2)^2\right]$. I have chosen a diagonal background whose trace is normalized, $\mathrm{tr}\, U^2 = 1$, such that

$$U = \begin{pmatrix} 1/\sqrt{2} & \\ & -1/\sqrt{2} \end{pmatrix} . \tag{2.193}$$

The result is lengthy and therefore not reproduced here in full. Instead, I only note its important properties. The crucial observation is that a term linear in the double trace coupling f appears at the two loop level. This is precisely the term of the sort, which we have argued to be necessary to have the potential being regularized and turned around in section 2.4.3.

The effective potential obtained in this way still explicitly depends on the renormalization scale. To compare with the Coleman-Weinberg potential as obtained in [106], I want to identify the scale Q with the value of the scalar field ϕ. The correct way to do this is to ensure that the effective potential is invariant under RG-transformations, yielding β-functions for the couplings and anomalous dimensions for the fields (cf. section 7 in [125]) via

$$Q\frac{\mathrm{d}\,V}{\mathrm{d}Q} = \left(Q\frac{\partial}{\partial Q} + \beta_g \frac{\partial}{\partial g} + \beta_f \frac{\partial}{\partial f} - \sum_{i=1}^{6} \gamma_i \Phi^i \right) V_{\mathrm{eff}} = 0 \, . \tag{2.194}$$

The β functions and anomalous dimensions can be extracted order by order (cf. appendix B.4)

$$Q\frac{\partial}{\partial Q} V^{(1)} \quad + \left(\beta_g^{(1)} \tfrac{\partial}{\partial g} + \beta_f^{(1)} \tfrac{\partial}{\partial f} - \textstyle\sum_i \gamma_i^{(1)} \Phi^i \tfrac{\partial}{\partial \Phi^i}\right) V^{(0)} \quad = 0 \, , \tag{2.195}$$

$$Q\frac{\partial}{\partial Q} V^{(2)} \quad + \left(\beta_g^{(1)} \tfrac{\partial}{\partial g} + \beta_f^{(1)} \tfrac{\partial}{\partial f} - \textstyle\sum_i \gamma_i^{(1)} \Phi^i \tfrac{\partial}{\partial \Phi^i}\right) V^{(1)} \tag{2.196}$$

$$+ \left(\beta_g^{(2)} \tfrac{\partial}{\partial g} + \beta_f^{(2)} \tfrac{\partial}{\partial f} - \textstyle\sum_i \gamma_i^{(2)} \Phi^i \tfrac{\partial}{\partial \Phi^i}\right) V^{(0)} \quad = 0 \, . \tag{2.197}$$

The first order RG-equation for the double trace coupling is found to be

$$-\frac{29}{4\pi^2} f(Q)^2 - Q f'(Q) = 0 \, , \tag{2.198}$$

which we solve to find the renormalized coupling

$$f(Q) = -\frac{4\pi^2}{4\pi^2 C - 29\log(Q)} \, , \tag{2.199}$$

where C is a constant of integration to be determined below. For the second order RG-equation, we neglect the logarithmic terms for simplicity and find

$$-\frac{1}{4\pi^4}\phi^4 \left[84 g^4 f(Q) + \left(130 g^2 + 29\pi^2\right) f(Q)^2 + 4\pi^4 Q f'(Q)\right] = 0 \, , \tag{2.200}$$

which is solved to obtain the two-loop renormalized double trace coupling

$$f(Q) = -\frac{84g^4}{e^C\left(130g^2 + 29\pi^2\right) - Q^{\frac{21g^4}{\pi^4}}} \,. \tag{2.201}$$

We substitute these renormalized couplings back into the one- and two-loop expressions for the effective potential, respectively. Note that we have made an adiabatic approximation for the Yang-Mills coupling: Since there is a term in the effective potential, which contains both couplings, we know that also the single-trace coupling has a non-trivial β-function. However, we assume that g flows much more slowly than f. In principle, we would have to solve the coupled set of RG-equations for both couplings. It is, however, not possible to extract the β-functions for both couplings and the anomalous dimensions from the effective potential. Those are rather needed as an extra input, which can for instance be extracted from the expansion in Feynman diagrams. As an approximation, this assumption is, however, justified, as we can also see from the coupled RG-flow as presented in figures 2.14 and 2.15.

I have argued in section 2.4.3 that for a turnaround to happen, the coefficient of the linear term in the β-function of f must be bigger than 4. To check this condition, I read off this coefficient from equation (2.200) to be

$$-\frac{21g^4}{\pi^4} \tag{2.202}$$

which generalizes to $-\frac{84}{\pi^4}\frac{(g^2N)^2}{N^4}$ for arbitrary N. We see that this condition is only met if the Yang-Mills coupling is larger than

$$g \geq \frac{\pi}{21^{1/4}} \approx 1.5 \,. \tag{2.203}$$

Although this is bigger than unity, we can still trust perturbation theory, because the expansion is actually not in the coupling g, only, but rather in the ratio $\frac{g}{4\pi} \approx 0.12$, which is still smaller than unity. To make sure that this condition is met, we can just tune the 't Hooft coupling to be large enough. This, however, appears not to be necessary. As long as the Yang-Mills sector is free in the IR, g will be driven to a larger value at higher scales and the absolute value of the coefficient will eventually be large enough at some scale. The other condition, which has to be met for the linear term in f to make the potential turn around is that the sign of the coefficient is negative. This appears to be correct in the example calculation performed, but it should be robust on general grounds. We know that the double trace operator is marginally irrelevant. This means that the theory should be asymptotically free in the double-trace coupling, which a $1/N$ suppressed effect should not change. Therefore, the sign of this term in the β-function should be negative, as required.

To finally obtain the result, we need to determine the constant of integration C in (2.198) and (2.200). We do this by matching the one-loop and two-loop potentials at a point for a specific value for ϕ. Then, the one-loop and two-loop RG-equations are solved consistently. Since the double-trace coupling is asymptotically free, this matching can best be done at a large value of ϕ, infinity, say, where the two-loop effective potential approaches zero. However, the region, where we can best trust the perturbative treatment is for small values of ϕ.

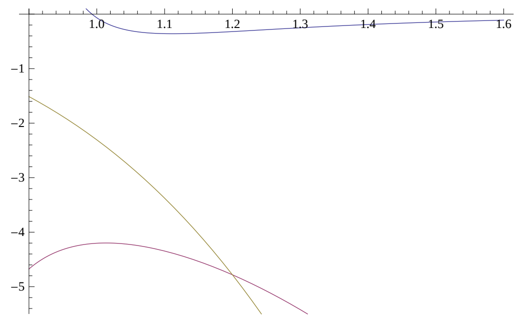


Figure 2.16: *A plot of the tree level (green), one-loop (red) and two-loop (blue) contributions to the effective potential. The value chosen for the Yang-Mills coupling is $g = 2.5$, which results in a coefficient for the term linear in f in the β-function of about $8.4 > 4$, as required for a regularization of the potential. We see, that the potential, indeed, turns around and that quantum corrections generate a ground state for the scalar sector at the two-loop level. In this plot, I have ignored the contributions of the logarithms ins the effective potential.*

When examining the results, we encounter a hitch. Since we have ignored the logarithmic terms in the two-loop β-function it seems consistent to also ignore the logarithmic terms in the two-loop effective potential after replacing the double trace coupling by its renormalized version. The result of this procedure is illustrated in figures 2.16 and 2.17. In the former plot, I also show the one-loop and tree-level potentials. It was surprisingly not possible to find a value for the integration constant such that the one- and two-loop potentials would intersect. Apart from that, the effective potential shows the expected behavior: It starts out negative around the renormalization scale, turns around and asymptotes to zero owing to the asymptotic freedom.

One could conclude that neglecting the logarithmic terms was not appropriate. In figure 2.18, I plot the result keeping those contributions after replacing the bare with the renormalized coupling. Surprisingly, the logarithmic terms seem to have a large impact on the effective potential, which is now positive and diverges at infinity. Conveniently, it still shows a turnaround, though. A possible point of further investigation is, if it leads to a better control of the potential to include the logarithmic terms also when solving the two-loop RG-equations.

At the end of the day, the important criterion which determines whether the field theory constitutes a well-defined way to describe the dual cosmological singularity is whether it has a ground state. If the effective potential turns around, we can still draw useful conclusions from our result. We see, indeed, that the potential turns around and now features a minimum in both cases. This leads to a stable ground state for the field theory and defeats the problem of unitarity loss in the evolution of the scalar field, because the scalar will in fact thermalize around that ground state.

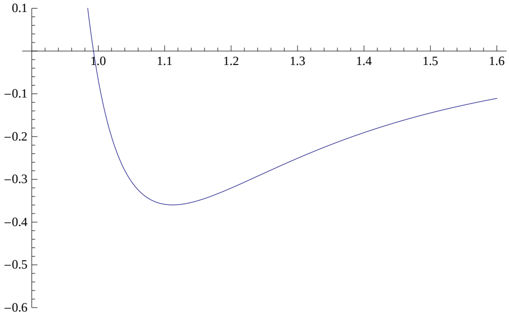


Figure 2.17: *Same plot as 2.16 zoomed in to the region, where the two-loop potential turns around.*

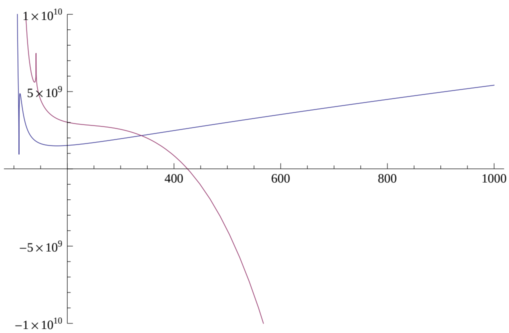


Figure 2.18: *This plot shows the one-loop (red) and two-loop (blue) effective potentials including their logarithmic terms. We see that the two-loop potential is still bounded from below.*

2.6 THE BULK INTERPRETATION

Now that we have found that taking into account $1/N$ corrections bound the effective potential of the boundary theory, which is well under calculational control, we are in a position to ask ourselves how the resolution of the cosmological, space-like singularity works in the gravitational bulk itself. It has been argued in [105] that a regularization of the effective potential by adding a higher dimensional operator in the boundary field theory will lead to the formation of a black hole with a horizon which covers up the singularity. Such a regularization has the disadvantage that the field theory becomes non-renormalizable. In our case, however, the regularization arises naturally without adding any irrelevant operators. Rather, the corrections contained in the limit of the marginal operator, which bound the potential. Therefore, the theory remains renormalizable and the situation remains well-understood all the way to the cut-off.

We review the behavior of the bulk in the present context (cf. [106]). The appropriate initial data for any boundary conditions is obtained by slicing an $O(5)$-invariant Euclidean instanton of the form

$$\mathrm{d}s^2 = \frac{\mathrm{d}\rho^2}{b^2(\rho)} + \rho^2 \mathrm{d}\Omega_4 \tag{2.204}$$

with $\phi = \phi(\rho)$ through its center. The instanton field equations with boundary conditions $\alpha_f = f\beta$ determine b to be

$$b^2(\rho) = \rho^2 + 1 + \frac{2\alpha^2(\log\rho)^2}{2\rho^2} + \frac{\alpha(4\beta - \alpha)\log\rho}{3\rho^2} + \frac{8\beta^2 - 4\alpha\beta + \alpha^2}{12\rho^2} \,, \tag{2.205}$$

and the scalar field obeys

$$b^2\phi'' + \left(\frac{4b^2}{\rho} + bb'\right)\phi' - R_{\mathrm{AdS}}^2 V_{,\phi} = 0 \,, \tag{2.206}$$

with $' = \partial_\rho$. The mass of this initial data for the Lorentzian solution is

$$M = -\frac{\pi^2 R_{\mathrm{AdS}}^2 f^2 \beta^2}{4} \,. \tag{2.207}$$

Hence, the instanton specifies negative mass initial data.

Numerically integrating the Einstein equations, one can show that the theory admits static, spherical black holes with the chosen boundary conditions [106]. In particular, there is precisely one black hole with scalar hair for any given horizon size. Its mass is given by

$$M_{\mathrm{hbh}} = 2\pi^2 R_{\mathrm{AdS}}^2 \left[\frac{3}{2}M_0 + \beta^2\left(1 - \frac{1}{2}f\right)\right] \,, \tag{2.208}$$

where M_0 is the mass of an equally large, bald, usual Scharzschild black hole corresponding to the standard vacuum with $\langle\mathcal{O}\rangle = 0$. We see that the scalar hair adds some mass to it and the smaller f, the bigger the mass increase. In particular, the mass M_{hbh} is always positive. If the potential is unbounded, there is no black hole to conceal the singularity and it extends all the way to the boundary. The hairy black hole then represents an excitation about the local maximum of the field theory potential.

We can change the boundary conditions such that negative mass solutions become available. The example treated in [106] is to adjust the boundary conditions to

$$\alpha_{f,\epsilon} = f\beta - \epsilon\beta^3 \ . \tag{2.209}$$

For a sufficiently small parameter ϵ a negative mass black hole exists with the same mass as the instanton $M \sim -\epsilon R^2_{\mathrm{AdS}}$ and hence is the natural final state of the bulk evolution. Changing the boundary conditions for the bulk scalar field, however, will source a different operator in the field theory. Inspecting relation (2.83)

$$\alpha = -\frac{\delta W}{\delta \beta} \ , \tag{2.210}$$

we see that the corresponding deformation of the field theory is

$$W = -\frac{f}{2}\beta^2 + \frac{\epsilon}{4}\beta^4 \ , \tag{2.211}$$

where $\beta = \langle \mathcal{O} \rangle$ is the expectation value of the single trace operator. This means that changing the boundary conditions in the bulk such that negative mass black holes are accessible corresponds to regularizing the field theory potential by an irrelevant quadruple trace operator. In the limit $\epsilon \to 0$, we recover the case of the unbounded potential, for which the singularity spreads to the boundary.

In the previous section 2.5, I have shown that it is not necessary to regularize the potential by hand. Rather, the contributions subleading in $1/N$ of the double trace operator automatically regularize the potential. In the large-N limit, the potential, however, is unbounded just as in the case of (2.211) for $\epsilon \to 0$. The double trace operator is marginally irrelevant and as such resembles the regularization in (2.211). Yet, since the contribution which ensures the turnaround of the potential is suppressed by $1/N^2$, it should correspond to higher curvature and string loop corrections in the bulk, which are not captured by the supergravity approximation. Therefore, the expectation is that the singularity in the bulk is resolved by a "small" black hole, whose horizon is only supported by quantum corrections, with scalar hair.

This also matches with our expectation on the field theory side, in which the final state will be thermal. On the Poincaré patch, the geometry can only be Euclidean AdS with zero temperature or an AdS black hole, which is thermal. Indeed, we see in figure 2.19 that the Penrose diagrams of a black hole in AdS space and of a space-like singularity stretching all the way to the boundary are almost the same. The black hole case can be seen as the limit of the cosmological case in which the time at which the singularity hits the boundary becomes larger and larger until it is infinite for the case of an eternal black hole.

So far, our picture explains how a big crunch singularity is resolved into a black hole, which forms the thermal endpoint of the evolution. We should keep in mind that we were using a setup in supergravity, while the $1/N$ corrections in the dual field theory suggest that string

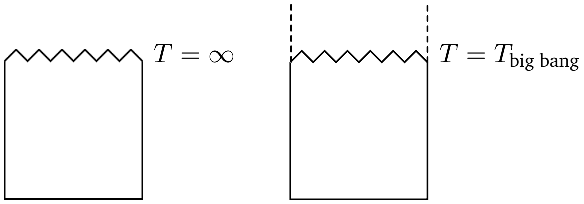


Figure 2.19: *In their Penrose diagrams, we see that an AdS space with a black hole singularity and with a big bang singularity are very similar. Only the upper corners of the diagram are changed. Whereas the time around the singularity at infinity goes on forever, the big bang singularity hits the boundary at some finite time.*

loop corrections are important and this picture is necessarily incomplete. In particular, the truncation to a theory with a fixed dilation breaks down as the singularity approaches the boundary. It remains an interesting question if it is possible to re-interpret this black hole in a more complete picture as the starting point of the evolution of the universe, the resolved big bang.

2.7 Conclusions

The aim of this study was to examine, if effects of quantum gravity resolve cosmic singularities using a holographic Super-Yang-Mills description, which is a well-defined quantum theory of type IIB supergravity on AdS_5. If a low-energy effective description of gravity descends from a UV complete theory, one expects in general that it does not contain any singularities. The supergravity theory I have examined is related to type IIB string theory in ten dimensions compactified on an S^5 and full string theory is considered to be a consistent theory of quantum gravity. Quantum gravity must resolve cosmological singularities.

In the specific example I have studied, the scalar field potential in the bulk is unbounded from below. As the bulk scalar field rolls down the potential to infinity, a space-like singularity forms. If quantum corrections do not bound this potential, the theory remains ill-defined and needs to be discarded. Here, this can be seen as follows. The scalar field describes one of the quadrupole distortions of the S^5 on which the bulk is compactified. Since it is driven down an unbounded potential, this means that the sphere becomes highly squashed, signaling a breakdown of the low-energy effective description. The mass of the background scalar considered here satisfies the Breitenlohner-Friedman bound and the background is perturbatively stable. Therefore it appears as if the gravity side is not fundamentally flawed and we think that quantum effects will play an important rôle.

The unbounded potential in the bulk is replicated by an unbounded potential in the dual field theory. The conformal field theory is deformed by a double trace deformation, which breaks the superconformal symmetries. This implies that the coupling constant of the double trace coupling flows. Although supersymmetry is broken by the double trace interaction, we have seen that the resident R-symmetry has prolonged the protection of the Yang-Mills coupling to the one-loop level, but it is hard to conceive, that it extends to even higher orders. Therefore, one expects that the RG-flows of the two couplings mix and also the Yang-Mills coupling gets renormalized. In the β-function of any one coupling, the other one appears as a coefficient. This means that, in fact, the RG-flow of the two couplings is described by a system of coupled differential equations and the flow of one coupling influences the flow of the other. The main point is that by neglecting the running of the gauge coupling the well-established results on the leading qualitative UV behavior of the double trace coupling can be misleading. Inclusion of the running gauge coupling can result in the gauge theory becoming an effective field theory, only valid up to some UV cut-off, by turning the double trace operator into an irrelevant deformation. Or it could lead to a turn-around in the effective potential, stabilizing the dynamics. What happens crucially depends on the coefficients in the beta-functions, which are determined by perturbative analysis of the deformed gauge theory.

We have seen in the calculation performed here on the field theory side, that at the two loop level, indeed, the renormalization of the single- and double-trace couplings mix. In particular, a term linear in the double trace coupling survives at the two-loop level, which can turn the effective potential around and, thus, provide the field theory with a ground state generated by quantum effects. The critical condition for this to happen is that the modulus of the coefficient of this term is big enough, namely bigger than four. This is not at all ensured a priori. As mentioned above, this coefficient contains in particular a power of the Yang-Mills coupling g, whose value is arbitrary at the conformal fixed point. As such, it is at first not possible to determine the value of this coefficient. However, conformal invariance is broken at the one-loop level for the double trace coupling and at the two-loop level for the single trace coupling. The latter is growing with the RG-flow until it will inevitably be big enough to ensure the turnaround of the effective potential to happen. Inclusion of $1/N$ corrections bounds the effective potential.

This could have been expected in retrospect, since the unavoidable running of the gauge coupling signals the presence of a dynamical dilaton field in the bulk, which one would indeed expect to become an important factor as one approaches the (spreading) crunch singularity. The background solution we have chosen has a dilaton fixed at its expectation value. This means, that the string coupling, which corresponds to the Yang-Mills coupling, is constant. As soon as this coupling flows and the dilaton becomes dynamical, the truncation used is too restrictive. As the dilaton grows it also influences the scalar field potential in the bulk. Hence, the truncation to supergravity with only one scalar field is rendered invalid. String loop corrections can no longer be neglected.

Since quantum corrections regularize the potential the theory has now a ground state. The evolution of the scalar fields ends in a thermalization around this well-defined minimum and unitarily loss in the field theory is avoided. For the bulk gravity, this means that the spreading of the singularity towards the boundary will stop and that it is covered up by a huge black hole, which conceals the singularity. Thus, I have explained using a specific example, how $1/N$ corrections of a field theory dual to a string theory in an unstable bulk can resolve a cosmological singularity.

Appendix A

Two-loop renormalization of ϕ^4-theory

We are looking at a theory with a single scalar field ϕ and a potential $-\frac{\lambda}{4}\phi^4$. This is the harmonic oscillator potential upside down, so the configuration is unstable and the field rolls down this potential to infinity. The counterterm necessary to renormalize this interaction is $\frac{\delta_\lambda}{4}\phi^4$ and the Lagrangian for the renormalized fields and parameters is

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}\phi^4 - \frac{1}{2}\delta_Z(\partial_\mu\phi)^2 - \frac{1}{2}\delta_m\phi^2 + \frac{\delta_\lambda}{4}\phi^4 \,. \tag{A.1}$$

Note that with this definition of the interaction term, the vertex Feynman rule should be $6\mathrm{i}\lambda$. The metric used is

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} . \tag{A.2}$$

In the following we are calculating the contribution of one- and two-loop diagrams to δ_λ, which we denote as $\delta_\lambda^{(1)}$ and $\delta_\lambda^{(2)}$, respectively.

A.1 One-loop renormalization

On the level of one-loop interactions the following diagrams contribute to the four-point function:

The three loop diagrams just represent different channels corresponding to the Mandelstam variables $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$ and $u = (p_1 - p_4)^2$, respectively. The last diagram is the vertex counterterm defined to cancel the divergences of the first three diagrams.

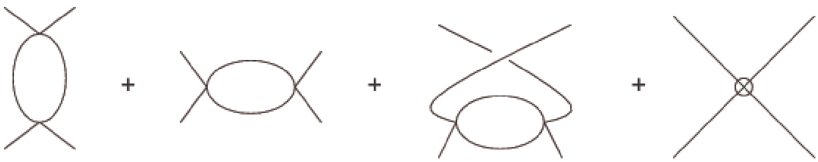


Figure A.1: *One-loop diagrams and the counterterm for the four-point function.*

The first three diagrams can be evaluated at once. We obtain their amplitude in 4 dimensions by integrating over the free momentum k and multiplying with the two interaction vertices

$$(6\mathrm{i}\lambda)^2 \mathrm{i}V(p^2) \equiv \frac{(6\mathrm{i}\lambda)^2}{2} \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{-\mathrm{i}}{k^2+m^2} \frac{-\mathrm{i}}{(k+p)^2+m^2} \,, \tag{A.3}$$

where $\frac{1}{2}$ is the symmetry factor of the diagram. We denote the momentum flow in the loop with two insertions of momentum p and $-p$, respectively, with $V(p^2)$. For each diagram, p^2 should be replaced by the corresponding Mandelstam variable and the three contributions are added. The vertex is worth $6\mathrm{i}\lambda$ and this is the contribution we have to add on top for the tree level diagram to get the entire amplitude at one-loop level:

$$\mathrm{i}\mathcal{M}^{(1)} = 6\mathrm{i}\lambda + (6\mathrm{i}\lambda)^2[\mathrm{i}V(s) + \mathrm{i}V(t) + \mathrm{i}V(u)] + 6\mathrm{i}\delta_\lambda^{(1)} \,, \tag{A.4}$$

where the last term $\delta_\lambda^{(1)}$ denotes the order λ contribution to the counterterm.

We impose the following renormalization conditions:

$$\mathrm{i}\mathcal{M}(p_1p_2 \to p_3p_4) = 6\mathrm{i}\lambda \quad \text{at} \quad s = t = u = \mu^2 \,, \tag{A.5}$$

which relates the physical coupling to a renormalization scale μ, which later on will be related to the value of the scalar field. Due to this condition the second and third terms should cancel each other if the incoming momenta equal the renormalization scale. There, the value of the square bracket just becomes $3\mathrm{i}V(\mu^2)$ and the counterterm can be read of as

$$\begin{aligned} \delta_\lambda^{(1)} &= (6\mathrm{i}\lambda)^2 \cdot \left(-\frac{1}{2}V(\mu^2)\right) \\ &= 18\lambda^2 V(\mu^2) \,. \end{aligned} \tag{A.6}$$

We now evaluate (A.3) using dimensional regularization. This implies that every dimensionful parameter must be expressed in terms of a dimensionless number times the appropriate power of the renormalization scale, in order to ensure that the overall dimension of the Lagrangian works out. So in the following, we replace $m^2 = a\mu^2$.[1] Combining the two denominators by use of a Feynman-parameter (A.3) looks like

$$\mathrm{i}V(p^2) = -\frac{1}{2}\int_0^1 \mathrm{d}x \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{1}{[k^2+2xkp+xp^2+a\mu^2]^2} \,. \tag{A.7}$$

[1] We actually also would have to adjust the interaction term by including a factor of μ^{4-d}, however, this factor will vanish in the expansion around $d = 4$ to first order in ϵ, since $\lambda\mu^{4-d} \simeq \lambda(1 + \ln \mu\epsilon)$.

We shift the integration variable to $l = k + xp$ and obtain

$$\mathrm{i}V(p^2) = -\frac{1}{2}\int_0^1 \mathrm{d}x \int \frac{\mathrm{d}^d l}{(2\pi)^d} \frac{1}{[l^2 + x(1-x)p^2 + a\mu^2]^2} \ , \tag{A.8}$$

rotate to Euclidean space by performing a Wick rotation

$$l_E^0 = -\mathrm{i}l^0 \ , \qquad \overrightarrow{l_E} = \overrightarrow{l} \ , \tag{A.9}$$

which leads to

$$\begin{aligned}\mathrm{i}V(p^2) &= -\frac{\mathrm{i}}{2}\int_0^1 \mathrm{d}x \int \frac{\mathrm{d}^d l_E}{(2\pi)^d} \frac{1}{[l_E^2 + x(1-x)p^2 + a\mu^2]^2} &\text{(A.10)}\\ &= -\frac{\mathrm{i}}{2}\int_0^1 \mathrm{d}x \frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{\frac{d}{2}}} \left[\frac{1}{x(1-x)p^2 + a\mu^2}\right]^{2-\frac{d}{2}} &\text{(A.11)}\\ &= -\frac{\mathrm{i}}{2}\int_0^1 \mathrm{d}x \frac{\Gamma(\frac{\epsilon}{2})}{(4\pi)^{2-\frac{\epsilon}{2}}} \frac{1}{[a\mu^2 + x(1-x)p^2]^{\frac{\epsilon}{2}}} \ , &\text{(A.12)}\end{aligned}$$

where in the last line we have replaced $\epsilon = 4 - d$. In the limit $d \to 4$ or $\epsilon \to 0$, the Gamma-function has a pole and its approximation looks

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma + \frac{1}{2}\left(\gamma^2 + \frac{\pi^2}{6}\right)\epsilon + \mathcal{O}(\epsilon^2) \ , \tag{A.13}$$

where γ is the EULER-MASCHERONI constant. Altogether we therefore get

$$\begin{aligned}\cdots \xrightarrow{d\to 4} -\frac{\mathrm{i}}{32\pi^2}\left(\frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon)\right)\left(1 + \frac{\log(4\pi)}{2}\epsilon\right)\\ \times\left[1 - \frac{\epsilon}{2}\int_0^1 \mathrm{d}x \log(a\mu^2 + x(1-x)p^2)\right] \ ,\end{aligned} \tag{A.14}$$

where the logarithmic terms come from a Taylor expansion of the denominators around $d = 4$. The integral evaluates to

$$-2 + \frac{2\sqrt{-4a\mu^2 - p^2}}{p} \arctan \frac{p}{\sqrt{-4a\mu^2 - p^2}} + \log a\mu^2 \ . \tag{A.15}$$

Assembling this into (A.14) we obtain the one-loop amplitude in dimensional regularization up to first order in ϵ

$$\begin{aligned}\mathrm{i}V(p^2) = -\frac{\mathrm{i}}{32\pi^2}\Bigg(&\frac{2}{\epsilon} + \log 4\pi - \gamma + 2 - \log a\mu^2\\ &-\frac{2\sqrt{-4a\mu^2 - p^2}}{p} \arctan \frac{p}{\sqrt{-4a\mu^2 - p^2}} + \mathcal{O}(\epsilon)\Bigg)\end{aligned} \tag{A.16}$$

Inserting this result into equation (A.6) we obtain for the shift of the coupling constant at one-loop order

$$\delta_\lambda^{(1)} = -\frac{9\lambda^2}{16\pi^2}\left(\frac{2}{\epsilon} + \log 4\pi - \gamma + 2 - \log a\mu^2 - 2\sqrt{-4a-1}\arctan\frac{1}{\sqrt{-4a-1}}\right) \ . \tag{A.17}$$

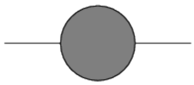
Note that the result agrees with the one given in [106], appendix B,[2]

$$\delta_\lambda^{(1)} = -\frac{9\lambda^2}{16\pi^2}\left(\frac{2}{\epsilon} - \log m^2 + \text{finite}\right) . \tag{A.18}$$

Eventually, we will be interested in the massless theory. We can then discard off the mass term in the denominator in (A.12), which simplifies the evaluation. We obtain

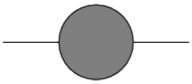
$$\mathrm{i}V_{\mathrm{m=0}}(p^2) = -\frac{\mathrm{i}}{32\pi^2}\left(\frac{2}{\epsilon} + 2 + \log 4\pi - \gamma - \log p^2\right) . \tag{A.19}$$

In order to calculate the two-loop renormalization, we also need the mass and field renormalization constants. Those we obtain from the one loop corrections to the two-point functions. We define the sum of all one-particle-irreducible insertions into the propagator as $\mathbf{1PI} = -\mathrm{i}M^2(p^2)$. The second renormalization condition



$$= \frac{-\mathrm{i}}{p^2 - a\mu^2} \tag{A.20}$$

defines the pole of the propagator to be at $m^2 = a\mu^2$ and having residue 1. On the other hand, this propagator is defined by a geometric series as



$$= \mathbf{1PI} + \mathbf{1PI}^2 + \cdots = \frac{-\mathrm{i}}{p^2 - a\mu^2 - M^2(p^2)} . \tag{A.21}$$

With this at hand the renormalization condition can be restated as

$$M^2(p^2)\big|_{p^2=\mu^2} = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}p^2}M^2(p^2)\bigg|_{p^2=\mu^2} = 0 , \tag{A.22}$$

respectively. The one-loop divergence of the two-point function is cancelled by the $\delta_Z^{(1)}$ and $\delta_m^{(1)}$ terms in the action, which leads to the following relation for the two-point function:

$$\mathbf{1PI} = -\mathrm{i}M^2(p^2) = (6\mathrm{i}\lambda)\frac{1}{2}\int\frac{d^dk}{(2\pi)^d}\frac{-\mathrm{i}}{k^2 + a\mu^2} - \mathrm{i}(p^2\delta_Z^{(1)} + \delta_m^{(1)}) , \tag{A.23}$$

with $\frac{1}{2}$ being the symmetry factor of the one-loop diagram. Note that the sign of the δ_Z-term differs from standard renormalization procedure due to the different sign of the term in the action (A.1) chosen. We perform the integral and get

$$\cdots = -\frac{3\mathrm{i}\lambda}{(4\pi)^{\frac{d}{2}}}\frac{\Gamma(1-\frac{d}{2})}{(-a\mu^2)^{1-\frac{d}{2}}} - \mathrm{i}(p^2\delta_Z^{(1)} + \delta_m^{(1)}) . \tag{A.24}$$

Since the first term is independent of p^2 we conclude that

$$\delta_Z^{(1)} = 0 \qquad \text{and} \qquad \delta_m^{(1)} = -\frac{3\lambda}{(4\pi)^{\frac{d}{2}}}\frac{\Gamma\left(1-\frac{d}{2}\right)}{(-a\mu^2)^{1-\frac{d}{2}}} , \tag{A.25}$$

ensuring that $M^2(p^2)$ vanishes for all momenta as required by the renormalization conditions (A.22). So there is no contribution to $M^2(p^2)$ at the one-loop level. The vanishing of the field renormalization at one-loop order is a common feature of ϕ^4-theories.

[2] Note that $\arctan z = \frac{\mathrm{i}}{2}\left(\log(1+\mathrm{i}z) - \log(1+\mathrm{i}z)\right)$ and thus imaginary for imaginary z and well defined for $|z| < 1$. Hence, the combination $\sqrt{-4a-1}\arctan\frac{1}{\sqrt{-4a-1}}$ is real and well defined for $a > -\frac{1}{2}$.

A.2 Two-loop renormalization

A.2.1 Renormalization of the coupling

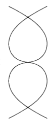
We are now going one loop further. The contributing diagrams are listed in [128], page 339, Fig. 10.5. We divide the two loop diagrams into three groups. So that the two-loop counterterm has three contributions

$$\delta_\lambda^{(2)} \sim d_{\rm I} + d_{\rm II} + d_{\rm III} \; . \tag{A.26}$$

To begin with, we only calculate the s-channel diagrams. The t and u-channels are then just the same under our renormalization conditions.

The first double loop diagram has the value

$$d_{\rm I} =$$



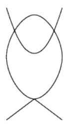
$$= (6{\rm i}\lambda)^3 \cdot [{\rm i}V(p^2)]^2 \; , \tag{A.27}$$

which, using the result (A.16), reads

$$d_{\rm I} = -(6{\rm i}\lambda)^3 \frac{1}{1024\pi^4} \left(\frac{4}{\epsilon^2} + \frac{4}{\epsilon}(2 + \log 4\pi - \gamma - \log p^2) + (2 + \log 4\pi - \gamma - \log p^2)^2 \right) \; . \tag{A.28}$$

The second diagram we have to calculate is

$$d_{\rm II} =$$



$$= (6{\rm i}\lambda)^3 \frac{1}{2} \int \frac{{\rm d}^d k}{(2\pi)^d} \frac{-{\rm i}}{k^2 + a\mu^2} \frac{-{\rm i}}{(k+p)^2 + a\mu^2} V((k+p_3)^2) \; , \tag{A.29}$$

in which we combine the first two propagator terms by multiplying and simplifying them after dropping the mass term and use our previous one-loop result (A.12).

$$\begin{aligned} d_{\rm II} &= (6{\rm i}\lambda)^3 \frac{1}{2} \int \frac{{\rm d}^d k}{(2\pi)^d} \int \frac{{\rm d}^d l}{(2\pi)^d} \frac{-{\rm i}}{k^2} \frac{-{\rm i}}{(k+p)^2} \frac{-{\rm i}}{l^2} \frac{-{\rm i}}{(l+k+p_3)^2} \\ &= (6{\rm i}\lambda)^3 \frac{1}{2} \int \frac{{\rm d}^d k}{(2\pi)^d} \frac{1}{(k^2+kp)^2} \int_0^1 {\rm d}x {\rm i} \int \frac{{\rm d}^d r_E}{(2\pi)^d} \frac{1}{[r_E^2 + x(k^2 + p_3^2 + 2kp_3) - x^2(k+p_3)^2]^2} , \end{aligned} \tag{A.30}$$

where we do the integral and obtain

$$d_{\rm II} = (6{\rm i}\lambda)^3 \frac{\rm i}{2(4\pi)^{\frac{d}{2}}} \Gamma(2 - \frac{d}{2}) \int \frac{{\rm d}^d k}{(2\pi)^d} \frac{1}{(k^2+kp)^2} \left(\frac{1}{(k+p_3)^2} \right)^{2-\frac{d}{2}} \int_0^1 {\rm d}x \left(\frac{1}{x(1-x)} \right)^{2-\frac{d}{2}} . \tag{A.31}$$

The last integral is just an Euler beta-function. We combine the two denominators by use of

(A.32)

[(10.56) in [128]] and get

$$\begin{aligned} d_{\rm II} &= (6\mathrm{i}\lambda)^3 \frac{\mathrm{i}}{2(4\pi)^{\frac{d}{2}}} \Gamma(4-\frac{d}{2}) B\left(\frac{d}{2}-1, \frac{d}{2}-1\right) \\ & \int_0^1 \mathrm{d}y \int \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{y^{1-\frac{d}{2}}(1-y)}{[y(k+p_3)^2+(1-y)(k^2+kp)]^{4-\frac{d}{2}}} \ . \end{aligned} \tag{A.33}$$

We introduce a shifted momentum $s = k + \frac{1}{2}(p(1-y)+2yp_3)$, Wick rotate to s_E and perform the momentum integral

$$\begin{aligned} d_{\rm II} &= -(6\mathrm{i}\lambda)^3 \frac{1}{2(4\pi)^d} \frac{\Gamma(4-d)\Gamma(\frac{d}{2}-1)^2}{\Gamma(d-2)} \int_0^1 \mathrm{d}y \frac{y^{1-\frac{d}{2}}(1-y)}{[yp_3^2 - \frac{1}{4}(p(1-y)+2yp_3)^2]^{4-d}} \\ &= -(6\mathrm{i}\lambda)^3 \frac{1}{2(4\pi)^{4-\epsilon}} \frac{\Gamma(\epsilon)\Gamma(1-\frac{\epsilon}{2})^2}{\Gamma(2-\epsilon)} \int_0^1 \mathrm{d}y \frac{y^{\frac{\epsilon}{2}-1}(1-y)}{[yp_3^2 - \frac{1}{4}(p(1-y)+2yp_3)^2]^{\epsilon}} \ . \end{aligned} \tag{A.34}$$

The latter integral over y can be done analytically and results in

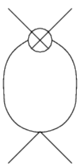
$$\begin{aligned} d_{\rm II} &= -(6\mathrm{i}\lambda)^3 \frac{1}{2(4\pi)^{4-\epsilon}} \frac{\Gamma(\epsilon)\Gamma(1-\frac{\epsilon}{2})^2}{\Gamma(2-\epsilon)} \frac{1}{(-p)^{2\epsilon}} \\ & \left(\frac{(-2)^\epsilon \Gamma(\frac{1-\epsilon}{2})\Gamma(\frac{\epsilon}{2})\, {}_2F_1(\frac{\epsilon}{2}, \epsilon, 1-\frac{\epsilon}{2}, \frac{(p-2p_3)^2}{p^2})}{\sqrt{\pi}} \right. \\ & \left. - \frac{(-4)^\epsilon \Gamma(1-\epsilon)\Gamma(1+\frac{\epsilon}{2})\, {}_2F_1(\frac{2+\epsilon}{2}, \epsilon, 2-\frac{\epsilon}{2}, \frac{(p-2p_3)^2}{p^2})}{\Gamma(2-\frac{\epsilon}{2})} \right) \ . \end{aligned} \tag{A.35}$$

To expand this expression around $\epsilon = 0$, we rewrite $p_3 = bp$ with $0 < b < 1$. Up to constant terms, we then find

$$\begin{aligned} d_{\rm II} &= -(6\mathrm{i}\lambda)^3 \frac{1}{2(4\pi)^4} \left[\frac{2}{\epsilon^2} + \frac{1-2\gamma+\log[256]-2(\log[-p]+\log[p])+2\log[\pi]}{\epsilon} \right. \\ & + \left(-\frac{1}{2} + \gamma^2 + \frac{\pi^2}{4} + 16\log[2]^2 + \log[16] + \frac{4(-1+b)b(\log[4-4b]+\log[b])}{(1-2b)^2} \right. \\ & + (\log[-p]+\log[p]) \left(-1 + 2\gamma + \log[p] + \log\left[-\frac{p}{256\pi^2}\right] \right) \\ & \left. + \log\pi(1+\log[256]+\log[\pi]) - \gamma(1+\log[256]+2\log[\pi]) + \mathrm{Li}_2\left((1-2b)^2\right) \right) + \mathcal{O}(\epsilon) \end{aligned} \tag{A.36}$$

We find a nonlocal divergence in the latter expression which should be cancelled by the corresponding counterterm, which we calculate in the diagrams containing the one-loop counterterms.

$$d_{\rm III} = \quad = (6\mathrm{i}\lambda)\mathrm{i}V(p^2)6\mathrm{i}\delta_\lambda^{(1)} \tag{A.37}$$



$$= (6\mathrm{i}\lambda)^3 3V(p^2)V(\mu^2) \tag{A.38}$$

$$= -648\mathrm{i}\lambda^3 V(p^2)V(\mu^2) \ , \tag{A.39}$$

in which we insert the expression (A.19) but this time expanded to second order in ϵ, since there are two factors of $1/\epsilon$ to render those finite and obtain

$$\begin{aligned} d_{\rm III} = (6{\rm i}\lambda)^3 \frac{3}{(4\pi)^4} \quad & \Bigg[\frac{1}{\epsilon^2} + \frac{1}{\epsilon}(2 - \gamma + \log 4\pi - \log p - \log \mu) \\ & + 3 - 2\gamma + \frac{\gamma^2}{2} - \frac{\pi^2}{24} + \log 16 + 2\log \pi - \gamma \log 4\pi + \frac{1}{2}(\log 4\pi)^2 \\ & + (-2 + \gamma - \log 4\pi)(\log p + \log \mu) \\ & + \frac{1}{2}[(\log p)^2 + (\log \mu)^2] + \log p \log \mu \Bigg] \end{aligned} \tag{A.40}$$

This diagram, indeed, also contains a nonlocal divergence from the logarithm of p^2 in the $\frac{1}{\epsilon}$-term, which just cancels with the corresponding term in (A.36). The remaining divergences can be included in the tree level counterterm.

Note that the diagrams of type $d_{\rm II}$ and $d_{\rm III}$ also need to be included "upside down", i.e. with initial and final momenta interchanged (group III of [128]). This means in particular

$$p \to -p \quad \text{and} \quad p_3 \to -p_3 \ . \tag{A.41}$$

We see from (A.29) that $d_{\rm II}$ remains invariant under this exchange if we also reparametrise $k \to -k$, which we can do since this just changes the arbitrary orientation of the internal momentum. Then

$$k^2 \quad \to \quad (-k)^2 = k^2, \tag{A.42}$$

$$(k+p)^2 \quad \to \quad (-k-p)^2 = (k+p)^2 \quad \text{and} \tag{A.43}$$

$$(k+p_3)^2 \quad \to \quad (-k-p_3)^2 = (k+p_3)^2 \ . \tag{A.44}$$

For $d_{\rm III}$ (A.37) we see that it only depends on p^2, anyway. Hence, the contributions (A.36) and (A.40) of these diagrams can just be doubled. Then we have accounted for all the s-channel diagrams on the two-loop level. In order to include also the t- and u-channels we have to take this contribution three times, so that schematically

$${\rm i}\mathcal{M}^{(2)} = {\rm i}\mathcal{M}^{(1)} + 3 \cdot (d_{\rm I} + 2d_{\rm II} + 2d_{\rm III}) + \delta_\lambda^{(2)} \ . \tag{A.45}$$

We first investigate how the interplay between the three types of diagrams removes all the divergences and the dependence of the regulator from the final result. For that, we need to split the contribution from the diagram d_{III} into the contribution from one channel, s, say, for the first order counterterm, and the two others. Then the two loop result for one channel is

$$\begin{aligned} (6{\rm i}\lambda)^{-3}(d_{\rm I} + 2d_{\rm II} + 2d_{\rm III}) &= ({\rm i}V(p^2))^2 + 2d_{\rm II} + 2\cdot 3V(p^2)V(\mu^2) \\ &= -V^2(p^2) + \underbrace{2V(p^2)V(\mu^2)}_{\frac{2}{3}d_{\rm III}} + \underbrace{4V(p^2)V(\mu^2)}_{\frac{4}{3}d_{\rm III}} + 2d_{\rm II} \\ &= -\underbrace{\left(V(p^2) - V(\mu^2)\right)^2}_{\text{finite}} + V^2(\mu^2) + 2d_{\rm II} + \frac{4}{3}d_{\rm III} \ . \end{aligned}$$

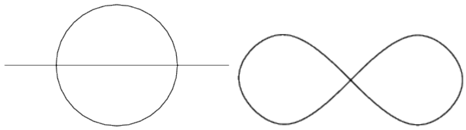


Figure A.2: *Two-loop prototype diagrams*

The combination $V(p^2) - V(\mu^2)$ has been made finite at the one loop level. In the following, we extract only the divergent terms

$$(d_{\rm I} + 2d_{\rm II} + 2d_{\rm III})_{\text{divergent}} = \frac{(6{\rm i}\lambda)^3}{(4\pi)^4}\left[\frac{3}{\epsilon^2} + \frac{9 - 3\gamma + (5 - \frac{1}{2}\gamma)\log[4\pi] - \log\left[256\pi^2\right] - 3\log\mu^2}{\epsilon}\right]$$

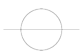
This is the momentum independent divergent part, which we include in the third order vertex counterterm $\delta_\lambda^{(2),\text{vertex}}$

Furthermore, we need to determine the counterterm for the remaining divergences. Since the renormalization conditions were already fulfilled at the one-loop level ${\rm i}\mathcal{M}^{(1)}$, the two-loop counterterm needs to cancel the two-loop contribution completely and hence

$$\delta_\lambda^{(2)} = -3\cdot\left(d_{\rm I} + 2d_{\rm II} + 2d_{\rm III} - \delta_\lambda^{(2),\text{vertex}}\right)\Big|_{p^2=\mu^2} \tag{A.46}$$

A.2.2 Two-loop vacuum diagram

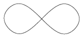
In order to do the resummation of the perturbation ordered by the number of loops á la Coleman-Weinberg [116], we need to evaluate the two-loop "prototype diagrams". Those are shown in fig. A.2. The first one is evaluated as follows:



$$\begin{aligned}
&= \frac{(6{\rm i}\lambda)^2}{3!}\int\frac{{\rm d}^d l}{(2\pi)^d}\frac{-{\rm i}}{l^2 + a\mu^2}{\rm i}V\left((p+l)^2\right)\\
&= \frac{(6{\rm i}\lambda)^2}{3!}\frac{-1}{2}\int\frac{{\rm d}^d l}{(2\pi)^d}\frac{1}{l^2 + a\mu^2}\int_0^1{\rm d}x\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}\frac{1}{(x(1-x)(p+l)^2 + a\mu^2)^{2-\frac{d}{2}}}\\
&= \frac{(6{\rm i}\lambda)^2}{3!}\frac{-1}{2}\frac{\Gamma(3-\frac{d}{2})}{(4\pi)^{\frac{d}{2}}}B\left(\frac{d}{2}-1,\frac{d}{2}-1\right)\int_0^1{\rm d}y\,y^{1-\frac{d}{2}}{\rm i}\int\frac{{\rm d}^d s_E}{(2\pi)^d}\frac{1}{[s_E^2 + p^2 y(1-y)]^{3-\frac{d}{2}}}\\
&= -{\rm i}\frac{(6{\rm i}\lambda)^2}{3!}\frac{3\pi\csc((4-\epsilon)\pi)}{(4\pi)^{4-\epsilon}}\frac{\Gamma(2-\frac{\epsilon}{2})\Gamma(1-\frac{\epsilon}{2})^2}{\Gamma(4-\frac{3\epsilon}{2})\Gamma(2-\epsilon)}\left(p^2\right)^{1-\epsilon}\\
&= {\rm i}\frac{(6{\rm i}\lambda)^2}{3!}\frac{p^2}{2(4\pi)^4}\left[\frac{1}{\epsilon} + \frac{13}{4} - \gamma + 2\log 2 + \log\pi - \log p^2\right]
\end{aligned} \tag{A.47}$$

where we have introduced a shifted momentum variable $s = l + yp$ and Wick rotated afterwards.

We now turn to the second diagram, which evaluates as follows.



$$= \frac{6\mathrm{i}\lambda}{2^2}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\int\frac{\mathrm{d}^d l}{(2\pi)^d}\frac{-\mathrm{i}}{k^2+a\mu^2}\frac{-\mathrm{i}}{l^2+a\mu^2} \tag{A.48}$$

$$= \frac{6\mathrm{i}\lambda}{2^2}\int\frac{\mathrm{d}^d k_E}{(2\pi)^d}\frac{1}{k_E^2-a\mu^2}\int\frac{\mathrm{d}^d l_E}{(2\pi)^d}\frac{1}{l_E^2-a\mu^2} \tag{A.49}$$

$$= \frac{6\mathrm{i}\lambda d^2}{16(4\pi)^d}\Gamma^2\left(-\frac{d}{2}\right)\left(\frac{1}{-a\mu^2}\right)^{2-d} \tag{A.50}$$

Appendix B

Diagrammatics of the double trace interaction

The theory we actually want to discuss is $\mathcal{N} = 4$ Super-Yang-Mills theory with gauge group $\mathrm{SU}(N)$, which has a double trace deformation [106]. In this section I explain the different diagrams having single trace and double trace interactions. The field ϕ is now an $N \times N$ matrix valued scalar and the two interaction terms in the Lagrangian look like

$$\mathcal{L}_{\text{int}}(N, \lambda) = \frac{g^2}{4} \mathrm{tr}\,(\phi^4) + \frac{f}{2} \left(\frac{1}{N} \mathrm{tr}\,\phi^2 \right)^2 \tag{B.1}$$

respectively. The different interactions correspond to different flow of the gauge group degrees of freedom. If the fields are written with their gauge group indices, e.g. ϕ^a_b, the traces can be represented as appropriate contraction of indices and the two different interactions just differ in some δ-functions. The single and double trace vertices then read

$$Tr\phi^4 = \phi_a{}^b \phi_c{}^d \phi_e{}^f \phi_g{}^h \ \delta^c_b \delta^e_d \delta^g_f \delta^a_h \quad \text{and} \quad (Tr\phi^2)^2 = \phi_a{}^b \phi_c{}^d \phi_e{}^f \phi_g{}^h \ \delta^c_b \delta^d_a \ \delta^g_f \delta^e_h \ , \tag{B.2}$$

respectively. The indices contracted run over gauge group degrees of freedom $a \ldots h = 1 \ldots N$ and do not affect the momentum flow in the diagrams calculated above.

For each interactions of the theory, there is a coupling, namely the single-trace g and the double-trace couplings f. Each get renormalized by an appropriate set of diagrams. If a diagram renormalizes the one or the other is determined by the structure of the gauge group flow. If it corresponds to the flow of a single trace vertex, the diagram renormalizes g, if it corresponds to a double trace structure, it renormalizes f. It is not important, which kind the internal vertices of the diagram are and a diagram can even contain vertices of different kinds.

The diagram including its gauge group structure is evaluated by taking the result from section A and replacing the dummy coupling λ by the appropriate form of the single and double trace

Figure B.1: *A straight and a twisted propagator in double line notation.*

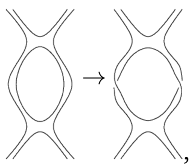
coupling, g^2 and $\frac{2f}{N^2}$, respectively. After the appropriate order of the 't Hooft couplings $g^2 N$ and f are split off, the order of N gets extracted as a pre-factor and the symmetry factor gets adjusted. Finally, we add up the diagrams order by order in N to get our final answer up to the desired order in N, i.e. the first sub-leading order.

B.1 Remark about twisting

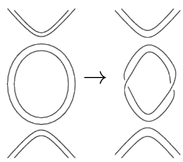
If we use double line notation, we notice that all the propagators can be "twisted", i.e. the line connects the upper vertex on the left hand side with the lower index on the right hand side and vice versa, as indicated in fig. B.1 Usually, this implies non-planarity and hence, these diagrams are suppressed in the large N limit. Since, here, we are taking into account next to leading order in $1/N$ diagrams, we must carefully examine, if diagrams with a twist do contribute in our approximation.

A single twist would not be compatible with the group structure, because arrows couldn't be placed in opposite directions on the same propagator any more. A Twist in two or more propagators can, depending on the diagram at hand and the specific propagator, that is twisted, or the combination thereof, have one or more of the following effects:

1. The trace structure is changed from single to double trace, e.g.



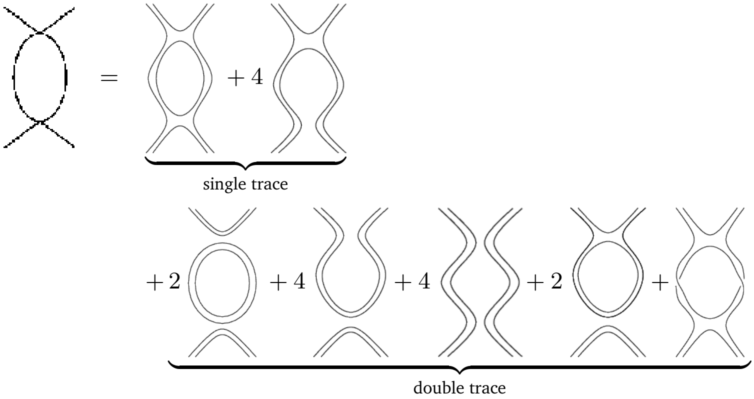
2. Nothing changes, e.g.



Because of the significance of the problem at hand, twists with the first effect are drawn as separate diagrams, to visualize them clearly. Diagrams of the second kind, that just double an untwisted diagram, are not drawn separately in order to maintain readability, but are taken into account by placing a combinatorial factor in front of the untwisted diagram which is multiplied.

B.2 One-loop diagrams with trace structure

At the one-loop level, we have an overall symmetry factor of $\frac{1}{2}$ due to interchange of the two propagators. Then, for one channel, we obtain the group structure



$$= (6\mathrm{i})^2 \mathrm{i} V(p^2) \frac{1}{2} \left[\underbrace{\frac{1}{2} g^4 N + 4 g^2 \left(\frac{2f}{N^2}\right)}_{\text{single trace}} + \underbrace{2\frac{1}{2}\left(\frac{2f}{N^2}\right)^2 N^2 + 4\left(\frac{2f}{N^2}\right)^2 + 4\frac{1}{2}\left(\frac{2f}{N^2}\right)^2 + 2g^2\left(\frac{2f}{N^2}\right) N + \frac{1}{2} g^4}_{\text{double trace}} \right], \tag{B.3}$$

where the first two terms renormalize the single-trace coupling. Using 't Hooft couplings, we obtain

$$= (6\mathrm{i})^2 \mathrm{i} V(p^2) \left\{ \underbrace{\frac{1}{4}(g^2 N)^2 \frac{1}{N}}_{\text{single trace}} + \underbrace{\left[2f(g^2 N) + 2f^2 + \frac{1}{4}(g^2 N)^2\right] \frac{1}{N^2}}_{\text{double trace}} + \underbrace{4f(g^2 N)\frac{1}{N^3}}_{\text{single trace}} + \underbrace{12 f^2 \frac{1}{N^4}}_{\text{double trace}} \right\}, \tag{B.4}$$

from which we keep diagrams up to $\mathcal{O}(N^{-2})$. This group structure contains the the momentum structure (A.19) and, hence, is inherited also by the one-loop counterterm, which we are going to determine subsequently.

To calculate the counterterm, we need to refine the renormalization conditions. For each of the two couplings we have a tree level diagram. Hence, the renormalization conditions now read

$$\mathrm{i}\mathcal{M}(p_1 p_2 \to p_3 p_4) = 6\mathrm{i}\left(g^2 + \frac{2f}{N^2}\right) \quad \text{at} \quad s = t = u = \mu^2 \,. \tag{B.5}$$

The one-loop counterterm now cancels the diagrams respecting their group structure in

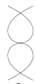
$$\mathrm{i}\mathcal{M}^{(1)}_{f,g} = 6\mathrm{i}\left(g^2 + \frac{2f}{N^2}\right) + (6\mathrm{i})^2[\text{grp. struct.(B.4)}]\, 3\mathrm{i}\cdot 2V(\mu^2) + 6\mathrm{i}\left(\delta^{(1)}_g + \delta^{(1)}_f\right) \,, \tag{B.6}$$

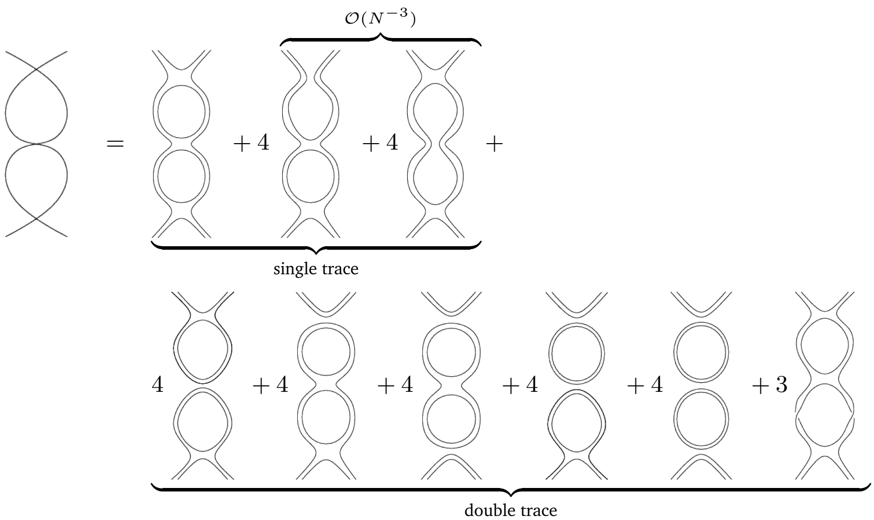
where we insert eqation (B.4). The second and third term need to cancel, but only single trace and double trace terms contribute to $\delta^{(1)}_g, \delta^{(1)}_f$, respectively. The factor 2 in front of the momentum factor removes the symmetry factor which is now absorbed into the group structure. We read off the counter terms up to second order in $1/N$

$$\delta^{(1)}_g = -(6\mathrm{i})^2(g^2N)^2\frac{1}{4N}V(\mu^2)\,, \tag{B.7}$$

$$\delta^{(1)}_f = -(6\mathrm{i})^2\left[2f(g^2N) + 2f^2 + \frac{1}{4}(g^2N)^2\right]\frac{1}{N^2}V(\mu^2)\,. \tag{B.8}$$

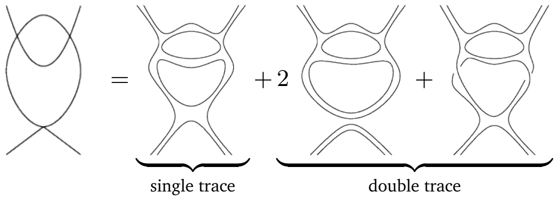
B.3 2-loop diagrams with double-trace structure

At the two-loop level, we have two diagrams. For  , suppressing all diagrams with $\mathcal{O}(N^{-3})$ in the 't Hooft limit, we obtain the group structure for one channel of this diagram



$$= (6\mathrm{i})^3[\mathrm{i}V(p^2)]^2\left[\underbrace{\frac{1}{8}(g^2N)^3\frac{1}{N}}_{\text{single trace}} + \underbrace{\left(3f(g^2N)^2 + 6f^2(g^2N) + 4f^3 + \frac{3}{4}(g^2N)^3\right)\frac{1}{N^2}}_{\text{double trace}}\right]\,. \tag{B.9}$$

For the other diagram , we only have one single and two double trace diagrams contributing up to order $\mathcal{O}(N^{-2})$



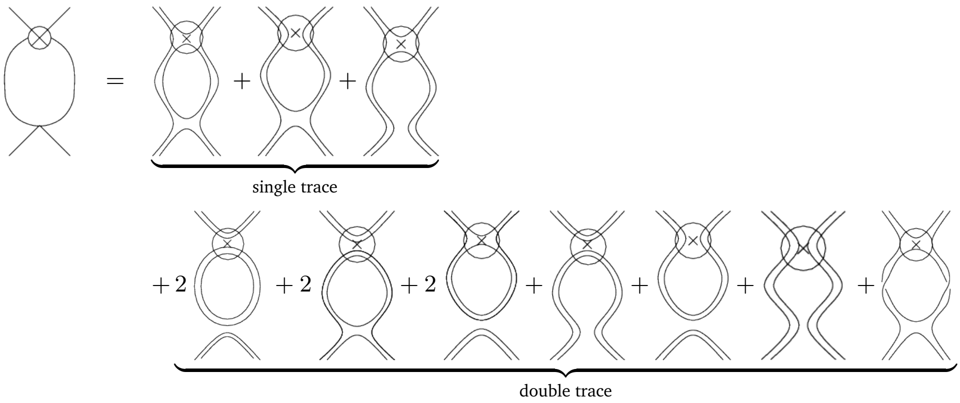
(B.10)

$$= (6\mathrm{i})^3 \left[\underbrace{(g^2N)^3 \frac{1}{N}}_{\text{single trace}} + \underbrace{\left(\frac{1}{2} f (g^2N)^2 + \frac{1}{8} (g^2N)^3 \right) \frac{1}{N^2}}_{\text{double trace}} \right] \cdot [\text{momentum}] \,. \tag{B.11}$$

The last type of diagram is a one-loop diagram with the second order counterterm replacing one vertex. Just as the first order counterterm, the second order counterterms inherits the above trace structure. The results (B.7) and (B.8) are now included into the calculation of the two-loop counterterm. Here, for the diagrams containing counterterms, its trace structure needs to fit the rest of the diagram, again. Such diagrams have an overall single or double trace structure coming from the one-loop diagrams, in which every one vertex is replaced by its corresponding counterterm, i.e. a single trace coupling with $\delta_g^{(1)}$ and a double trace coupling with $\delta_f^{(1)}$. The number of diagrams thus doubles but some of them are identical.[1] Since the index structure of the counterterm is the same as the one of the coupling, index loops and the

[1] Those diagrams count double, which is the same as removing one symmetry factor of $\frac{1}{2}$ because of the counterterm removing the symmetry between the two vertices. Such diagrams are only listed once.

order in N don't change. Hence, we find for a one-loop diagram with counterterm



$$
\begin{aligned}
&= (6\mathrm{i})^2 \mathrm{i}V(p^2)\Bigg[\underbrace{\frac{1}{2}(g^2N)\delta_g^{(1)} + \frac{1}{2}(g^2N)\frac{1}{N}\delta_f^{(1)} + f\frac{1}{N^2}\delta_g^{(1)}}_{\text{single trace}} \\
&\quad + \underbrace{2f\delta_f^{(1)} + (g^2N)\delta_f^{(1)} + 2f\frac{1}{N}\delta_g^{(1)} + 3f\frac{1}{N^2}\delta_f^{(1)} + \frac{1}{2}(g^2N)\frac{1}{N}\delta_g^{(1)}}_{\text{double trace}}\Bigg] \\
&= (6\mathrm{i})^2 \mathrm{i}V(p^2)\Bigg[\underbrace{\frac{1}{2}(g^2N)\delta_g^{(1)}}_{\text{single trace}} \\
&\quad + \underbrace{\big(2f + (g^2N)\big)\,\delta_f^{(1)} + \left(2f + \frac{1}{2}(g^2N)\right)\frac{\delta_g^{(1)}}{N}}_{\text{double trace}} + \mathcal{O}(N^{-3})\Bigg] \\
&= \underbrace{(6\mathrm{i})^3 3V(p^2)V(\mu^2)}_{d_{\mathrm{III}}\text{-momentum}} \cdot 2\Bigg[\underbrace{\frac{1}{8}(g^2N)^3\frac{1}{N}}_{\text{single trace}} \\
&\quad + \underbrace{\left(4f^3 + 6f^2(g^2N) + 3f(g^2N)^2 + \frac{3}{8}(g^2N)^3\right)\frac{1}{N^2}}_{\text{double trace}}\Bigg] \,. \qquad \text{(B.12)}
\end{aligned}
$$

In the second line, we have suppressed terms of higher order than $\mathcal{O}(N^{-2})$. After having filled in the expressions of the counterterms (B.7), (B.8), we see that diagrams contributing to the double trace structure are sub-leading in $1/N$ as expected earlier.

Finally, we have to determine the second order vertex counterterms $\delta_{g,f}^{(2)}$ with the respective trace structure. These need to cancel all the contributions of the two-loop diagrams which can not be absorbed into the one loop counterterms $\delta^{(1)}_{g,f}$. Considering only single and double trace diagrams, respectively, equation (A.46) looks

$$\delta^{(2)}_{g,f} = -3 \cdot \left(d^{g,f}_{\mathrm{I}} + 2d^{g,f}_{\mathrm{II}} + 2d^{g,f}_{\mathrm{III}} \right)\Big|_{p^2=\mu^2} . \tag{B.13}$$

We extract the corresponding terms from equations (B.9), (B.11) and (B.12) and sort them by their different couplings to obtain

$$\delta^{(2)}_{g} = -3(6\mathrm{i})^3 (g^2N)^3 \frac{1}{N} \left[8\frac{\text{eq. (A.28)}}{(6\mathrm{i}\lambda)^3} + 2\frac{\text{eq. (A.36)}}{(6\mathrm{i}\lambda)^3} + \frac{1}{2}\frac{\text{eq. (A.40)}}{(6\mathrm{i}\lambda)^3} \right] \tag{B.14}$$

and

$$\begin{aligned}
\delta^{(2)}_{f} \quad = \quad & -3(6\mathrm{i})^3 \frac{1}{N^2} \\
& \left[f(g^2N)^2 \left(3\frac{\text{eq. (A.28)}}{(6\mathrm{i}\lambda)^3} + 1\frac{\text{eq. (A.36)}}{(6\mathrm{i}\lambda)^3} + 12\frac{\text{eq. (A.40)}}{(6\mathrm{i}\lambda)^3} \right) \right. \\
& + f^2(g^2N) \left(6\frac{\text{eq. (A.28)}}{(6\mathrm{i}\lambda)^3} + 24\frac{\text{eq. (A.40)}}{(6\mathrm{i}\lambda)^3} \right) \\
& + f^3 \left(4\frac{\text{eq. (A.28)}}{(6\mathrm{i}\lambda)^3} + 16\frac{\text{eq. (A.40)}}{(6\mathrm{i}\lambda)^3} \right) \\
& \left. + (g^2N)^3 \left(\frac{3}{4}\frac{\text{eq. (A.28)}}{(6\mathrm{i}\lambda)^3} + \frac{1}{4}\frac{\text{eq. (A.36)}}{(6\mathrm{i}\lambda)^3} + \frac{3}{2}\frac{\text{eq. (A.40)}}{(6\mathrm{i}\lambda)^3} \right) \right] .
\end{aligned}$$

Now, we have all information at hand, which we need to extract any contribution to either the single- or double-trace β-function at any order in $1/N$ present up to two loops. I will explain the schematics of this process in the following section.

B.4 Extraction of the β-functions

B.4.1 Schematics

The β-function can be extracted from the Callan-Symanzik equation (2.157) applied to the 4-point Greens' function as noted earlier. As an ingredient thereof, we need to determine the anomalous dimension γ_ϕ of the scalar field, which is non-vanishing at the two-loop order. We obtain it from the Callan-Symanzik equation applied to the two-point Greens' function

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_g\frac{\partial}{\partial g} + \beta_f\frac{\partial}{\partial f} + 2\gamma_\phi \right) G^{(2)} = 0 , \tag{B.15}$$

which reduces to

$$\mu\frac{\partial}{\partial\mu} G^{(2),\text{2nd order}} + 2\gamma^{(2)}_\phi G^{(2),\text{0th order}} = 0 , \tag{B.16}$$

at second order, because the first contribution to the anomalous dimension comes at the second order and the first order contribution to the β-function vanishes. Here, $G^{(2),\text{0th order}} = \frac{-\mathrm{i}}{p^2}$

denotes the uncorrected propagator and $G^{(2),\text{2nd order}}$ contains the terms of second order in the couplings of the Green's function. From here, the anomalous dimension can readily be extracted to be

$$\gamma_\phi = -\frac{48f^2 + 48fg^2N + g^4N^4}{3072\,(N^2\pi^4\mu)} \ . \tag{B.17}$$

We only regard 4 point vertices at the moment. It is therefore convenient to abbreviate the single trace coupling with $e = g^2$, which we will do in the following. The functions $\beta_{e,f}$ are extracted from the single and double trace 4-point Green's functions, respectively,

$$G^{(4)}_{s/d}(p_1, p_2, p_3, p_4) = \mathrm{i}\mathcal{M}_{s,d}(p_1, p_2 \to p_3, p_4) \prod_{k=1}^{4} \frac{-\mathrm{i}}{p_k^2} \tag{B.18}$$

by evaluating the Callan-Symanzik equation (2.157) successively order by order. It is important to distinguish between the single and double trace Greens' functions, since this doubles the number of equations.

We identify the different terms by matching the coefficients of each term in the polynomial expansion of the Green's function. The first contribution to the β function comes in at the one-loop order, which is second order in the couplings for the four point Greens' function

$$\left(\mu\frac{\partial}{\partial\mu} + \beta_e\frac{\partial}{\partial e} + \beta_f\frac{\partial}{\partial f} + 4\gamma_\phi\right)\left(G^e_{4,s/d}e + G^f_{4,s/d}f + G^{e^2}_{4,s/d}e^2 + G^{ef}_{4,s/d}ef + G^{f^2}_{4,s/d}f^2\right) = 0 \ , \tag{B.19}$$

where we denote the single and double coefficient of the term x with $G^x_{4,s/d}$, respectively. Applying the operators to the Greens' function, we note, that only the counterterms introduce a dependence on μ to the Greens' function, such that the derivative with respect to μ kills the terms first order in the couplings. Hence, we find in terms of coefficients

$$\begin{aligned}
&\mu\frac{\partial}{\partial\mu}\left(G^{e^2}_{4,s/d}e^2 + G^{ef}_{4,s/d}ef + G^{f^2}_{4,s/d}e^2\right)\\
+\quad &\beta_g\left(G^e_{4,s/d} + 2G^{e^2}_{4,s/d}e + G^{ef}_{4,s/d}e\right)\\
+\quad &\beta_f\left(G^f_{4,s/d} + G^{ef}_{4,s/d}e + 2G^{f^2}_{4,s/d}f\right)\\
+\quad &4\gamma_\phi\left(G^e_{4,s/d}e + G^f_{4,s/d}f + G^{e^2}_{4,s/d}e^2 + G^{ef}_{4,s/d}ef + G^{f^2}_{4,s/d}f^2\right) = 0 \ ,
\end{aligned} \tag{B.20}$$

which needs to be satisfied term by term. The anomalous dimension doesn't have any linear contribution as seen in (B.17), so at the one loop order, it just drops out. The β-functions contain a term for each order in the coupling constants, which we denote with $\beta^x_{e/f}$ in the following. Since there is no linear contribution to the β function, those terms drop out. Furthermore, we observe that there are no tree level terms double and single trace terms with coupling e and f, respectively. Hence, the relations simplify considerably. We read off the relations for the individual coefficients after having restored the trace structure of the Greens' functions

$$\beta_e^{e^2} = -\tfrac{\mu}{G_s^e}\tfrac{\partial}{\partial\mu}G_s^{e^2}\ , \quad \beta_f^{e^2} = -\frac{\mu}{G_d^f}\frac{\partial}{\partial\mu}G_d^{e^2} \tag{B.21}$$

$$\beta_e^{f^2} = -\tfrac{\mu}{G_s^e}\tfrac{\partial}{\partial\mu}G_s^{f^2} = 0\ , \quad \beta_f^{f^2} = -\frac{\mu}{G_d^f}\frac{\partial}{\partial\mu}G_d^{f^2} \tag{B.22}$$

$$\beta_e^{ef} = -\tfrac{\mu}{G_s^e}\tfrac{\partial}{\partial\mu}G_s^{ef} = 0\ , \quad \beta_f^{ef} = -\frac{\mu}{G_d^f}\frac{\partial}{\partial\mu}G_d^{ef}\ , \tag{B.23}$$

where $\beta_e^{f^2}$ and β_e^{ef} vanish because single trace Greens' functions can only contain single trace vertices, at the one loop level.

At the two loop order, the Callan Symanzik equation reads

$$\begin{aligned}
&\mu\frac{\partial}{\partial\mu}\Big(G^e e + G^f f + G^{e^2}e^2 + G^{ef}ef + G^{f^2}f^2 \\
&\quad + G^{e^3}e^3 + G^{e^2 f}e^2 f + G^{ef^2}ef^2 + G^{f^3}f^3\Big) \\
+\ &\beta_e\Big(G^e + 2G^{e^2}e + G^{ef}f + 3G^{e^3}e^2 + 2G^{e^2f}ef + G^{ef^2}f^2\Big) \\
+\ &\beta_f\Big(G^f + G^{ef}e + 2G^{f^2}f + G^{e^2f}e^2 + 2G^{ef^2}ef\Big) \\
+\ &4\gamma_\phi\Big(G^e e + G^f f + G^{e^2}e^2 + G^{ef}ef + G^{f^2}f^2 \\
&\quad + G^{e^3}e^3 + G^{e^2f}e^2f + G^{ef^2}ef^2 + G^{f^3}f^3\Big) = 0\ ,
\end{aligned} \tag{B.24}$$

from which we extract, again, the relations third order in the couplings for the β functions. Their second order coefficients should be inserted from the one loop equations. Again, after restoring the trace structure of the Greens' functions, we can read off the eight third order coefficients. For that we note, that there are no single trace diagrams with only double trace couplings, so $G_s^{f^2} = 0$ and alike. We note that in our approximation, taking into account only terms up to third order in $\frac{1}{N}$, the single trace Greens' functions contain single trace couplings, only. Hence, the expressions for the third order coefficients of the β functions simplify to

$$\begin{aligned}
\beta_e^{e^3} &= -\frac{1}{G_s^e}\left(\mu\frac{\partial}{\partial\mu}G_s^{e^3} + 2\beta_e^{e^2}G_s^{e^2}\right) - 4\gamma^{e^2} \\
\beta_f^{e^3} &= -\frac{1}{G_d^f}\left(\mu\frac{\partial}{\partial\mu}G_d^{e^3} + 2\beta_e^{e^2}G_d^{e^2} + \beta_f^{e^2}G_d^{ef}\right) \\
\beta_e^{e^2f} &= -4\gamma^{ef} \\
\beta_f^{e^2f} &= -\frac{1}{G_d^f}\left(\mu\frac{\partial}{\partial\mu}G_d^{e^2f} + \beta_e^{e^2}G_d^{ef} + 2\beta_f^{e^2}G_d^{f^2} + \beta_f^{ef}G_d^{ef}\right) - 4\gamma^{e^2} \\
\beta_e^{ef^2} &= -4\gamma^{f^2} \\
\beta_f^{ef^2} &= -\frac{1}{G_d^f}\left(\mu\frac{\partial}{\partial\mu}G_d^{ef^2} + 2\beta_f^{ef}G_d^{f^2} + \beta_f^{f^2}G_d^{ef}\right) - 4\gamma^{ef} \\
\beta_e^{f^3} &= 0 \\
\beta_f^{f^3} &= -\frac{1}{G_d^f}\left(\mu\frac{\partial}{\partial\mu}G_d^{f^3} + 2\beta_f^{f^2}G_d^{f^2}\right) - 4\gamma^{f^2}\ .
\end{aligned}$$

The relations contained in this section represent a schematic way to extract all the information about the β-functions up to the two-loop order. Upon integration, the β-functions yield the renormalized couplings, which we use in section 2.5 to replace the RG-scale with the value of the scalar field by replacing the coupling in the effective potential.

APPENDIX C

REMARKS ON THE BACKGROUND FIELD METHOD

We first reproduce the one- and two-loop results of scalar ϕ^4 theory. In order to use the background field method, we decompose the scalar field

$$\Phi = \phi + \frac{R + \mathrm{i}C}{\sqrt{2}} , \tag{C.1}$$

where ϕ is the background field and R, C are the real and complex perturbations around it. We can choose $C = 0$ to examine the case of a real scalar field first.

Expanding the potential term of the Lagrangian

$$\mathcal{L}_{\text{pot.}} = \frac{1}{2} m_\Phi^2 \Phi^2 + \frac{\lambda}{4} \Phi^4 \tag{C.2}$$

we read of the effective mass of the real scalar R with the value ϕ of the background field

$$m_R^2 = m^2 - 9\lambda\phi^2 . \tag{C.3}$$

We find then for the effective one-loop potential

$$V^{(1)} = \frac{(m^2 - 9\lambda\phi^2)^2}{8} \left(-3 + 2 \ln \frac{m^2 - 9\lambda\phi^2}{Q^2} \right) . \tag{C.4}$$

C.1 THE IR CUTOFF

We observe that the effective potential calculated in the previous subsection acquires an imaginary part above a certain field value. In general, this corresponds to information loss and is not expected for an effective potential. However, since we are indeed expanding the potential around an unstable point, the apparent loss of information is to be expected (see [106, 129]). When expanded around the turnaround point, the reality of the effective potential should be recovered.

Technically, the imaginary part comes from integrating out also tachyonic modes, i.e. such modes for which the effective mass is negative. Since we are only interested in the UV behavior of the theory, we can use an IR cutoff such that only non-tachyonic modes are integrated over and a real effective potential is obtained. In our case, we see that for $\phi^2 \leq \frac{m^2}{9\lambda}$, the field becomes tachyonic. Hence, we apply an IR cutoff, which prevents the effective mass from becoming negative. We see, that this cutoff must depend on the field value and we can choose $\mu_{IR} = 3\lambda\phi^2 + \epsilon^2$.

To see how the IR cutoff takes effect in the background field method, we compare the treatment of the cutoff in appendix B of [106] (equation B.26 onwards) with C.4. We have

$$V^{(1)}_{\mathrm{CHT}} = \frac{1}{32\pi^2}\left(-\frac{9\lambda^2\phi^4}{4} + \frac{9\lambda^2\phi^4}{2}\ln\frac{-3\lambda\phi^2}{\Lambda^2} \underbrace{-3\lambda\Lambda^2\phi^2}_{\text{included in ct. in Martin}}\right) \tag{C.5}$$

$$V^{(1)}_{\mathrm{Martin}} = \frac{1}{32\pi^2}\left(-\frac{3m^4_{\text{eff.}}}{4} + \frac{m^4_{\text{eff.}}}{2}\ln\frac{m^2_{\text{eff.}}}{Q^2}\right) . \tag{C.6}$$

Comparing the two expressions, we see that the effective mass in [106] is $m^2_{\text{eff.}} = -3\lambda\phi^2$, which is consistent with the fact, that their bare mass is zero. We see a mismatch of a factor 3 in the term quartic in the effective mass, but this is merely a scheme dependence.

In the background field method, we see that we need to replace (C.4) with

$$-\frac{m^2_{\text{eff.}}}{4} - \frac{\mu^4_{IR}}{2}\ln\frac{\mu^2_{IR}+m^2_{\text{eff.}}}{\mu^2_{IR}} - \frac{m^2_{\text{eff.}}.\mu^2_{IR}}{2} + \frac{m^4_{\text{eff.}}}{2}\ln\frac{\mu^2_{IR}+m^2_{\text{eff.}}}{\Lambda^2} . \tag{C.7}$$

We can check that for $\mu_{IR} \to$ this reproduces (C.4).

Note that the IR cutoff is still ok due to perturbation theory, because for that we need

$$\mu^2_{\mathrm{IR}} \ll Q^2 \ll \Lambda^2 , \tag{C.8}$$

which with the above IR cutoff and after the replacement $Q \to \phi$ reads

$$\lambda\phi^2 \ll \phi^2 , \tag{C.9}$$

which is fulfilled automatically as long as $\lambda \ll 1$.

Chapter 3

Inflationary cosmology in supergravity

3.1 Problems of FRW cosmology and cosmic inflation as their solution

Whilst in chapter 2 we have dealt with the beginning of the universe, we are now turning our attention to its further development. Already in chapter 2, I have established that the universe is expanding as seen from the red-shifting of galaxies around us. Now we are going to look at which form this expansion has taken, which turns out to be not at all uniform. In fact, the latest physics Nobel prize at the time of writing was awarded for the precision observation at distant supernovae that the universe is currently expanding at an increasing rate [130–132], for which there has since also been further evidence [133, 134]. These observations mean that, today, there is a non-zero vacuum density or "dark energy". Here, we are interested in the form of the acceleration in the early universe, shortly after the big bang.

There are a number of theoretical and observational problems with the standard FRW cosmology as described by equation (2.1). An in-depth treatment of the following material can be found e.g. in [135–138]. It is already clear from the second Friedman equation (2.5) that the Hubble parameter is constant only for specific combinations of the cosmological constant Λ and the equation of state parameter w. In general, the rate of expansion will change, $\ddot{a} \neq 0$. We will now introduce models that interpret the cosmological constant as some vacuum energy, which changes in time. It is given by a potential, which has to obey certain restrictions.

The problems of FRW cosmology are

- the flatness problem,

- the horizon problem,
- the missing monopole problem
- and the explanation of the power spectrum of the cosmic background radiation (CMB).

In the following I will explain these problems before introducing *cosmic inflation* as a solution to them.

The flatness problem

One of the parameters of the first Friedman equation (2.4) is the curvature k of the universe. The spatial curvature of the universe is a quantity, measured to be extremely small. For the universe to be flat, its energy density must have a critical value for a given Hubble parameter. If we take the cosmological constant to be (approximately) zero, we can read off the *critical density* for a given Hubble parameter

$$\rho_c = \frac{3H^2}{8\pi G} \ . \tag{3.1}$$

We now introduce the ratio between the actual and the critical energy density

$$\Omega = \frac{\rho}{\rho_c} \ . \tag{3.2}$$

In terms of this, the Friedman equation can be rewritten as

$$\frac{1-\Omega}{\Omega}\rho a^2 = -\frac{3kc^2}{8\pi G} \ , \tag{3.3}$$

where the right hand side is constant. The scale factor a increases with the expansion, while the energy density ρ decreases. For matter and radiation dominated universes this decrease is quicker than the increase of the scale factor squared a^2 (cf. equation (2.9)). This means that the left hand side of equation (3.3) decreases rapidly. The order of magnitude for this decrease within one Planck time is 10^{60}. Figure 3.1 depicts how quickly the universe deviates from a flat initial configuration.

The CMB provides a wealth of information about the early universe. Its anisotropies can for example be used to measure the flatness of the universe. The typical angular distance between a cold and a hot spot, i.e. the first peak in the angular power spectrum, depends on the curvature of the universe. Besides this, comparing the distance of type 1a supernovae, which act as standard candles, to their redshift can be used to measure the expansion rate of the universe at different times. From those measurements combined, we arrive at the current value of Ω to be within 1% of unity [133, 134], or $|1-\Omega| \leq 0.01$. This implies that it was fine-tuned to 10^{-62} during the Planck era. This fine-tuning problem is called the *flatness problem*.

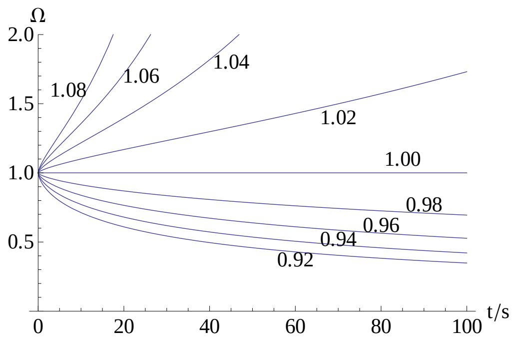


Figure 3.1: *A flat universe is not stable. Small perturbations quickly drive the universe away from the critical mass density, unless the initial configuration is fine tuned to be precisely unity.*

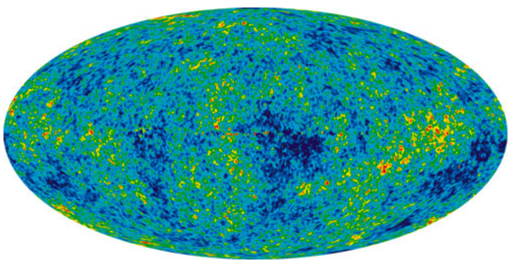


Figure 3.2: *The seven year WMAP scan of the CMB. The relative size of anisotropies is only* 10^{-5}. *Figure courtesy of the WMAP science team from the LAMBDA archive.*

Horizon problem

Although the CMB has much sought-after anisotropies, it is remarkably homogeneous. In fact, the temperature fluctuations have a relative size of only $\frac{\Delta T}{T} \approx 10^{-5}$, and the average temperature of the CMB is uniform over the whole sky. A plasma would only be so homogeneous in a region, which has been in causal contact for a long enough time to equilibrate. The CMB is conjectured to have formed 360.000 years after the big bang, when electrons and ions in the early plasma combined to form neutral hydrogen and photons were no longer scattered such that they decoupled from matter. Therefore, the Hubble horizon at this time is the maximum distance at which two points in the sky could still have been in causal contact. Assuming a standard expansion, this patch would have blown up to what now appears under one degree at the sky (see figure 3.3). A fluid dynamical equilibration process therefore cannot account for the isotropy of the CMB, which must have been pre-imposed by another mechanism. Unless there is a natural explanation of that, this is a fine-tuning problem known as the *horizon problem*.

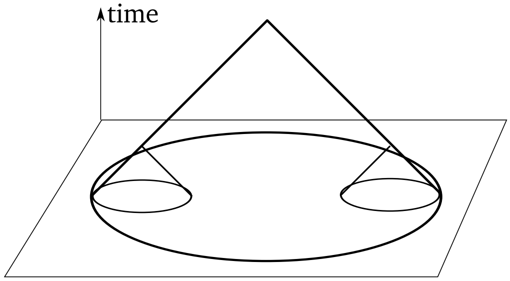


Figure 3.3: *The graphic shows that what appears in the current backward light-cone could not have been in causal contact in the past in a cosmological model with standard expansion.*

MISSING MONOPOLE PROBLEM

In the very early universe at very high temperatures, the electro-magnetic, weak and strong nuclear forces are supposed to be described by one unified gauge theory. Those grand unified theories (GUTs) [139, 140] are a generalization of the Glashow-Salam-Weinberg (GSW) model for the electro-weak unification [141–143]. In such models, the different forces become unified above a critical temperature, which is about $3 \cdot 10^{15} K$ for the GSW-model and about $10^{27} K$ for GUTs. While cooling down with the expansion, the temperature drops below this critical value and the unified symmetry gets broken. This phase transition leads to the production of topological defects such as magnetic monopoles and domain walls. Since one domain must have been around the size of the Hubble radius at the time of the freeze-out, the number of defects can be estimated [144–146]. The fact that we have not yet been able to observe them is referred to as the *missing monopole problem*.

INFLATION AS A SOLUTION

It seems as if all those problems could be solved simultaneously if there was a mechanism which made the universe expand at a very large rate very early on, such that a causally connected region would be stretched out to match the size of the nowadays observable universe, space-time would be flattened out and the monopole remnants would be diluted. Such a mechanism was proposed by Alan Guth in 1980 under the name of *inflation* [147] and later improved by Andrei Linde [148], Andreas Albrecht and Paul Steinhard [149]. It should be noted that since the above problems concern the fine tuning of initial conditions, any mechanism solving them is only a valid improvement if it is somehow more natural and requires less tuning. One of the advantages of the inflationary paradigm is that it erases the dependence on the initial state.

The way inflation can solve the cosmological conundrums is that a contribution to the energy density with an equation of state parameter $w \leq -\frac{1}{3}$, such as a cosmological constant, makes the causal radius grow faster then the Hubble radius. The region which is causally connected to a point can be determined from the *comoving particle horizon*, which is the distance a null

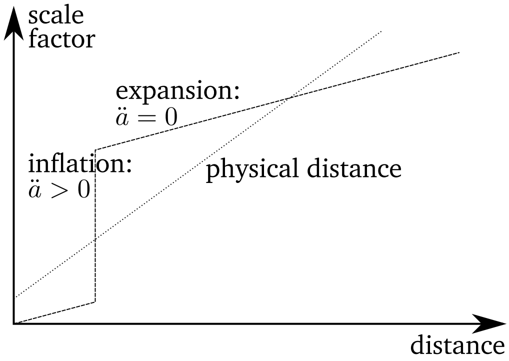


Figure 3.4: *During the inflationary phase, the Hubble horizon grows much faster than the physical distance between two points. After the end of inflation, the scales re-enter the horizon. Inflation magnifies a formerly causally connected patch to fill the visible universe.*

particle can have travelled since the big bang with $a(0) = 0$. It is equal to the comoving time

$$\tau = \int_0^t \frac{\mathrm{d}t'}{a(t')} = \int_0^a \frac{\mathrm{d}a'}{a'} \frac{1}{a' H(a')} \,. \tag{3.4}$$

For a spacially flat universe, the scale factor behaves as

$$a(t) = a_0 t^{\frac{2}{3(w+1)}} \,. \tag{3.5}$$

With the definition of the Hubble parameter $H = \dot{a}/a$, we see that the comoving time depends on the equation of state parameter as

$$\tau \sim a^{\frac{1}{2}(1+3w)} \,. \tag{3.6}$$

This means, that the horizon indeed shrinks for $w < -\frac{1}{3}$, and the horizon problem is solved (cf. fig. 3.4). Such a configuration will lead to an accelerated expansion

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{1}{aH} < 0 \quad \Rightarrow \quad \ddot{a} > 0 \,. \tag{3.7}$$

Hence the name inflation.

The easiest incarnation of inflation is *single-field slow-roll inflation*, in which inflation is driven by a single scalar field ϕ, the so-called *inflaton*. This is added to the gravitational action

$$\mathcal{S}_{\text{inflation}} = \int \mathrm{d}^4 x \sqrt{g} \left(\frac{M_{\text{pl}}^2}{2} R + \frac{1}{2} (\partial \phi)^2 - V(\phi) \right) \,. \tag{3.8}$$

The equations of motion and the Friedman equations derived from that action are

$$0 = \ddot{\phi} + 3H\dot{\phi} + V'(\phi) \,, \tag{3.9}$$

$$H^2 = \frac{1}{3} \left(\frac{1}{2} \dot{\phi}^2 + V(\phi) \right) \,, \tag{3.10}$$

$$\frac{\ddot{a}}{a} = (\rho + 3p) \,, \tag{3.11}$$

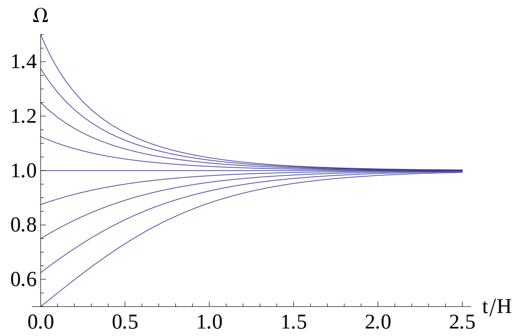


Figure 3.5: *The energy density of a universe which is dominated by a cosmological constant or some potential energy is driven towards the critical value. Therefore, inflation erases the initial condition and yields a flat universe.*

where

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) \ , \tag{3.12}$$

$$p = \frac{1}{2}\dot{\phi}^2 - V(\phi) \ . \tag{3.13}$$

We see that to get $w < -\frac{1}{3}$, the system needs to be dominated by potential energy. For the "no-roll"-case $\dot{\phi}^2 = 0$, we recover a cosmological constant.

To be a solution, inflation needs to last long enough to extend space-time by a sufficient amount. To test this for a given model the two slow-roll parameters can be used. They are defined to be

$$\epsilon = \frac{M_{\rm pl}^2}{2}\left(\frac{V'}{V}\right)^2 \ , \tag{3.14}$$

$$\eta = M_{\rm pl}^2 \frac{V''}{V} \tag{3.15}$$

and are required to be small $\epsilon \ll 1, \eta \ll 1$ to have a good model of slow-roll inflation. In that limit, the Friedman equations can be easily solved to give

$$H^2 = \frac{1}{3}V(\phi) \sim \text{constant} \ , \tag{3.16}$$

$$\dot{\phi} = -\frac{V'}{3H} \ , \tag{3.17}$$

$$a(t) \sim \mathrm{e}^{Ht} \tag{3.18}$$

and the expansion is exponential, indeed.

An accelerated expansion also solves the flatness problem. Since the scale factor is exponential, it now wins agains the scaling of the matter density, and the critical density naturally becomes an attractor for a variety of initial values of Ω (see figure 3.5). Potentially present monopoles

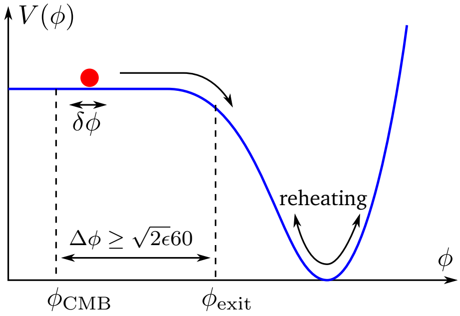


Figure 3.6: *A possible form of an inflaton potential. In the flat piece, the universe will inflate. The CMB is formed at ϕ_{CMB}. Its anisotropies reflect the fluctuations $\delta\phi$ of the inflaton. After inflation has come to an end at ϕ_{end} the energy of the inflaton field must be transferred to the standard model particles, which is called reheating.*

and topological defects are diluted by the vast expansion such that they are hardly visible at present.

The shape of a possible slow-roll inflationary potential is depicted in figure 3.6. Inflation will end if $\epsilon \sim 1$. It is customary to express the amount of inflation in so-called *e-foldings*, which is the powers of e with which the scale factor a has grown. It can be expressed in terms of the first slow-roll parameter as

$$N \equiv \ln \frac{a_{\text{final}}}{a_{\text{initial}}} = \int_{t_i}^{t_f} H \mathrm{d}t = \int_{\phi_i}^{\phi_f} = \int_{\phi_i}^{\phi_f} \frac{H}{\dot{\phi}} \mathrm{d}\phi \approx \int_{\phi_i}^{\phi_f} \frac{V'}{V} \mathrm{d}\phi = \int_{\phi_i}^{\phi_f} \frac{\mathrm{d}\phi}{\sqrt{2\epsilon}} \,, \tag{3.19}$$

where we have employed the slow-roll approximation. To solve the cosmological problems, ϵ needs to be such that inflation lasts for at least $N > 60$ e-foldings.

I do not want to leave unmentioned that there are alternative explanations for the conundrums of the early universe. Amongst such models are string gas cosmology [150, 151], the ekpyrotic or cyclic universe [152–154] or, taking the holographic lessons from chapter 1 more seriously, a holographic model [155, 156].

3.2 Inflation in string theory and supergravity

So far, we have got to know inflation as a merely phenomenological model which can reduce the fine-tuning problems in the early universe. Nothing has been said so far about what the inflaton should be and where its potential would come from. These ingredients, if at all real, must finally come from a fundamental, UV-complete theory of gravity.

One of the reasons why we would like to see inflation be backed up by a UV-safe theory is the objective of this chapter, namely the *η-problem*. The second slow-roll parameter, which is basically the inflaton mass measured in Hubble units

$$\eta = M_{\rm pl}^2 \frac{V''}{V} \approx \frac{m_\phi^2}{3H^2} \tag{3.20}$$

needs to be much smaller than unity to ensure successful inflation. However, when going to higher energies, this parameter will receive UV-corrections. In fact, integrating out Planck scale degrees of freedom will add a generic dimension six operator as a correction to the Lagrangian

$$\frac{\mathcal{O}_6}{M_{\rm pl}^2} = \frac{\mathcal{O}_4}{M_{\rm pl}^2}\phi^2 \ . \tag{3.21}$$

If the dimension four operator has a vacuum expectation value comparable to the scale of inflation

$$\langle \mathcal{O}_4 \rangle \sim V \ , \tag{3.22}$$

this term corrects the inflaton mass by an order $\frac{\sqrt{V}}{M_{\rm pl}} \sim H$, which would correspond to correcting η by order unity. This reflects the fact that generically, in an effective theory the mass would run all the way to the cut-off. Therefore, to ensure that inflation remains valid, we need to have good control over the UV physics.

An effective way of removing an irrelevant operator is by employing some kind of symmetry which forbids it. In the case at hand, this could be a shift symmetry. If the inflaton was taken to be e.g. the phase φ of a complex field ϕ with a U(1) symmetry $\phi \to e^{i\alpha}\phi$, shifting the inflaton $\varphi \to \varphi + \alpha$ doesn't change the theory. A flat potential breaks this symmetry only weakly during the inflationary era and the symmetry could remove the dimension six coupling. Yet, quite general arguments seem to suggest that global continuous symmetries are not allowed in a generic theory of quantum gravity [157]. Therefore a shift symmetry is not a natural thing to assume, and how to deal with the η-problem depends crucially on how the details of the fundamental theory are reflected in the low energy action. On top, the natural scale of inflation is some 10^{14} GeV such that we expect new physics to enter the picture.

For a long time, string theory has been building up hopes of providing such a UV complete understanding of gravity. Examining the observational and theoretical constraints of inflationary models built from string theory is therefore of extreme interest. In addition, string theory comes with a lot of new degrees of freedom, like moduli fields, branes, extra dimensions and warp factors, that can be used as an inflaton and for building inflationary models. It is widely known that a consistent formulation of string theory requires ten space-time dimensions, whereas our observations determine the number of extended dimensions to be only four. This means that the extra dimensions would be accessible only at high energies but need to be compactified at low energies. Then, if the energy scale is low as compared to the string tension, four-dimensional supergravity is the appropriate effective theory to describe cosmology because only string zero modes enter the description and the effects of the extra dimensions can be integrated out. For an introduction to string theory, I refer the reader to one of the many textbooks [158–161].

When compactifying the extra dimensions one inevitably tampers with the symmetries of the theory. The more internal symmetries the compactification manifold has, the more symmetries are preserved by the compactification. For instance, compactifying supergravity on a torus leads to maximal supergravity, whereas compactifying on a Calabi-Yau manifold will lead to an $\mathcal{N} = 2$ theory. Adding even more ingredients like fluxes, branes or orientifolds, the effective theory can be down to $\mathcal{N} = 1$ [162–167]. Calabi-Yau manifolds are widely used to construct low dimensional models. They come with a lot of symmetries [168] which manifest themselves as massless scalar degrees of freedom, so-called *moduli* in the effective theory. The proliferation of extra degrees of freedom is a vice as much as a virtue of the theory. Whereas some of them come in handy as inflaton fields, most of these massless fields are not observed and hence must be made heavy by creating some potential. This procedure is known as moduli stabilization. Another problem when looking at cosmology is the fact that in de Sitter space, which describes our universe, supersymmetry is broken. This makes it much more difficult to find string theory solutions. A first attempt to solve both problems was done in the famous KKLT paper [169]. It builds on the insight that three-form fluxes can stabilize the complex structure modulus and the dilaton [170–173]. On top, when including non-perturbative corrections, the volume modulus is also stabilized, and adding the potential of a small number of anti-D3 branes lifts the vacuum to being de Sitter. Large volume scenarios provide another way of stabilizing the moduli [174, 175].

The KKLT procedure has been extended also to allow for inflation with the potential of a $D3-\overline{D3}$-brane pair [176] and a large class of models inspired thereby. These models only allow for a small scale for inflation as determined by the Lyth bound, which constrains the variation of a field to be sub-Planckian [177]. A higher scale for inflation is allowed by axion monodromy inflation [178–180]. With a large supply of degrees of freedom, it is also possible to drive inflation with multiple fields as e.g. in [181–184]. Multiple fields will generically interact with each other, which will lead to observable signatures in the CMB, the most important of which are non-Gaussianities and isocurvature modes [185, 186]. A comprehensive re- and overview of string theoretic inflationary models is given in [187, 188] and references therein.

For the purpose of this chapter, the important common property of all such models is the following. To construct an inflationary potential, some of the degrees of freedom are chosen, while the others are supposed to remain silent at the minimum of their stabilizing potential. This is not only assumed at some point in field space but along the whole inflationary trajectory, such that the physics is assumed to be only described by the fields that have been picked to play a rôle in inflation. I group the fields participating in inflationary dynamics to form the inflationary sector. From the pint of view of inflation, the dynamics of the other fields is not visible and thus I call them the hidden sector. In a realistic model, they comprise in particular the standard model fields. They become dynamical only during reheating, when the inflation's energy is transferred to them (cf. figure 3.7). This truncation of the theory to a sub-sector is justified by the observation that the coupling is only of gravitational strength. Revisiting the η-problem for a simple setup with multiple fields in supergravity is one example study, which shows that the gravitational strength coupling must not be underestimated.

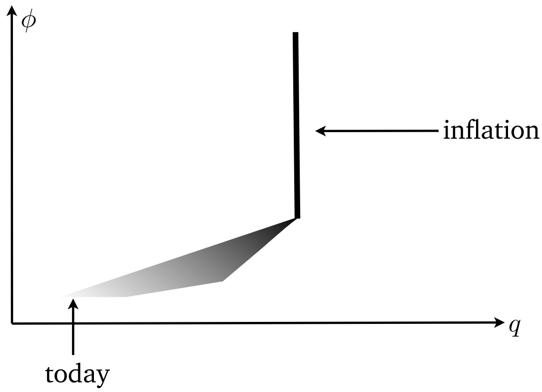


Figure 3.7: *A schematic picture of the evolution of a two-sector cosmology during and after inflation in configuration space.*

3.3 Inflationary theories with multiple sectors

The construction of realistic models of slow-roll inflation in supergravity is a longstanding puzzle. Supersymmetry can alleviate the fine-tuning necessary to obtain slow-roll inflation — if one assumes that the inflaton is a modulus of the supersymmetric ground state — but cannot solve it completely. This is most clearly seen in the supergravity η-problem: if the inflaton is a lifted modulus, then its mass in the inflationary background is proportional to the supersymmetry breaking scale. Therefore, the slow-roll parameter $\eta \simeq V''/V$ generically equals unity rather than a small number [189].

We will show here, however, that the η-problem is more serious than a simple hierarchy problem. In the conventional mode of study, the inflaton sector is always a sub-sector of the full supergravity theory presumed to describe our Universe. When the inflationary sub-sector of the supergravity is studied *an sich*, tuning a few parameters of the Lagrangian to order 10^{-2} will generically solve the problem. We will clarify that this split of the supergravity sector into an inflationary sector and other hidden sectors implicitly makes the assumption that all the other sectors are in a 'supersymmetric' ground state: i.e. if the inflaton sector which must break supersymmetry is decoupled, the ground state of the remaining sectors is supersymmetric. If this is not the case, the effect on the η-parameter or on the inflationary dynamics in general can be large, even if the sypersymmetry breaking scale in the hidden sector is small. Blind truncation in supergravities to the inflaton sector alone, if one does not know whether other sectors preserve supersymmetry, is therefore an inconsistent approach towards slow-roll supergravity inflation. Coupling the truncated sector back in completely spoils the naïve solution found. This result, together with recent qualitatively similar findings for sequestered supergravities (where only the potential has a two-sector structure) [190], provides strong evidence that to find true slow-roll inflation in supergravity one needs to know the global ground state of the system. The one obvious class of models where sector-mixing is not yet considered is the newly discovered manifest embedding of single field inflationary models in supergravity [191]. If these models are also sensitive to hidden sectors, it would arguably certify the necessity of a global analysis for cosmological solutions in supergravity and string theory.

We will obtain our results on two-sector supergravities by an explicit calculation. The gravitational coupling between the hidden and the inflaton sectors is universal, which can be described by a simple F-term scalar supergravity theory. As in most discussions on inflationary supergravity theories, we will ignore D-terms as one expects its VEV to be zero throughout the early Universe [192]. Including D-terms (which themselves always need to be accompanied by F-terms) only complicates the F-term analysis, which is where the η-problem resides. Furthermore, although true inflationary dynamics ought to be described in a fully kinetic description [193], we can already make our point by simply considering the mass eigenmodes of the system. In a strict slow-roll and slow-turn approximation the mass eigenmodes of the system determine the dynamics of the full system.

Specifically we shall show the following for two-sector supergravities where the sectors are distinguished by independent R-symmetry invariant Kähler functions:

- Given a naïve supergravity solution to the η-problem, this solution is only consistent if the other sector is in its supersymmetric ground state.
- If it is not in its ground state, then the scalar fields of that sector cannot be static but *must* evolve cosmologically as well.
- In order for the naïve solution to still control the cosmological evolution these fields must move very slowly. This translates in the requirement that the contribution to the first slow-roll parameter of the hidden sector must be much smaller than the contribution from the naïve inflaton sector, $\epsilon_{\text{hidden}} \ll \epsilon_{\text{naïve}}$.
- There are two ways to ensure that ϵ_{hidden} is small: Either the supersymmetry breaking scale in the hidden sector is very small or a particular linear combination of first and second derivatives of the generalized Kähler function is small.
  - In the latter case, one finds that the second slow-roll parameter $\eta_{\text{naïve}}$ receives a very large correction $\eta_{\text{true}} - \eta_{\text{naïve}} \gg \eta_{\text{naïve}}$, unless the supersymmetry breaking scale in the hidden sector is small. This returns us to the first case.
  - In the first case, one finds that the hidden sector always contains a light mode, because in a supersymmetry breaking (almost) stabilized supergravity sector there is always a mode that scales with the scale of supersymmetry breaking. This light mode will overrule the naïve single field inflationary dynamics.

Thus for *any* nonzero supersymmetry breaking scale in the hidden sector — even when this scale is very small — the true mass eigenmodes of the system are linear combinations of the hidden sector fields and the inflaton sector fields. We compute these eigenmodes. By assumption, the true value of the slow-roll parameter η is the smallest of these eigenmodes. Depending on the values of the supersymmetry breaking scale and the naïve lowest mass eigenstate in the hidden sector, we find that

1. The new set of mass eigenmodes can have closely spaced eigenvalues, and thus the initial assumption of single field inflation is incorrect. Then a full multi-field re-analysis is required.

2. The relative change of the value of η from the naïve to the true solution can be quantified and shows that for a supersymmetry breaking hidden sector, the naïve model is only reliable if the naïve lowest mass eigenstate in the hidden sector is much larger than the square of the scale of hidden sector supersymmetry breaking divided by the inflaton mass. This effectively excludes all models where the hidden sector has (nearly) massless modes.

3. The smallest eigenmode can be dominantly determined by the hidden sector, and thus the initial assumption that the cosmological dynamics is constrained to the inflaton sector is incorrect. Again a full multi-field re-analysis is required.

One concludes that in general one needs to know/assume the ground states and the lowest mass eigenstates of *all* the hidden sectors to reliably find a slow-roll inflationary supergravity.

The structure of the rest of this chapter is the following. Section 3.4 reviews some definitions in supergravity and explains how sectors are coupled in supergravity. This leads directly to the first result that in a stabilized supergravity sector there always is a mode that scales with the scale of supersymmetry breaking. In section 3.5 we discuss the η-problem in a single sector theory and then consider the effect of a hidden sector qualitatively and quantitatively. The quantitative result is analyzed in section 3.6 both in terms of effective parameters and direct supergravity parameters. As a notable example of our result, we show that if the hidden sector is the Standard Model, where its supersymmetry breaking is not caused by the inflaton sector but otherwise, spoils the naïve slow-roll solution in the putative inflaton sector. The chapter is supplemented with two appendices in which some of the longer formulae are given.

3.4 A STABILIZED SECTOR IN A SUPERGRAVITY TWO-SECTOR SYSTEM

We shall start by recalling how two sectors are gravitationally coupled in supergravity. Although this coupling is universal, the definition differs from regular gravity in an important way: the superpotentials multiply rather than add.

We will then consider one of the two sectors to be a stable hidden sector. We show that a light mode develops, which indicates that the hidden sector obtains a flat direction and is not stable any more. This extends the result of [194], in which it is shown that non-supersymmetric

Minkowski minima always develop at least one light mass mode, to de Sitter and Anti-de Sitter vacua.

3.4.1 The supergravity action

The action for the scalar sector of $\mathcal{N}=1$ supergravity is

$$S = M_{\rm pl}^2 \int \mathrm{d}^4x\sqrt{g}\left[\frac{1}{2}R - g^{\mu\nu}G_{\alpha\overline{\beta}}\partial_\mu X^\alpha \partial_\nu \overline{X}^{\overline{\beta}} - VM_{\rm pl}^2\right] , \tag{3.23}$$

in which $G_{\alpha\overline{\beta}}$ is the field space metric and $g_{\mu\nu}$ is the spacetime metric with associated Riemann scalar R. The Greek indices run over all fields $\{\alpha, \overline{\beta}\}$ or over spacetime coordinates $\{\mu, \nu\}$. For calculational convenience we have defined the scalar fields X and functions V, K and W to be dimensionless. The (F-term) potential V of the scalar sector is defined as

$$V = e^G\left(G_\alpha G^\alpha - 3\right) . \tag{3.24}$$

Through $G_\alpha = \partial_\alpha G$, $G_{\alpha\overline{\beta}} = \partial_\alpha\partial_{\overline{\beta}}G$, the action (3.23) is completely specified by the real Kähler function $G(X, \overline{X})$, which is related to global supersymmetry quantities through

$$G(X, \overline{X}) = K(X, \overline{X}) + \log\left(W(X)\right) + \log\left(\overline{W}(\overline{X})\right) \tag{3.25}$$

in terms of the real Kähler potential $K(X, \overline{X})$ and the holomorphic (dimensionless) superpotential $W(X)$.-[1] The definition for G is convenient as it is invariant under Kähler transformations, i.e. it is invariant under the simultaneous transformation of $K(X, \overline{X}) \to K(X, \overline{X}) + f(X) + \overline{f}(\overline{X})$ and $W(X) \to e^{-f(X)}W(X)$ for an arbitrary holomorphic function $f(X)$.

3.4.2 Canonical coupling

To describe a two-sector system we consider a class of minimally coupled scenarios [195–197]

$$G(\phi, \overline{\phi}, q, \overline{q}) = G^{(1)}(\phi, \overline{\phi}) + G^{(2)}(q, \overline{q}) , \tag{3.26}$$

with ϕ, q denoting the fields in the two sectors respectively. In the following, we will take the indices $\{i, \overline{\jmath}\}$ to run over the ϕ-fields, while $\{a, \overline{b}\}$ denote the fields in the q-sector. Later in this chapter we will take the ϕ-fields to drive inflation, while the q-fields reside in another sector which is naïvely assumed not to take part in the inflationary dynamics and is hence called the hidden sector. This split of the Kähler function $G(\phi, \overline{\phi}, q, \overline{q})$ (3.26) is invariant under Kähler transformations in each sector separately [198–202] and thus defines a sensible way of splitting up the action in multiple sectors. Amongst other properties, this split guarantees that a BPS solution in one particular sector is a BPS solution of the full theory. In terms of K and W,

[1] Note that this definition requires $W \neq 0$. For $W = 0$ a Kähler function G cannot be defined. In this paper we will assume that $W \neq 0$.

this definition has a conventional separation of the Kähler potential, but the superpotentials in each sector combine multiplicatively rather than add

$$K(\phi,\overline{\phi},q,\overline{q})+\log|W(\phi,q)|^2=K^{(1)}(\phi,\overline{\phi})+K^{(2)}(q,\overline{q})+\log|W^{(1)}(\phi)W^{(2)}(q)|^2\,. \tag{3.27}$$

Let us illustrate the importance of this multiplicative superpotential in the situation in which the hidden sector resides in a supersymmetric vacuum, i.e. $\partial_a V(q_0)=0$ and $\partial_a G^{(2)}(q_0)=0$. We write the superpotential of the hidden sector as $W^{(2)}(q)=W_0^{(2)}+W_{\text{global}}^{(2)}(q-q_0)$. The second term in this expression is what determines the potential for fluctuations around the minimum of the hidden sector, while the first constant term is just an overall contribution and hence not interesting for the internal hidden sector dynamics at energies much less than the Planck scale. However, for the gravitational dynamics and the remaining ϕ-sector this 'vacuum energy contribution' $W_0^{(2)}$ is of crucial importance as it sets the scale of the potential

$$V=e^{K^{(2)}}|W_0^{(2)}|^2e^{G^{(1)}}\left(G_i^{(1)}G^{(1)i}-3\right)\,, \tag{3.28}$$

which is evaluated at $q=q_0$ such that all terms depending on $W_{\text{global}}^{(2)}$ vanish. The normal practice of setting $W_0^{(2)}$ to zero as an overall contribution to the hidden sector is neglecting the fact that gravity also feels the constant part of the potential energy, as opposed to field theory. The inflationary sector feels the presence of the hidden sector through this coupling and as such it may be more intuitive to regard $W_0^{(2)}$ to contain information about the inflationary sector rather than the hidden sector. Making a similar split in $W^{(1)}$, the constant part $W_0^{(1)}$ is the overall contribution to the hidden sector due to the inflaton sector.

The multiplicative superpotential also means that the zero-gravity limit to a global supersymmetry is more subtle than just taking $M_{\text{pl}}\to\infty$, as is usually done [2]. One must first determine a ground state which sets $W_0^{(1)}$ and $W_0^{(2)}$, and then send both $W_0^{(1)}\to 0$ and $W_0^{(2)}\to 0$ in such a way that the combinations $W_0^{(1)}M_{\text{pl}}$ and $W_0^{(2)}M_{\text{pl}}$ remain constant. Instating the canonical dimensions for the fields and the Kähler potential and rescaling the couplings such that $W_{\text{eff}}^{(2)}=W_0^{(1)}W_{\text{global}}^{(2)}$ and $W_{\text{eff}}^{(1)}=W_0^{(2)}W_{\text{global}}^{(1)}$ scale as M_{pl}^{-3}, the total superpotential

$$W=W_0^{(1)}W_0^{(2)}+W_0^{(1)}W_{\text{global}}^{(2)}+W_0^{(2)}W_{\text{global}}^{(1)}+W_{\text{global}}^{(1)}W_{\text{global}}^{(2)}\,, \tag{3.29}$$

then consists of a constant term which scales as $W_0^{(1)}W_0^{(2)}\sim M_{\text{pl}}^{-2}$, cross-terms which scale as $W_0^{(1)}W_{\text{global}}^{(2)}+W_0^{(2)}W_{\text{global}}^{(1)}\sim M_{\text{pl}}^{-3}$ and a multiplicative term which scales as $W_{\text{global}}^{(1)}W_{\text{global}}^{(2)}\sim M_{\text{pl}}^{-4}$. Considering the dimensionful superpotential this results in an overall infinite contribution, a finite sum of two terms and a vanishing product. In this decoupling limit one recovers the two independent global supersymmetry sectors with the naïve additive behavior in both the superpotential and the Kähler potential,

$$\begin{aligned}K(\phi,\overline{\phi},q,\overline{q})&=K^{(1)}(\phi,\overline{\phi})+K^{(2)}(q,\overline{q})\,,\\ W(\phi,q)&=W_{\text{eff}}^{(1)}(\phi)+W_{\text{eff}}^{(2)}(q)\,.\end{aligned} \tag{3.30}$$

However, one cannot use this split (3.30) and couple gravity back in [203]. As explained, in supergravity the definition (3.30) is not invariant under Kähler transformations in each sector separately and is valid only in a specific Kähler frame or, say, gauge dependent [201]. Another way to understand the result is to realize that the definition (3.30) does not lead to a Kähler metric and mass matrix that can be made block diagonal in the same basis [202], and thus there is no sense of 'independent' sectors.

Insisting on the separate Kähler invariance of (3.26), the two-sector action (3.23) reads

$$S = M_{\rm pl}^2 \int \mathrm{d}^4x\sqrt{g}\left[\frac{1}{2}R - g^{\mu\nu}(G^{(1)}_{i\bar{\jmath}}\partial_\mu\phi^i\partial_\nu\overline{\phi}^{\bar{\jmath}} + G^{(2)}_{a\bar{b}}\partial_\mu q^a\partial_\nu\overline{q}^{\bar{b}}) - VM_{\rm pl}^2\right] , \tag{3.31}$$

with

$$V(\phi,\overline{\phi},q,\overline{q}) = e^{G^{(1)}+G^{(2)}}\left(G^{(1)}_i G^{(1)i} + G^{(2)}_a G^{(2)a} - 3\right) . \tag{3.32}$$

We will allow ourselves to drop the sector label from G in the remainder, since $G^{(1)}_\phi = G_\phi$ and similarly for q. For a short overview of relevant conventions and identities in supergravity, we refer the reader to appendix D.

3.4.3 Zero mass mode for a stabilized sector

Anticipating the situation for an inflationary scenario we will analyze the mass spectrum of a stabilized q-sector in a de Sitter background. For Minkowski spaces it is known that the lightest mass in a stabilized sector scales with the supersymmetry breaking VEV G_a [194]. Here we extend the analysis to de Sitter vacua as the zeroth order approximation of slow-roll inflation. Already in this zeroth order approach we will show that a similar light mode develops in the stabilized sector. Throughout this discussion we assume that the potential V is kept positive by the presence of the 'inflationary' sector. In the next section we show that this result can be translated directly into an inflationary setting, where this light mode will affect the slow-roll dynamics.

Given that we insist the q-sector to be stabilized, we have $\partial_a V = 0$. In terms of the Kähler function $G(\phi,\overline{\phi},q,\overline{q})$ this means

$$(\nabla_a G_b)G^b = -G_a(1+e^{-G}V) . \tag{3.33}$$

If the q-ground state breaks supersymmetry, i.e. $G_a \neq 0$, we may rewrite it in terms of the supersymmetry breaking direction $f_a = G_a/\sqrt{G^b G_b}$,

$$(\nabla_a G_b)f^b = -f_a(1+e^{-G}V) . \tag{3.34}$$

For simplicity we will assume that the q-sector consists of only a single complex scalar field q, in which case we may write this equation as

$$\nabla_q G_q = -G_{q\overline{q}}(1+e^{-G}V)\widehat{G}_q^2 . \tag{3.35}$$

A hat $\widehat{q}$ on a complex number denotes the 'phase'-part of the number, $z = |z|\widehat{z} = |z|e^{i\,\arg(z)}$. As such $\widehat{G}_q = \sqrt{G^{q\overline{q}}}f_q$. Note that in an arbitrary supersymmetric configuration $G_a = 0$ there are no restrictions on $\nabla_a G_b$, but on a supersymmetry broken configuration this is no longer true. Were one to turn on supersymmetry breaking, one would first have to reach a surface in parameter space where this restriction can be imposed at the onset of supersymmetry breaking.

We will now compute the mass spectrum for the23 two modes of the complex scalar field q, at the hyper-surface defined by (3.35). The mass modes are given by the eigenvalues of the matrix

$$M^2 = \begin{pmatrix} V^q{}_q & V^q{}_{\overline{q}} \\ V^{\overline{q}}{}_q & V^{\overline{q}}{}_{\overline{q}} \end{pmatrix} , \tag{3.36}$$

which in our case means

$$m_q^{\pm} = \left(V^q{}_q \pm |V^q{}_{\overline{q}}|\right) = G^{q\overline{q}} \left(V_{q\overline{q}} \pm |V_{qq}|\right) . \tag{3.37}$$

Expanding the second derivatives of the potential (cf. appendix E) to first order in $|G_q|$, these eigenvalues are

$$m_q^- = e^G G^{q\overline{q}} \mathrm{Re}\big\{(\nabla_q \nabla_q G_q) \widehat{G^q}^3\big\} |G^q| + \mathcal{O}(|G_q|^2) , \tag{3.38}$$

$$m_q^+ = e^G \left[2(2 + e^{-G}V)(1 + e^{-G}V) - G^{q\overline{q}} \mathrm{Re}\big\{(\nabla_q \nabla_q G_q) \widehat{G^q}^3\big\} |G^q|\right] + \mathcal{O}(|G_q|^2) . \tag{3.39}$$

We see from (3.38) that in the limit of vanishing supersymmetry breaking the lightest mass mode becomes massless, just as in the case of Minkowski space [194].[2] It is important to note that this result depends crucially on taking the limit G_q to zero in the supersymmetry breaking direction. When supersymmetry is restored and both $G_q = 0$ and $G_{\overline{q}} = 0$, the phases of these vectors have no meaning. In fact, we see that then a new degree of freedom arises: $\nabla_q G_q$ becomes unrestricted which allows one to choose the masses freely.

The geometrical picture is that there is a whole plane of supersymmetric solutions where arbitrary masses are allowed. However, when supersymmetry is broken, the supersymmetry breaking direction has to align with its complex conjugate fixing one point on this plane where supersymmetry can be broken. In this point, the lightest mode becomes massless.

3.5 Two-sector inflation in supergravity

Generally, when inflation is described in supergravity, realistic matter resides in a hidden sector.[3] Supergravities descending from string theory often have additional hidden sectors as well. These sectors are always gravitationally coupled. In the previous section we have seen that for de Sitter vacua the hidden sector develops a light direction. In this section we will consider how

[2] The result can also be extended to hold for anti-de Sitter vacua. However, for $-2 < e^{-G}V < -1$, also a tachyonic mode develops.

[3] The supersymmetric partners of the Standard Model are not good inflaton candidates, as these partners are charged under the Standard Model gauge group and gauge fields taking part in inflation would lead to topological defects, e.g. [204, 205]. The exception could be a gravitationally non-minimally coupled Higgs field, e.g. [206, 207].

this light mode of the hidden sector can affect the naïve dynamics of the inflationary sector. We will show that despite the weakness of gravity, these effects can be large. Realistic slow-roll inflation is characterized by small numbers, the slow-roll parameters ϵ and η, and even small absolute changes to these numbers can be of the order of 100% in relative terms.

We will first briefly review the η-problem in the context of single field inflation in supergravity. Then we will explain what effects are to be expected when including an additional (hidden) sector. The section ends with calculating the relevant objects to determine the true dynamics of the full system.

3.5.1 Inflation and the η-problem in supergravity

In single scalar field models of inflation the spectrum of density perturbations is characterized by the two slow-roll parameters ϵ and η. To ensure that this spectrum matches the observed near scale invariance, both $\epsilon \ll 1$ and $\eta \ll 1$. Inflationary supergravity in its simplest form consists of a single complex scalar field, the inflaton, whose potential is generated by F-terms (3.24). The definition of η may be phrased as the lightest direction of the mass matrix in units of the Hubble rate $3H^2 = V$, i.e. η is the smallest eigenvalue of the matrix [208]

$$\widetilde{N}^I{}_J = \frac{1}{V}\begin{pmatrix} \nabla^i \nabla_j V & \nabla^i \nabla_{\bar{\jmath}} V \\ \nabla^{\bar{\imath}} \nabla_j V & \nabla^{\bar{\imath}} \nabla_{\bar{\jmath}} V \end{pmatrix} , \tag{3.40}$$

where the tilde on $\widetilde{N}$ indicates that this value of η is defined with respect to the inflaton sector only and $I \in \{i, \bar{\imath}\}$, $J \in \{j, \bar{\jmath}\}$, respectively.[4] From the second ϕ-derivative of V,

$$V_{i\bar{\jmath}} = G_{i\bar{\jmath}} V + G_i V_{\bar{\jmath}} + G_{\bar{\jmath}} V_i - G_i G_{\bar{\jmath}} V + e^G \left[R_{i\bar{\jmath}k\bar{l}} G^k G^{\bar{l}} + G^{k\bar{l}} \nabla_i G_k \nabla_{\bar{\jmath}} G_{\bar{l}} + G_{i\bar{\jmath}} \right] , \tag{3.41}$$

we see that a natural value for η is $V^i_{\ j}/V \sim \nabla^i G_j \sim 1$ is unity. Therefore, we must tune G_i, $\nabla_i G_j$ and $R_{i\bar{\jmath}k\bar{l}}$ so that $V^i_{\ j} = \mathcal{O}(10^{-3})V$. The necessity of this tuning is known as the η-problem.

As shown in [209], successful inflation is achievable if one tunes the Kähler function G such that

$$R_{i\bar{\jmath}k\bar{l}} f^i f^{\bar{\jmath}} f^k f^{\bar{l}} \lesssim \frac{2}{3}\frac{1}{1+\gamma} , \tag{3.42}$$

where $\gamma = e^{-G}V/3$ is inversely proportional to an overall mass scale $m_{3/2} = e^{G/2}$, which is related to the gravitino mass and $R_{i\bar{\jmath}k\bar{l}}$ is the Riemann tensor of the inflaton sector. As $f^i f_i = 1$ the above equation defines the normalized sectional curvature along the direction of supersymmetry breaking. The constraint becomes stronger as $\gamma \gg 1$, thus as $H \gg m_{3/2}$. When the bound is met, one can always tune η to be small by tuning G_i, $\nabla_i G_j$ and $R_{i\bar{\jmath}k\bar{l}}$.

Finding a suitably tuned supergravity potential from a (UV-complete) string theoretical set-up has proven to be incredibly difficult [210, 211], but possible [169, 173, 212]. Currently, in

[4] A careful definition based on the kinetic behavior of the inflaton field is done in [181, 182]. In the slow-roll, slow-turn limit, it reduces to the definition of η given here.

models with correctly tuned slow-roll parameters it is typically assumed that the 'hidden sector' does not affect the fine-tuning of parameters. The subject of this work is to examine whether such an assumption is justified and hence how relevant tuned models are that only consider the inflationary sector.

3.5.2 Stability of the hidden sector during inflation

Having reviewed the η-problem in single sector supergravity theories, we will now consider if and how the fields in the hidden sector can affect the inflationary evolution. From the diagonalization of the kinetic terms in (3.23) the distinction between ϕ-fields and q-fields is explicit, leading naturally to an inflationary and a hidden sector. We will again assume these sectors to both consist of only one complex scalar field, ϕ and q respectively. The argument we shall present can already be made in a two-field system. It carries through to multi-field models because the field ϕ is viewed as the inflaton in an effective single field inflationary model, while the field q can be seen as the lightest mode in the hidden sector. Following the usual practice [187, 188, and references therein], we assume that inflation is solved by tuning the inflationary sector only, including obtaining satisfactory values for the slow-roll parameters from a phenomenological viewpoint. As a result all data in the inflationary sector are fixed and known. Contrarily, the hidden sector is left unspecified and the restrictions we find on it are a function of model specific parameters of the inflaton sector only.

To ensure that the hidden sector does not take part in the inflationary dynamics, one generally assumes that the fields in the hidden sector are stabilized in a ground state at a constant field value $q = q_0$ throughout inflation

$$\partial_q V|_{q_0} = 0 \tag{3.43}$$

and, hence, are not dynamical. Clearly this is true if $G_q = 0$, i.e. when the ground state of the hidden sector preserves supersymmetry. As was shown in detail in [191, 199–202, 213–215], when $G_q = 0$ the ground state of the hidden sector decouples gravitationally from the inflationary sector and the inflationary sector truly determines the inflationary evolution without any contributions from the hidden sector.

The case we examine here is when supersymmetry is broken in the hidden sector, $G_q \neq 0$. The first thing to note is that the stability assumption (3.43) cannot be met anymore. In supergravity the position $q = q_0$ of the minimum of the potential is given by

$$V_q = G_q V(\phi, \overline{\phi}, q, \overline{q}) + e^{G(\phi,\overline{\phi},q,\overline{q})} \left((\nabla_q G_q) G^q + G_q \right) = 0 \; , \tag{3.44}$$

which shows that for $G_q \neq 0$ the ground state q_0 depends on the inflaton field ϕ, through $V(\phi, \overline{\phi}, q, \overline{q})$ and $G(\phi, \overline{\phi}, q, \overline{q})$. In the situation of unbroken supersymmetry, $G_q = 0$, all ϕ-dependence drops out, but for $G_q \neq 0$ we see that it is impossible to keep the position of the minimum constant during inflation. As the inflaton ϕ rolls down the inflaton direction, the 'stabilized' hidden scalar q will change its value. It is clear that the assumption of a vanishing $V_q = 0$ for all q is incompatible with $G_q \neq 0$ and we should therefore abandon it. This in turn means that the hidden sector field q must be dynamical, through its equation of motion. Since we still want to identify the field ϕ as the inflaton in the sense that it drives the cosmological dynamics, we have to assume that q moves very little. We must therefore also assume a slow-roll, slow-turn approximation to the solution of the q equation of motion

$$\dot{q} = \frac{G^{q\bar{q}} V_{\bar{q}}}{3H} \ . \tag{3.45}$$

The statement that the cosmological dynamics is driven by the ϕ-sector means that $\|\dot{q}\| \ll \|\dot{\phi}\|$, where $\|\dot{q}\| \equiv \sqrt{G_{q\bar{q}}\dot{q}\dot{\bar{q}}}$, etc. Through both slow-roll equations of motion this equates to $\|V_q\| \ll \|V_\phi\|$ or $\epsilon_q \ll \epsilon_\phi$,

As the hidden sector has now become dynamical, we have to treat the system as a multi-field inflationary model. Since it is impossible to diagonalize the Kähler transformations and mass matrix simultaneously, the fields will mix in the case of a hidden sector with broken supersymmetry [201]. In the next section we will study the consequences of this mixing by explicitly diagonalizing the mass matrix of the full two-field system. From the result we shall find three possible effects on the inflationary dynamics.

First, the lightest masses of fields from the different sectors can be too close together. It is obvious that one cannot consider an effective single field model if this is the case, since for the dynamics to be independent of initial conditions, the lightest field needs to be much lighter than the other fields. When the masses of the two fields are similar, both of them contribute to the dynamics, resulting into a multi-field rather than a single field inflationary scenario. As is known from the literature, a multi-field inflationary model will produce effects such as isocurvature modes, e.g. [216–230], features in the power spectrum, e.g. [183, 186, 193, 231] and non-Gaussianities, e.g. [232–241], pointing to a qualitatively different model.

Second, a change of the true value of η can occur. We have assumed the inflaton sector to be tuned in such a way that it agrees with observed values for the slow-roll parameters. If the effects of the hidden sector on the total dynamics are such that η will change significantly, the initial naïve tuning would be of no meaning and one would have to start the tuning process all over again after the hidden sector has been added. Again we note that there is no contribution in the case of unbroken supersymmetry in the hidden sector, since we shall show that the contribution to η from the hidden sector is mostly determined by the cross terms in the mass matrix,

$$V_{\phi q} = G_\phi V_q + G_q V_\phi - G_\phi G_q V \ , \tag{3.46}$$

which vanish when $G_q = 0$.

Third, a complete change of the sector that determines η is possible. It is possible that the eventual η-parameter is still within the limits of its naïve tuned value, satisfying the second bound, but instead it is determined by the hidden sector rather than the inflationary sector. Any initial control obtained by tuning the inflationary sector is superseded by the sheer coincidental configuration of the hidden sector.

3.5.3 The mass matrix of a two-sector system

To investigate when effects from the hidden sector are to be expected, we need to calculate the eigenvalues of the mass matrix of the full two-field system. Since we assume the inflationary evolution to be in the slow-roll, slow-turn regime, the dynamics is completely potential energy dominated. The mass matrix of the full two-field system determines which directions are stable or steep, as characterized by the eigenvalues of this matrix. Normalizing by $1/V$ to obtain the value of η directly, the matrix we want to diagonalize is the 4×4-matrix

$$N^A{}_B = \frac{1}{V}\begin{pmatrix} \nabla^\alpha \nabla_\beta V & \nabla^\alpha \nabla_{\overline{\beta}} V \\ \nabla^{\overline{\alpha}} \nabla_\beta V & \nabla^{\overline{\alpha}} \nabla_{\overline{\beta}} V \end{pmatrix}, \tag{3.47}$$

where $A \in \{\alpha, \overline{\alpha}\}$ and $B \in \{\beta, \overline{\beta}\}$ run over both fields ϕ and q and their complex conjugates. Equation (3.47) is to be evaluated at a point near $q_0 = q_0(\phi_0)$, where q_0 is such that $\partial_q V(q_0) = 0$, with ϕ_0 indicating the beginning of inflation. As is clear from the discussion of section 3.5.2 we cannot truly expect the hidden sector to be stabilized throughout the inflationary evolution. Nevertheless we may consider $\partial_q V(q_0) = 0$ at a certain point $q_0 = q_0(\phi_0)$, with $\|\partial_q V\| \ll \|\partial_\phi V\|$ around q_0 in accordance with the restriction $\epsilon_q \ll \epsilon_\phi$.

The mass matrix is Hermitian and, considering again a two-field system, can be put in the form

$$N^A{}_B = \frac{1}{V}\begin{pmatrix} \nabla^\phi V_\phi & \nabla^\phi V_{\overline{\phi}} & \nabla^\phi V_q & \nabla^\phi V_{\overline{q}} \\ \nabla^{\overline{\phi}} V_\phi & \nabla^{\overline{\phi}} V_{\overline{\phi}} & \nabla^{\overline{\phi}} V_q & \nabla^{\overline{\phi}} V_{\overline{q}} \\ \nabla^q V_\phi & \nabla^q V_{\overline{\phi}} & \nabla^q V_q & \nabla^q V_{\overline{q}} \\ \nabla^{\overline{q}} V_\phi & \nabla^{\overline{q}} V_{\overline{\phi}} & \nabla^{\overline{q}} V_q & \nabla^{\overline{q}} V_{\overline{q}} \end{pmatrix}, \tag{3.48}$$

by a coordinate transformation. Diagonalizing the full matrix in general is involved. Therefore, we adopt the strategy to diagonalize the two sectors separately and then pick the lightest modes only. The first step yields

$$N^A{}_B = \begin{pmatrix} \frac{1}{V}(V^\phi{}_\phi - |V^\phi{}_{\overline{\phi}}|) & 0 & A_{11} & A_{12} \\ 0 & \frac{1}{V}(V^\phi{}_\phi + |V^\phi{}_{\overline{\phi}}|) & A_{21} & A_{22} \\ \overline{A}_{11} & \overline{A}_{21} & \frac{1}{V}(V^q{}_q - |V^q{}_{\overline{q}}|) & 0 \\ \overline{A}_{12} & \overline{A}_{22} & 0 & \frac{1}{V}(V^q{}_q + |V^q{}_{\overline{q}}|) \end{pmatrix}, \tag{3.49}$$

with

$$A = \frac{1}{2V}\begin{pmatrix} -\widehat{V}_{\overline{\phi}\overline{\phi}} & \widehat{V}_{\overline{\phi}\overline{\phi}} \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} V^\phi{}_q & V^\phi{}_{\overline{q}} \\ V^{\overline{\phi}}{}_q & V^{\overline{\phi}}{}_{\overline{q}} \end{pmatrix} \begin{pmatrix} -\widehat{V}_{\overline{q}\overline{q}} & \widehat{V}_{\overline{q}\overline{q}} \\ 1 & 1 \end{pmatrix}. \tag{3.50}$$

Here, the first matrix is the inverse of the similarity transformation of the ϕ-sector and the last matrix diagonalizes the q-sector.

In general the eigenmodes in the individual sectors will be different, one always being smaller than the other. Dynamically the most relevant direction is the lightest mode of each sector, but by restricting to these light directions, one assumes a hierarchy already within the sectors. For the inflationary sector this is phenomenologically justified if we assume that inflation is described by a single field, where we know that $V^{\phi}_{\ \phi}$ and $V^{\phi}_{\ \overline{\phi}}$ combine such that a light mode appears with mass ηV, much lighter than the other mass modes. For the hidden sector we will simply assume that a large enough hierarchy between mass modes exists. This will simplify matters without weakening our result. By including only the lightest mode of the hidden sector, we can already show that the true dynamics is in many cases not correctly described by the naïve inflaton sector. Our case would only be more strongly supported if we would include the heavy mode of the hidden sector, but this is technically more involved. Projecting on the light directions, we get a sub-matrix of light mass modes

$$N_{\text{light}} = \begin{pmatrix} \lambda_\phi & A_{11} \\ \overline{A}_{11} & \lambda_q \end{pmatrix}, \tag{3.51}$$

with

$$\lambda_\phi = \frac{1}{V}\left(V^{\phi}_{\ \phi} - |V^{\phi}_{\ \overline{\phi}}|\right) = \frac{G^{\phi\overline{\phi}}}{V}\left(V_{\phi\overline{\phi}} - |V_{\phi\phi}|\right), \tag{3.52}$$

$$\lambda_q = \frac{1}{V}\left(V^{q}_{\ q} - |V^{q}_{\ \overline{q}}|\right) = \frac{G^{q\overline{q}}}{V}\left(V_{q\overline{q}} - |V_{qq}|\right), \tag{3.53}$$

$$A_{11} = \frac{G^{\phi\overline{\phi}}}{2V}\left(\widehat{V}_{\overline{q}\overline{q}}\widehat{V}_{\phi\phi}V_{\overline{\phi}q} - \widehat{V}_{\overline{q}\overline{q}}V_{\phi q} + V_{\phi\overline{q}} - \widehat{V}_{\phi\phi}V_{\overline{\phi}\overline{q}}\right). \tag{3.54}$$

The eigenvalues of this two-field system are given by

$$\mu_\pm = \frac{1}{2}\left(\lambda_\phi + \lambda_q\right) \pm \frac{1}{2}\sqrt{\left(\lambda_q - \lambda_\phi\right)^2 + 4|A_{11}|^2}. \tag{3.55}$$

Since $\mu_- < \mu_+$ the second slow-roll parameter for the full system is given by $\eta = \mu_-$.

3.6 Dynamics due to the hidden sector

In slow-roll and slow-turn approximation, the mass modes $\mu_\pm$ from (3.55) determine the dynamics of the full system. In general the true dynamics will deviate from the naïve single sector evolution. As explained in section 3.5.2 it is necessary to put constraints on the full system for the true dynamics to still (largely) agree with the initial naïve dynamics. We will quantify these constraints in terms of the hidden sector light mode λ_q and the dynamical cross coupling $|A_{11}|$ between sectors. The results are graphically summarized in figures 3.8 and 3.9. In section 3.6.2 and figure 3.10 we will discuss the result again but then interpreted from the viewpoint of supergravity. Finally we will explain that a simple application of these bounds implies that the Standard Model cannot be ignored during cosmological inflation, if Standard Model supersymmetry breaking is independent of the inflaton sector.

3.6.1 Conditions on the hidden sector data

From (3.55) we see that the light modes λ_ϕ, λ_q from the two separate sectors mix through a cross coupling $|A_{11}|$ and combine to the true eigenvalues $\mu_\pm$ of the full two-sector system. As explained in 3.5.2, for the inflaton sector to still describe the cosmological evolution and the η-parameter reliably, the three constraints it must obey are (1) the bound arising from demanding a hierarchy between $\mu_\pm$ to prevent multi-field effects, (2) the bound arising from demanding the second slow-roll parameter $\mu_- = \eta$ to not change its value too much and (3) the bound from demanding that η is mostly determined by the ϕ-sector rather than the q-sector.

To prevent multi-field effects from setting in we take as a minimum hierarchy that μ_+ is at least five times as heavy as μ_- in units of the scale of the problem, $|\mu_-|$,

$$\frac{\mu_+ - \mu_-}{|\mu_-|} > 5. \tag{3.56}$$

This bound is rather arbitrary, but clearly a hierarchy between μ_+ and μ_- must exist. Calculations in [183] show that for a mass hierarchy $\lesssim 5$ multifield effects are typically important.

The second bound is given by the A_{11}-dependence of μ_-. The value of the second slow-roll parameter from the single field inflationary model only is $\eta_{\text{naïve}} = \lambda_\phi$. In the full two-sector system, μ_- takes over the role as the true second slow-roll parameter $\eta_{\text{true}} = \mu_-$. The contribution to the actual η-parameter from the presence of the hidden sector is therefore

$$\Delta\eta = \mu_- - \lambda_\phi = \frac{1}{2}\left[(\lambda_q - \lambda_\phi) - \sqrt{(\lambda_q - \lambda_\phi)^2 + 4|A_{11}|^2}\right] , \tag{3.57}$$

which is always negative. We argue that this difference should stay within $|\Delta\eta/\lambda_\phi| < 0.1$, i.e. η should not change by more than 10%. This choice for the range of η is given by current experimental accuracy. Current experiments can only determine $n_s = 1 - 6\epsilon + 2\eta$. WMAP has a 1σ error of 6.53% [134], Planck will have an error of 0.70% [242]. For $n_s - 1$, assuming 0.96, this gives a 17.5% error on the combination of $-6\epsilon + 2\eta$, which means an uncertainty of about 10% on the value of η.

We will examine λ_q, A_{11} in units of $|\lambda_\phi|$ and exclude regions in which the hidden sector affects the tuned inflationary sector too much. The analysis is best done separately for the cases $\lambda_\phi = \eta_{\text{naïve}} > 0$ and $\lambda_\phi = \eta_{\text{naïve}} < 0$ because of the qualitative differences between these cases.

The case $\eta_{\text{naïve}} > 0$

We first examine the hierarchy bound as explained above and focus first on the situation where $\mu_- > 0$. In this case (3.56) means that we demand

$$\frac{\mu_+ - 6\mu_-}{\lambda_\phi} = \frac{1}{2}\left[-5\left(\frac{\lambda_q}{\lambda_\phi} + 1\right) + 7\sqrt{\left(\frac{\lambda_q}{\lambda_\phi} - 1\right)^2 + 4\left(\frac{|A_{11}|}{\lambda_\phi}\right)^2}\right] > 0, \tag{3.58}$$

which allows us to solve λ_q/λ_ϕ as a function of $|A_{11}|/\lambda_\phi$,

$$\left(\frac{12}{35}\right)^2\left(\frac{\lambda_q}{\lambda_\phi} - \frac{37}{12}\right)^2 + \left(\frac{2\sqrt{6}}{5}\right)^2\left(\frac{|A_{11}|}{\lambda_\phi}\right)^2 = 1. \tag{3.59}$$

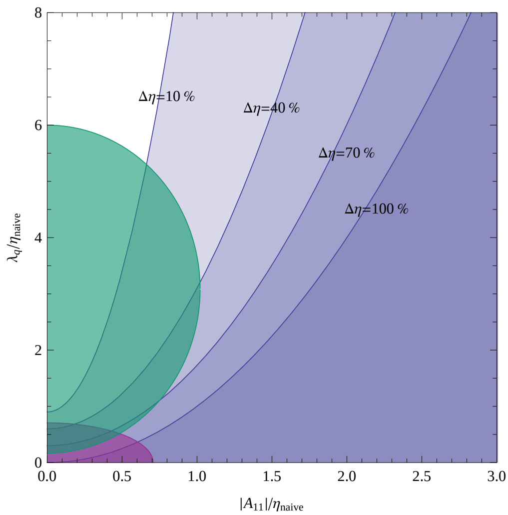


Figure 3.8: *Bounds from a dynamical hidden sector for $\eta_{\text{naïve}} > 0$. The multi-field constraint excludes an ellipse near the λ_q-axis (shaded in green). The bound from having too much effect on η excludes large $|A_{11}|$ (shaded with increasing intensities of blue for larger deviations). Around $\lambda_q = A_{11} = 0$ the hidden sector mode λ_q rather than λ_ϕ determines η, excluding that region as well (shaded in purple).*

This excludes everything inside the ellipse demarcating the green region in figure 3.8. The case $\mu_- < 0$ is not relevant as it is already excluded by the second bound.

For this second bound, to be somewhat more general than the observationally inspired constraint $\Delta\eta/\lambda_\phi > -0.1$, we give the bound $\Delta\eta/\lambda_\phi > -f$. Solving for λ_q this gives

$$\frac{\lambda_q}{\lambda_\phi} > 1 - f + \frac{1}{f}\left(\frac{|A_{11}|}{\lambda_\phi}\right)^2, \tag{3.60}$$

as is indicated in blue in figure 3.8. Note that since the true value of η is always lower than $\eta_{\text{naïve}}$ (see [180] for some specific examples), a change in η of 100% means that η changes sign from its naïve value. This shows that we were justified to only consider positive μ_- in the hierarchy bound earlier.

The third bound is given by a λ_q-dominance in μ_-. Since λ_ϕ and λ_q are treated on equal footing in μ_-, the true η is dominantly determined by the smallest eigenvalue, which is not necessarily λ_ϕ. When $\lambda_\phi \gg \lambda_q$ and $\lambda_\phi \gg |A_{11}|$ we see immediately that the true $\eta = \mu_-$ is determined by λ_q and is *independent* of λ_ϕ,

$$\mu_- = \frac{1}{2}\left[(\lambda_q + \lambda_\phi) - \lambda_\phi\left(1 - \frac{\lambda_q}{\lambda_\phi} + \mathcal{O}\left(\frac{\lambda_q^2}{\lambda_\phi^2}, \frac{|A_{11}|^2}{\lambda_\phi^2}\right)\right)\right]. \tag{3.61}$$

It is clear that this arguments excludes the lower left corner of parameter space. We will take the bound to be $1/\sqrt{2}$ such that $(\lambda_q/\lambda_\phi)^2, (|A_{11}|/\lambda_\phi)^2 < 1/2 \ll 1$, the radius of convergence

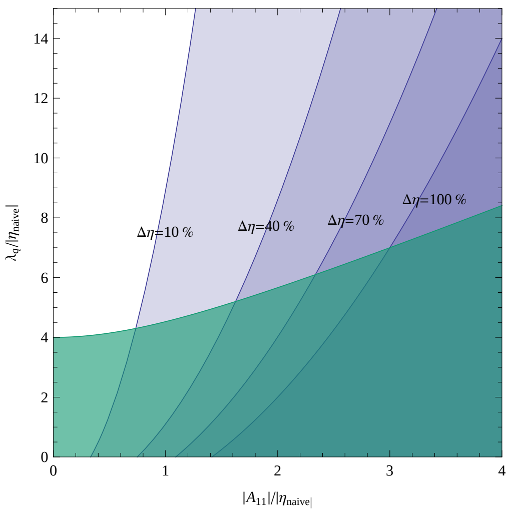


Figure 3.9: *Bounds from a dynamical hidden sector for $\eta_{\text{naïve}} < 0$. The multi-field bound excludes a hyperbola starting at $\lambda_q = 4|\lambda_\phi|$ and, in particularly small λ_q (shaded in green). The bound from having too much effect on η excludes the large $|A_{11}|$-region (shaded with increasing intensities of blue for larger deviations), but leaves open in particular the full range of λ_q.*

of this Taylor expansion. Contrarily to the somewhat debatable bounds imposed by $\Delta\eta/\lambda_\phi$, the points within this circle are truly excluded because they violate one of the core assumptions in the approach, viz. that the ϕ-sector is responsible for all cosmological dynamics including determining the value of η. The circle

$$\left(\frac{\lambda_q}{\lambda_\phi}\right)^2 + \left(\frac{|A_{11}|}{\lambda_\phi}\right)^2 = \frac{1}{2}, \tag{3.62}$$

is indicated as the purple region in the figure.

In figure 3.8 we have indicated in which regions of λ_q/λ_ϕ- and $|A_{11}|/\lambda_\phi$-parameter space the effects of a hidden sector can be rightfully ignored. We have shown that all negative values of λ_q are excluded and only in the region with large λ_q/λ_ϕ and small $|A_{11}|/\lambda_\phi$ there are no large effects from the hidden sector. This result is qualitatively easily understood, as the hidden sector with broken supersymmetry will still decouple if the masses in the hidden sector are truly large. We argue that this possibility is too easily assumed to be the case in the literature without considering the actual hidden constraints it imposes on the hidden sector. These hidden assumptions should be mentioned explicitly and one should show that they can be obtained.

THE CASE $\eta_{\text{NAÏVE}} < 0$

In the case that $\lambda_\phi = \eta_{\text{naïve}}$ is negative, the last bound of section 3.6.1 does not impose any condition on $\lambda_q/|\lambda_\phi|, |A_{11}|/|\lambda_\phi|$-parameter space. When $\lambda_\phi < 0$, i.e. when $\lambda_\phi = -|\lambda_\phi|$, the eigenvalues can be written as

$$\mu_\pm = \frac{|\lambda_\phi|}{2}\left[\left(\frac{\lambda_q}{|\lambda_\phi|}-1\right) \pm \sqrt{\left(\frac{\lambda_q}{|\lambda_\phi|}+1\right)^2 + 4\left|\frac{A_{11}}{\lambda_\phi}\right|^2}\,\right] , \tag{3.63}$$

which means that μ_- is not determined by λ_q to first order in $\lambda_q/|\lambda_\phi|$ but by λ_ϕ as should be,

$$\mu_- = \frac{|\lambda_\phi|}{2}\left[\left(\frac{\lambda_q}{|\lambda_\phi|}-1\right) - \left(1+\frac{\lambda_q}{|\lambda_\phi|}+\ldots\right)\right] . \tag{3.64}$$

However, by the hierarchy bound the small $\lambda_q/|\lambda_\phi|$-regime does get excluded. Since μ_- is always negative in this case,

$$\mu_- \leq \frac{|\lambda_\phi|}{2}\left[\left(\frac{\lambda_q}{|\lambda_\phi|}-1\right) - \left|\frac{\lambda_q}{|\lambda_\phi|}+1\right|\right] = -|\lambda_\phi| , \tag{3.65}$$

equation (3.56) translates into

$$\frac{\mu_+ + 4\mu_-}{|\lambda_\phi|} = \frac{1}{2}\left[5\left(\frac{\lambda_q}{|\lambda_\phi|}-1\right) - 3\sqrt{\left(\frac{\lambda_q}{|\lambda_\phi|}+1\right)^2 + 4\left|\frac{A_{11}}{\lambda_\phi}\right|^2}\,\right] > 0 . \tag{3.66}$$

This excludes everything beneath the upper branch of the hyperbola given by the line

$$\frac{\lambda_q}{|\lambda_\phi|} > \frac{17}{8} + \frac{1}{8}\sqrt{15^2 + 28\left|\frac{A_{11}}{\lambda_\phi}\right|^2} , \tag{3.67}$$

which is shaded green region in figure 3.9.

The final constraint on the parameter space comes from the bound on the change in η, see the previous paragraph on the $\eta_{\text{naïve}} > 0$-case for a discussion. In the blue region in figure 3.9 we have indicated the bound $|\Delta\eta/\lambda_\phi| < f$, which means

$$\frac{\lambda_q}{|\lambda_\phi|} > -1 - f + \frac{1}{f}\left|\frac{A_{11}}{\lambda_\phi}\right|^2 , \tag{3.68}$$

for different fractions of f.

In figure 3.9 we have indicated in which regions of $\lambda_q/|\lambda_\phi|$- and $|A_{11}|/|\lambda_\phi|$-parameter space the effects of a hidden sector can be rightfully ignored after imposing both constraints. As in the case for $\eta_{\text{naïve}} > 0$, the only allowed region is for large $\lambda_q/|\lambda_\phi|$ and small $|A_{11}|/|\lambda_\phi|$. Note that all values of $\lambda_q < 4$ are explicitly excluded by the imposed bounds.

3.6.2 Conditions on supergravity models

In principle, figures 3.8 and 3.9 provide all the information needed to verify whether the hidden sector of a given model may be neglected while studying the inflationary dynamics. Through equations (3.53–3.54) and the expressions for $V_{\alpha\beta}$ as summarized in appendix D, one can

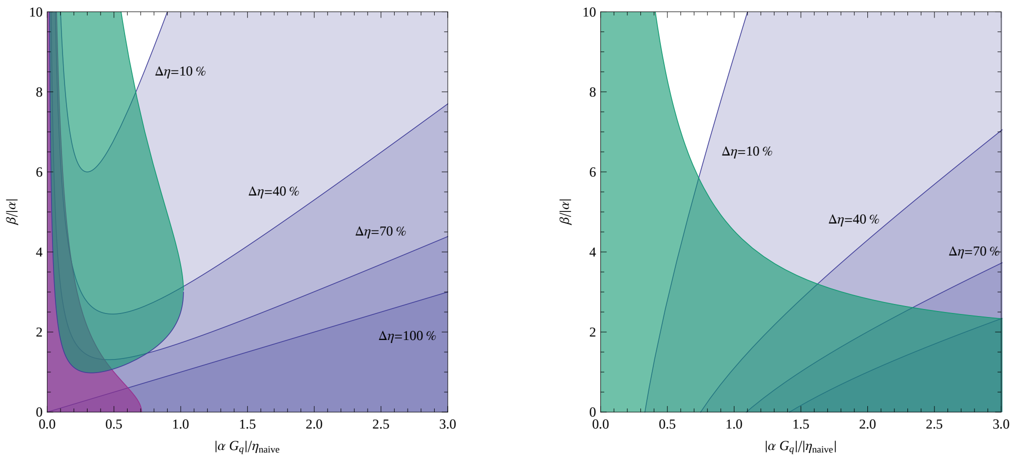


Figure 3.10: *Excluded regions for the supergravity parameter range for $|G_q|$ and β, which contains in particular $\nabla_q\nabla_q G_q$, in units of $|\eta_{\text{naïve}}|$ and $|\alpha|$, which contains ϵ_ϕ and G_ϕ. The indicated regions come from the multi-field bound (shaded in green), the correct identification of sectors (shaded in purple) and allowing only for small deviations of η (shaded in higher intensities of blue for larger deviations). The left (right) picture describes the case $\eta_{\text{naïve}} > 0$ ($\eta_{\text{naïve}} < 0$).*

explicitly calculate the corresponding λ_q and A_{11} for a given model and compare them with the figures. However, we would like to have some direct intuition about the dependence of the excluded regions on the supergravity data. In this section we will investigate how much we can say about this in general without having to specify a model. The main question to answer is whether the fact that λ_q and A_{11} are determined by a supergravity theory, provides any additional constraint on which regions are obtainable to begin with. The answer to this question turns out to be that a priori supergravity is not restrictive enough to exclude any of the regions in λ_q, A_{11}-parameter space.

The easiest way to translate figures 3.8 and 3.9 in terms of supergravity data would be to simply map the regions into supergravity parameter space. Unfortunately the expressions (3.53) and (3.54) are highly nonlinear and depend on too many supergravity variables to conveniently represent figures 3.8 and 3.9 in terms of supergravity data. However, for small $|G_q|$ this does turn out to be possible.

Using the expressions for $V_{\alpha\beta}$ in (3.54), yields

$$A_{11} = \alpha(\phi,\overline{\phi},q,\overline{q})|G_q|, \qquad \text{with} \tag{3.69}$$
$$\alpha(\phi,\overline{\phi},q,\overline{q}) = \frac{G^{\phi\overline{\phi}}}{2}\left(\widehat{G}_{\overline{q}} - \widehat{V}_{\overline{q}\overline{q}}\widehat{G}_q\right)\left(\left(\frac{V_\phi}{V} - G_\phi\right) - \widehat{V}_{\phi\phi}\left(\frac{V_{\overline{\phi}}}{V} - G_{\overline{\phi}}\right)\right).$$

From this equation we learn that A_{11} vanishes in the limit $G_q \to 0$, which makes sense as we know that the two sectors should decouple in the limit of restored supersymmetry. It is difficult to retrieve more information from this explicit expression of A_{11} in terms of supergravity data.

In principle $A_{11}(|G_q|,\ldots)$ may be inverted to give some function $|G_q|(A_{11},\ldots)$, but this is more tricky than (3.69) suggests. Although we have managed to extract one factor of G_q, the function $\alpha(\phi,\overline{\phi},q,\overline{q})$ still depends on G_q through the phases of $\widehat{V_{\overline{q}\overline{q}}}$ and $\widehat{V_{\phi\phi}}$, making it hard to perform the inversion explicitly.

The expression for λ_q looks even worse,

$$\lambda_q = \frac{G^{q\overline{q}}}{V}\left(V_{q\overline{q}} - \sqrt{V_{qq}V_{\overline{q}\overline{q}}}\right). \tag{3.70}$$

At this stage we have even refrained from substituting in the expressions for $V_{q\overline{q}}$, V_{qq} and its complex conjugate. The square root clearly shows that the dependence of λ_q on $|G_q|$ and the other supergravity data is extremely involved and difficult to invert. To get a useful expression we revert to the result of section 3.4.3 and consider λ_q in the small $|G_q|$-regime by performing a Taylor expansion. Copying from (3.38), we find

$$\lambda_q = \beta(\phi,\overline{\phi},q,\overline{q})|G^q| + \mathcal{O}(|G_q|^2), \qquad \text{with} \tag{3.71}$$
$$\beta(\phi,\overline{\phi},q,\overline{q}) = \frac{G^{q\overline{q}}}{e^{-G}V}\mathrm{Re}\big\{(\nabla_q\nabla_q G_q)\widehat{G^q}^3\big\}.$$

Having obtained the relations (3.69) and (3.71) we can now accommodate the reader with a graph of the allowed and excluded regions directly in terms of the supergravity data. For small $G_q \ll 1$ both λ_q and $|A_{11}|$ scale linearly with G_q, making it relatively easy to rewrite the bounds we found $\lambda_q/|\lambda_\phi| = \lambda_q/|\lambda_\phi|\,(|A_{11}|/|\lambda_\phi|)$ in terms of G_q, α and β as $\beta/|\alpha| = \beta/|\alpha|\,(|\alpha G_q|/|\lambda_\phi|)$. The resulting figure is depicted in 3.10. Note that α and β are still underdetermined — depending on $R_{q\overline{q}q\overline{q}}$ and $\nabla_q\nabla_q G_q$ at higher orders in $|G_q|$ — and are naturally of order 1. It is these numbers that determine where in figure 3.10 the model under investigation lies.

3.6.3 Inflation and the Standard Model

As a simple application of the previous section, we can consider to what extent the Standard Model ought to be included in any reliable supergravity model for cosmological inflation. Our current understanding of Nature includes a present-day supersymmetrically broken Standard Model after an inflationary evolution right after the big bang. As such the combined model is exactly that of a two-sector supergravity theory with an inflationary and a hidden sector whose ground state breaks supersymmetry in which it resides throughout the inflationary era.

Supersymmetry in the Standard Model sector can either have been broken by gravity mediation of the inflaton sector or by a mechanism in the Standard Model sector itself. The first situation would be consistent approach as far as our analysis goes: as $G_q = 0$ the sector decouples from the inflationary dynamics, can be stabilized and the slow-roll parameters are reliably determined from the inflaton sector alone. Nevertheless, from the point of view of our understanding of the Standard Model it would be unsatisfactory to not know the precise mechanism behind its supersymmetry breaking and (complete) models describing such mechanisms would still have to be analyzed to shed light on the situation.

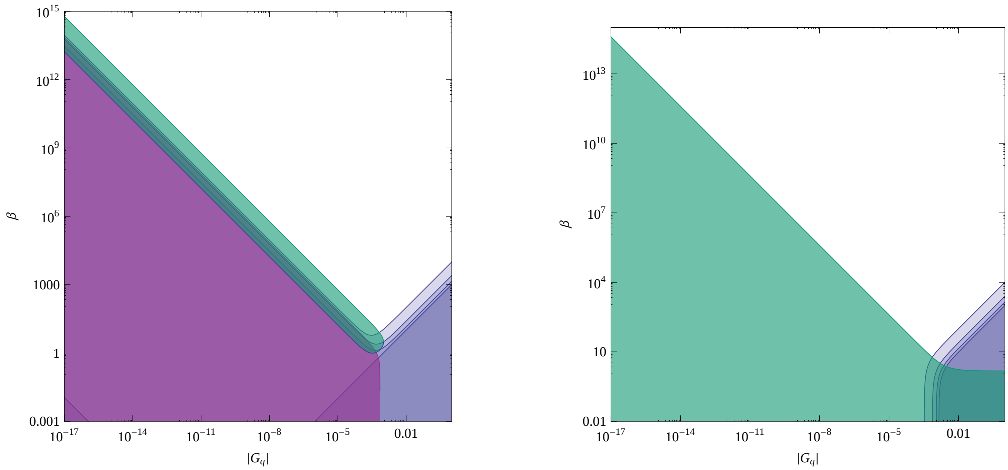


Figure 3.11: *The effects of the multi-field bound (shaded in green), the identification of the correct inflaton sector (shaded in purple) and the small deviations of η (shaded in blue) on a doubly logarithmic scale for $\eta_{\text{naïve}} > 0$ (left) and $\eta_{\text{naïve}} < 0$ (right). The approximate location of the Standard Model supergravity data is indicated with a red bar, showing that a large range of parameters is excluded. In this plot $\alpha = 1$ and $\lambda_\phi = \eta_{\text{naïve}} = 10^{-3}$.*

In the second situation $G_q \neq 0$ and we should apply the results of the previous sections. The field q may be seen as some light scalar degree of freedom in the (supersymmetrically broken) Standard Model. We assume the standard lore, that supersymmetry is broken in the Standard Model at a scale of about 1 TeV. In the F-term scalar potential, this scale enters via G_q. To determine the correct numerical value, we relate our dimensionless definition of the Kähler function to the standard dimensionful definition. Dimensionful quantities are denoted with a tilde in the following.[5] We recall from section 3.4.2 that in order to have a non-vanishing vacuum energy, the superpotential in both sectors must have a non-zero constant term $W_0^{(1)} = m_\Lambda^{(1)}/M_{\text{pl}}$, $W_0^{(2)} = m_\Lambda^{(2)}/M_{\text{pl}}$, which accounts for the always present gravitational coupling between the sectors. Hence, the dimensionful constant term in the total superpotential (3.29) has value $\widetilde{W}_0^{\text{tot}} = W_0^{(1)} W_0^{(2)} M_{\text{pl}}^3 = m_\Lambda^{(1)} m_\Lambda^{(2)} M_{\text{pl}}$. In contrast, the supergravity quantities $\widetilde{K}^{(2)}$ and $\widetilde{W}_{\text{eff}}^{(2)} = \widetilde{W}_0^{(1)} \widetilde{W}_{\text{global}}^{(2)}$ describing the Standard Model are naturally of the order of the TeV-scale, $[\widetilde{W}_{\text{eff}}^{(2)}] = \text{TeV}^3$, $[\partial_{\tilde{q}} \widetilde{K}^{(2)}] = \text{TeV}$. We relate the scale of supersymmetry breaking $\widetilde{G}_{\tilde{q}}$ to the superpotential via

$$\widetilde{G}_{\tilde{q}} = \frac{M_{\text{pl}}^2}{\widetilde{W}} \left(\partial_{\tilde{q}} \widetilde{W} + \frac{\partial_{\tilde{q}} \widetilde{K}^{(2)}}{M_{\text{pl}}^2} \widetilde{W} \right) , \tag{3.72}$$

[5] E.g. in dimensionful units $[\widetilde{G}] = \text{mass}^2$ and $[\tilde{q}] = \text{mass}$, while our conventions are $[G] = [q] = 0$. To relate G_q to $\widetilde{G}_{\tilde{q}}$ we can use the expression $[G_q] = \frac{[\widetilde{G}_{\tilde{q}}]}{M_{\text{pl}}}$.

which is naturally of order

$$\left[\widetilde{G}_{\tilde{q}}\right] = \frac{M_{\rm pl}^2}{m_\Lambda^{(1)} m_\Lambda^{(2)} M_{\rm pl} + \ldots} \left(\text{TeV}^2 + \frac{\text{TeV}}{M_{\rm pl}^2}(m_\Lambda^{(1)} m_\Lambda^{(2)} M_{\rm pl} + \ldots)\right) = \frac{M_{\rm pl}\text{TeV}^2}{m_\Lambda^{(1)} m_\Lambda^{(2)}} + \text{TeV} + \ldots\ , \tag{3.73}$$

where the $\ldots$ are of sub-leading order. We expect that $m_\Lambda^{(1)}$, the constant term of the inflaton sector, is of order $[H] = 10^{-5} M_{\rm pl}$, while $[m_\Lambda^{(2)}] = \text{TeV}$. Hence, translating back to dimensionless units, we find $G_q \sim 10^{-11}$.

Taking the kinetic gauge, i.e. a canonical Kähler metric $G_{\phi\overline{\phi}} = 1$, we can easily find the natural value of α. From (3.69) we see that α depends on ϵ_ϕ and G_ϕ via

$$\alpha \propto \sqrt{\epsilon_\phi} - G_\phi, \tag{3.74}$$

modulo some unknown but negligible phase factors. G_ϕ is of order $\sqrt{3}$ in order to have a potential $V > 0$. Since ϵ_ϕ is of order $\mathcal{O}(10^{-3})$, the value of $|\alpha|$ is of order unity. For a rough estimate for $\eta_{\text{naïve}} \sim 10^{-3}$ we can therefore pinpoint the Standard Model as indicated in figure 3.11. In both cases, $\eta_{\text{naïve}} > 0$ as well as $\eta_{\text{naïve}} < 0$, the lightest supersymmetric particle is too light for the single sector inflationary dynamics to truly describe the full system. Any tuned and working inflationary supergravity model in which the Standard Model is assumed to not take part considerably in the cosmic evolution, requires implicit assumptions on the Standard Model that either the inflaton sector is responsible for Standard Model supersymmetry breaking through gravity mediation or the masses of its scalar multiplets are unnaturally large in terms of the now independent Standard Model supersymmetry breaking scale.

3.7 Clarification of the rigid limit to supersymmetry

Multiple sectors are not only a common feature in supergravity cosmology but also in phenomenology. These sectors are necessary to either incorporate inflation or supersymmetry breaking or are a consequence of string model-building. It is therefore of general importance to understand the restrictions of combining several sectors with their individual actions to one theory. In particular to study inflation, it is desirable to separate the dynamics of all fields that do not contribute to the exponential expansion of the Universe from the inflaton fields that do. Since gravity is the weakest possible interaction, the inflationary sector is assumed to only couple gravitationally to an unknown "hidden" sector that may also break supersymmetry by itself. Whereas it is natural for a rigid supersymmetric theory to be separated into several sectors, the restrictive structure of supergravity forces the different sectors to couple not only non-locally through graviton exchange but also directly. For this reason embedding supersymmetric theories as sectors into a supergravity can be notoriously difficult, see e.g. [243–250].

Though multiple sector supergravities are a long studied subject, the context of cosmology has seriously sharpened the question. In supergravity models of inflation it is commonly noted that one seeks a consistent truncation of the scalar sector. This is necessary but not sufficient. Even with a consistent truncation one may have dominating instabilities towards the naïvely non-dynamical sectors, that can move them away from their supersymmetric critical points. One needs either a symmetry constraint or an energy barrier to constrain the dynamics to the putative inflaton sector.

During inflation, supersymmetry is broken and although it is frugal to consider scenarios where the inflaton sector is also responsible for phenomenological supersymmetry breaking (see e.g. [251–253]), this need not be so. For instance, in a generic gauge-mediation scenario, the mechanism responsible for supersymmetry breaking need not involve the fields that drive inflation. This example immediately shows that the generic cosmological set-up must be able to account for a sector that breaks supersymmetry *independently* of the inflationary dynamics.

This consideration is our starting point. We consider a multiple-sector supergravity that decouples in the strictest sense in the limit $M_{\rm pl} \to \infty$. In this limit the action must then be the sum of two independent functions[6]

$$S[\phi, \overline{\phi}, q, \overline{q}] = S[\phi, \overline{\phi}] + S[q, \overline{q}] \;, \tag{3.75}$$

such that the path integral factorizes. For a globally supersymmetric field theory with a standard kinetic term this can be achieved by demanding that the independent Kähler and superpotentials sum

$$K_{\rm susy}(\phi, \overline{\phi}, q, \overline{q}) = K^{(1)}(\phi, \overline{\phi}) + K^{(2)}(q, \overline{q}) \;, \qquad W_{\rm susy}(\phi, q) = W^{(1)}(\phi) + W^{(2)}(q) \;. \tag{3.76}$$

The issue we address here is that in supergravity complete decoupling in the sense of (3.75) appears to be impossible, even in principle. Even with block diagonal kinetic terms from a sum of Kähler potentials, the more complicated form of the supergravity potential

$$V_{\rm sugra} = {\rm e}^{K/M_{\rm pl}^2} \left(K^{a\overline{b}} {\rm D}_a W \overline{{\rm D}_b W} - \frac{3|W|^2}{M_{\rm pl}^2} \right) \;, \;\; {\rm D}W_{\rm sugra} = \partial W_{\rm sugra} + \partial K_{\rm sugra} \frac{W_{\rm sugra}}{M_{\rm pl}^2} \;, \tag{3.77}$$

implies that there are many *direct* couplings between the two sectors. It raises the immediate question: if the low-energy $M_{\rm pl} \to \infty$ globally supersymmetric model must consist of decoupled sectors, what is the relation between $K_{\rm sugra}, W_{\rm sugra}$ and $K_{\rm susy}, W_{\rm susy}$, or vice versa given a globally supersymmetric model described by $K_{\rm susy}, W_{\rm susy}$, what is the best choice for $K_{\rm sugra}, W_{\rm sugra}$ such that the original theory can be recovered in the limit $M_{\rm pl} \to \infty$?

In this section we shall show that the scaling implied by the explicit factors of $M_{\rm pl}$ in the supergravity potential (3.77) is an incomplete answer to this question. The direct communication between the sectors, controlled by $M_{\rm pl}$, has serious consequences for both the ground state

[6] As example we consider the simplest case, a model with two uncharged scalar supermultiplets $X^a = (\phi, q)$ that are singlets under all symmetries. Gauge interactions and global symmetries will not change this general argument provided the two sectors are not mixed by symmetries or coupled by gauge fields. Therefore, we will also ignore D-terms in the supergravity potential below.

structure (solutions to the equation of motion, i.e. the cosmological dynamics) and the interactions between the two sectors. To be explicit, the first guess at how the rigid supersymmetry and supergravity Kähler potentials and superpotentials are related

$$K_{\rm sugra}(\phi,\overline{\phi},q,\overline{q}) = K^{(1)}_{\rm susy}(\phi,\overline{\phi}) + K^{(2)}_{\rm susy}(q,\overline{q}) + \dots\ , \quad W_{\rm sugra}(\phi,q) = W^{(1)}_{\rm susy}(\phi) + W^{(2)}_{\rm susy}(q) + \dots\ , \tag{3.78}$$

with ... indicating Planck-suppressed terms and possibly a constant term, suffers from the drawback that the ground states of the full theory are no longer the product of the ground states of the individual sectors, except when both (rather than only one) ground states are supersymmetric [199, 200] (see also [201, 202, 213]). This directly follows from considering the extrema of the supergravity potential[7]

$$\nabla_i V = \frac{\mathrm{D}_i W}{W} V + \mathrm{e}^{K/M_{\rm pl}^2}|W|^2 \left(\nabla_i \left(\frac{\mathrm{D}_j W}{W}\right) \frac{\mathrm{D}^j \overline{W}}{\overline{W}} + \frac{1}{M_{\rm pl}^2}\frac{\mathrm{D}_i W}{W} + \nabla_i \left(\frac{\mathrm{D}_\beta W}{W}\right) \frac{\mathrm{D}^\beta \overline{W}}{\overline{W}} \right)\ , \tag{3.79}$$

$$\begin{aligned}\nabla_i \nabla_\alpha V = {} & \frac{\mathrm{D}_\alpha W}{W} \nabla_i V + \frac{\mathrm{D}_i W}{W} \nabla_\alpha V - \frac{\mathrm{D}_i W}{W}\frac{\mathrm{D}_\alpha W}{W} V + \mathrm{D}_i \left(\frac{\mathrm{D}_\alpha W}{W}\right)(V + \frac{2}{M_{\rm pl}^2}\mathrm{e}^{K/M_{\rm pl}^2}|W|^2) \\ & + \mathrm{e}^{K/M_{\rm pl}^2}|W|^2 \left(\nabla_i \nabla_\alpha \left(\frac{\mathrm{D}_\beta W}{W}\right)\frac{\mathrm{D}^\beta \overline{W}}{\overline{W}} + \nabla_\alpha \nabla_i \left(\frac{\mathrm{D}_j W}{W}\right)\frac{\mathrm{D}^j \overline{W}}{\overline{W}}\right)\ . \end{aligned} \tag{3.80}$$

Supersymmetric ground states, for which the covariant derivatives of W vanish on the solution, $\mathrm{D}_i W = 0$ and $\mathrm{D}_\alpha W = 0$, are still product solutions. But for Kähler - and superpotentials that sum (3.78), even if only one sector is in a non-supersymmetric ground state, by which we mean $\mathrm{D}_i W = 0$, $\mathrm{D}_\alpha W \neq 0$, we can neither conclude that sector 2, labeled by i, is in a minimum, for which $\nabla_i V$ would vanish, nor that the condition for sector 1, labeled by α, to be in a local ground state is independent of the sector 2 fields q^i, which would mean that $\nabla_i \nabla_\alpha V = 0$. The former is only true when

$$\nabla_i \left(\frac{\mathrm{D}_\beta W}{W}\right) \frac{\mathrm{D}^\beta \overline{W}}{\overline{W}} = 0\ . \tag{3.81}$$

The second requires, in addition,

$$\nabla_i \nabla_\alpha \left(\frac{\mathrm{D}_\beta W}{W}\right) \frac{\mathrm{D}^\beta \overline{W}}{\overline{W}} + \nabla_\alpha \nabla_i \left(\frac{\mathrm{D}_j W}{W}\right) \frac{\mathrm{D}^j \overline{W}}{\overline{W}} = 0\ , \tag{3.82}$$

and also sharpens the first condition (3.81) to[8]

$$\mathrm{D}_i \frac{\mathrm{D}_\alpha W}{W} = 0\ . \tag{3.83}$$

[7]To derive (3.80) note that, since $\mathrm{D}W/W$ is Kähler invariant and since the Levi-Civita connection ∇ of the field space manifold does not get cross-contributions in a product manifold,

$$\nabla_i \frac{\mathrm{D}_\alpha W}{W} = \partial_i \frac{\mathrm{D}_\alpha W}{W} = \mathrm{D}_i \frac{\mathrm{D}_\alpha W}{W}.$$

[8]These conditions are merely sufficient not necessary. However, it is clear that the restrictive nature of supergravity enforces conditions on the unknown sectors for the system to be separate.

Equations (3.81–3.83) are conditions for decoupling which apply not only to the ground state of the full system but also to other critical points of the potential, for instance along an inflationary valley. Generically these conditions are not met on the solution (the second derivative need not vanish at an extremum; recall that $\mathrm{D}_i W$ does not vanish identically but only on the solution). Hence, generically the ground states of hidden sectors mix and this spoils many cosmological supergravity scenarios that truncate the action to one or the other sector (see e.g. [254] and references therein). It is this issue that is particularly relevant for inflationary model building, where a very weak coupling between the inflaton sector and all other sectors has to persist over an entire *trajectory* in field space where the expectation values of the fields are changing with time (see e.g. [1, 193, 197, 203, 214]). At the same time, one is interested in the generic situation in which *both* sectors may contribute to supersymmetry breaking.[9]

3.8 Natural multi-sector supergravities

There is a well-known natural way to construct supergravity potentials for which the ground states (and critical points) do separate better. This obvious combination of superpotentials automatically satisfies (3.81–3.83) and hence does ensure that if one of the ground states is supersymmetric, the ground state of the other sector is a decoupled field theory ground state whether it breaks supersymmetry or not. This is if we choose a product of superpotentials, keeping the sum of Kähler potentials as before,

$$K_{\text{sugra}}(\phi,\overline{\phi},q,\overline{q}) = K^{(1)}_{\text{sugra}}(\phi,\overline{\phi}) + K^{(2)}_{\text{sugra}}(q,\overline{q}) \ , \quad W_{\text{sugra}}(\phi,q) = \frac{1}{M_{\text{pl}}^3} W^{(1)}_{\text{sugra}}(\phi) W^{(2)}_{\text{sugra}}(q) \ . \tag{3.84}$$

This is well-known [195, 196, 257] and has recently been emphasized in the context of cosmology [1, 201–203, 213, 214, 254, 258, 259]. This ansatz conforms to the more natural description

[9] This situation has to be contrasted to phenomenological models appropriate for studying gravity mediated supersymmetry breaking, such as an ansatz [255]

$$\begin{aligned} K(\phi,\overline{\phi},q,\overline{q}) &= K^{(0)}(\phi,\overline{\phi}) + q^a \overline{q}^{\overline{b}} K^{(1)}_{a\overline{b}}(\phi,\overline{\phi}) + (q^2+\overline{q}^2) K^{(2)}(\phi,\overline{\phi}) \ , \\ W(\phi,q) &= W_0(\phi) + q^i q^j W_{ij}(\phi) \ . \end{aligned}$$

or equivalently, if $W \neq 0$,

$$G(\phi,\overline{\phi},q,\overline{q}) = G^{(0)}(\phi,\overline{\phi}) + q^i \overline{q}^{\overline{j}} G^{(1,1)}_{i\overline{j}}(\phi,\overline{\phi}) + q^i q^j G^{(2,0)}_{ij}(\phi,\overline{\phi}) + \overline{q}^{\overline{i}} \overline{q}^{\overline{j}} G^{(0,2)}_{\overline{i}\overline{j}}(\phi,\overline{\phi}) + \dots \ .$$

In models like these, it is understood that $\dot{q} = 0$ and the q-sector can remain in its supersymmetric critical point throughout the evolution of the supersymmetry breaking fields. For inflation, such an expectation is unrealistic, as the supersymmetry-preserving sector can become unstable during the inflationary dynamics, see e.g. a recent discussion of the case in which the inflaton field ϕ is solely responsible for supersymmetry breaking during inflation ([253] and references therein). In this relatively simple case, and except for very fine-tuned situations, the generic scenario appears to be that one or more of the q-fields are destabilized somewhere along the inflationary trajectory and they trigger an exit from inflation (in other words, they become "waterfall" fields, and inflation is of the hybrid kind [256]). This implies that the pattern of supersymmetry breaking today is not related to the one during inflation, and also, since the waterfall fields are forced away from their supersymmetric critical points, that supersymmetry is broken by both sectors as the Universe evolves towards the current vacuum.

of supergravities in terms of the Kähler invariant function

$$G(X,\overline{X}) = \frac{1}{M_{\rm pl}^2} K_{\rm sugra}(X,\overline{X}) + \log\left(\frac{W_{\rm sugra}(X)}{M_{\rm pl}^3}\right) + \log\left(\frac{\overline{W}_{\rm sugra}(\overline{X})}{M_{\rm pl}^3}\right) , \tag{3.85}$$

which can be defined if W is non-zero in the region of interest.[10] This Kähler function in turn underlies a better description of multiple sectors in supergravity where G is a sum of independent functions, as we have already been advocating in equation (3.26)

$$G(\phi,\overline{\phi},q,\overline{q}) = G^{(1)}(\phi,\overline{\phi}) + G^{(2)}(q,\overline{q}) , \tag{3.86}$$

such that the two sectors are separately Kähler invariant. The sum implies the product superpotential put forward above. This is the simplest ansatz that still allows some degree of calculational control when both sectors break supersymmetry —as well as optimizing decoupling along the inflationary trajectory. One of the simplest models of hybrid inflation in supergravity, F-term inflation [260, 261], is in this class.

3.9 DECOUPLING

Given that we have just argued that a product of superpotentials is a more natural framework to discuss multiple sector supergravities, the obvious question arises how to recover a decoupled *sum* of potentials for a globally supersymmetric theory in the limit where gravity decouples, i.e. in which

$$V_{\rm sugra} = {\rm e}^{K/M_{\rm pl}^2}\left(|{\rm D}W|^2 - \frac{3|W|^2}{M_{\rm pl}^2}\right) \;\;\rightarrow\;\; V_{\rm susy} = \sum_j |\partial_j W^{(j)}|^2 . \tag{3.87}$$

For a two-sector supergravity defined by equations (3.84) one would not find this answer, if one takes the standard decoupling limit $M_{\rm pl}\to\infty$ with both $K = K^{(1)} + K^{(2)}$ and $W = M_{\rm pl}^{-3} W^{(1)} W^{(2)}$ fixed.[11] Instead, the product structure of the superpotential introduces a cross-coupling between sectors,

$$V_{\rm eff} = \frac{1}{M_{\rm pl}^3}\left(|W^{(2)}|^2|\partial_\alpha W^{(1)}|^2 + |W^{(1)}|^2|\partial_i W^{(2)}|^2\right) \neq V_{\rm susy} , \tag{3.88}$$

[10] We expect this condition to hold around a supersymmetry breaking vacuum with almost vanishing cosmological constant. It also holds in many models of supergravity inflation, although a notable exception is [191, 215].

[11] Strictly speaking the decoupling limit sends $M_{\rm pl}\to\infty$ while keeping the fields ϕ, q fixed with $W^{(j)}/M_{\rm pl}^3$ a holomorphic function of $\phi/M_{\rm pl}$ or $q/M_{\rm pl}$ and $K^{(j)}/M_{\rm pl}^2$ a real function of $\phi/M_{\rm pl}, \overline{\phi}/M_{\rm pl}$ or $q/M_{\rm pl}, \overline{q}/M_{\rm pl}$. The limit zooms in to the origin so K must be assumed to be non-singular there. Formally the decoupling limit does not exist otherwise. Physically it means that one is taking the decoupling limit w.r.t. an a priori determined ground state, around which K and W are expanded. If K is non-singular at the origin, the overall factor ${\rm e}^{K/M_{\rm pl}^2}$ yields an overall constant as $M_{\rm pl}\to\infty$, which may be set to unity, i.e. the constant part of K vanishes. In the decoupling limit, both K and W may then be written as polynomials. Letting the coefficients in W and K scale as their canonical scaling dimension such that W has mass dimension three and K has mass dimension two, then gives the rule of thumb that both K and W are held fixed as $M_{\rm pl}\to\infty$

whose behavior under the limit $M_{\rm pl} \to \infty$ is best examined at the level of the superpotential.

Supergravity is sensitive to the expectation value $W_0 = \langle W \rangle$ of W, which relates the scale of supersymmetry breaking to the expectation value of the potential, i.e. the cosmological constant

$$\Lambda^2 M_{\rm pl}^2 = \langle V \rangle \sim \langle {\rm D}W^2 \rangle - \frac{3}{M_{\rm pl}^2}\langle W^2 \rangle = m_{\rm susy}^4 - 3\frac{W_0^2}{M_{\rm pl}^2} \ . \tag{3.89}$$

The vacuum expectation value cannot vanish in a supersymmetry breaking vacuum with (nearly) zero cosmological constant, such as our Universe. Therefore, in the following we assume $\langle W \rangle \neq 0$ in the region of interest. Instead of the usual way to incorporate it, $W_{\rm sugra} = W_0 + W_{\rm dyn}$ with $W_{\rm dyn} = W_{\rm susy} + \ldots$, we include the vacuum expectation value for a two-sector product superpotential by writing

$$\begin{aligned} W(\phi, q) &= \frac{1}{M_{\rm pl}^3} W^{(1)} W^{(2)} = \frac{1}{M_{\rm pl}^3}\left(W_0^{(1)} + W_{\rm dyn}^{(1)}(\phi)\right)\left(W_0^{(2)} + W_{\rm dyn}^{(2)}(q)\right) \\ &= \frac{1}{M_{\rm pl}^3}\left(W_0^{(1)} W_0^{(2)} + W_0^{(2)} W_{\rm dyn}^{(1)}(\phi) + W_0^{(1)} W_{\rm dyn}^{(2)}(q) + W_{\rm dyn}^{(1)}(\phi) W_{\rm dyn}^{(2)}(q)\right) . \end{aligned} \tag{3.90}$$

This is physically equivalent to a sum of superpotentials except for the last term. Note again, that if one uses the standard scaling, $\frac{\phi}{M_{\rm pl}} \to 0$; $\frac{q}{M_{\rm pl}} \to 0$ with all couplings in $W^{(\rm total)}$ having the canonical scaling dimensions, this last term contains renormalizable couplings involving the scalar partner of the goldstino, and these are not Planck-suppressed: if supersymmetry is broken by the ϕ sector, terms of the form ϕq^2 are renormalizable and would survive the $M_{\rm pl} \to \infty$ limit, leading to a direct coupling between the two sectors.[12] If both sectors break supersymmetry then mass-mixing terms ϕq also survive. All such (relevant) terms are of course absent if none of the two sectors break supersymmetry, but this is not the case we are interested in. One would have expected that these cross-couplings naturally vanish in the decoupling limit.

The point of this note is simply to remark that the realization that each of the superpotentials $W^{(j)} = W_0^{(j)} + W_{\rm dyn}^{(j)}$ contains a constant term can resolve this conundrum by assuming a non-standard scaling for the constituent parts $W_0^{(j)}$, $W_{\rm dyn}^{(j)}$. To achieve a decoupling we need that the cross term $W_{\rm dyn}^{(1)} W_{\rm dyn}^{(2)}$, which contains the coupling between the two sectors, scales away in the limit $M_{\rm pl} \to \infty$. As a result the first term in (3.90) has to diverge, because its product with the cross term should remain finite. In particular we can choose an overall scaling

$$W = \frac{1}{M_{\rm pl}^3}(\underbrace{W_0^{(1)} W_0^{(2)}}_{\sim M_{\rm pl}^{3+r}} + \underbrace{W_0^{(1)} W_{\rm dyn}^{(2)}}_{\sim M_{\rm pl}^{3}} + \underbrace{W_0^{(2)} W_{\rm dyn}^{(1)}}_{\sim M_{\rm pl}^{3}} + \underbrace{W_{\rm dyn}^{(1)} W_{\rm dyn}^{(2)}}_{\sim M_{\rm pl}^{3-r}}) \ , \tag{3.91}$$

[12] For a product of superpotentials we can always choose a Kähler gauge *at every point* with $\langle K \rangle = \langle \partial_\phi K \rangle = \langle \partial_z K \rangle = 0$ without mixing the superpotentials. In that case F-term supersymmetry breaking is given by the linear terms in the expansion of $W^{(1)}$ and $W^{(2)}$: $\langle {\rm D}_\phi W \rangle \sim \langle \partial_\phi W^{(1)} \rangle$, $\langle {\rm D}_z W \rangle \sim \langle \partial_z W^{(2)} \rangle$.

with $r > 0$. Let us account for dimensions by introducing an extra scale m_Λ such that

$$W_0^{(1)} = m_\Lambda^{\frac{3-r}{2}-A} M_{\rm pl}^{\frac{3+r}{2}+A}\,, \qquad W_{\rm dyn}^{(1)} = M_{\rm pl}^3 \frac{W_{\rm susy}^{(1)}}{W_0^{(2)}}\,,$$
$$W_0^{(2)} = m_\Lambda^{\frac{3-r}{2}+A} M_{\rm pl}^{\frac{3+r}{2}-A}\,, \qquad W_{\rm dyn}^{(2)} = M_{\rm pl}^3 \frac{W_{\rm susy}^{(2)}}{W_0^{(1)}}\,, \tag{3.92}$$

with $W_{\rm susy}^{(j)}$ fixed as $M_{\rm pl} \to \infty$. Formally one can choose an inhomogeneous scaling with $A \neq 0$, but as we shall see it has no real consequences. For any A it is easily seen that with this scaling,

$$\begin{aligned} {\rm D}_\alpha W &= \partial_\alpha W_{\rm susy}^{(1)} + \frac{m_\Lambda^{r-3}}{M_{\rm pl}^r} W_{\rm susy}^{(2)} \partial_\alpha W_{\rm susy}^{(1)} \\ &\quad + \frac{\partial_\alpha K^{(1)}}{M_{\rm pl}^2}\left(m_\Lambda^{3-r} M_{\rm pl}^r + W_{\rm susy}^{(1)} + W_{\rm susy}^{(2)} + \frac{m_\Lambda^{r-3}}{M_{\rm pl}^r} W_{\rm susy}^{(1)} W_{\rm susy}^{(2)}\right) \to \partial_\alpha W_{\rm susy}^{(1)}\,, \end{aligned} \tag{3.93}$$

in the limit $M_{\rm pl} \to \infty$ if and only if $0 < r < 2$ and thus

$$V_{\rm sugra} = {\rm e}^{K/M_{\rm pl}^2}\left(|{\rm D}W|^2 - \frac{3|W|^2}{M_{\rm pl}^2}\right) \to \sum_j |\partial_j W_{\rm susy}^{(j)}|^2 - 3 m_\Lambda^{2(3-r)} M_{\rm pl}^{2(r-1)} + \mathcal{O}\left(\frac{1}{M_{\rm pl}}\right)\,. \tag{3.94}$$

For $r < 1$ the manifestly constant term in the potential vanishes as well and we recover the strict decoupled field theory result, with the gravitino mass going to zero as $m_{3/2} = \langle W \rangle M_{\rm pl}^{-2} = m_\Lambda^{3-r} M_{\rm pl}^{r-2} = \frac{m_{\rm susy}^2}{\sqrt{3} M_{\rm pl}}$. We see that the gravitino mass is independent of r in physical scales.

The parameter r should not be larger than unity for the new decoupling limit to be well defined. For the special case $r = 1$ [195], the potential has an additional overall "cosmological" constant. For a generic non-gravitational field theory in which $M_{\rm pl} \to \infty$ this is just an overall shift of the potential, which we can arbitrarily remove since it does not change the physics. Nevertheless from a formal point of view, we know that absolute ground state energy of a globally supersymmetric theory equals zero, as a result of the supersymmetry algebra $\{Q, Q\} = H$. For this reason it is more natural to restrict the value of r to the range $0 < r < 1$.

Finally, the novel scaling in (3.92) can be readily generalized to an arbitrary number of sectors. For s sectors, writing $W^{(j)} = W_0^{(j)} + W_{\rm dyn}^{(j)}$,

$$\begin{aligned} W &= \frac{1}{M_{\rm pl}^{3(s-1)}} \prod_{j=1}^s W^{(j)} \\ &= \frac{1}{M_{\rm pl}^{3(s-1)}}\left[\prod_{j=1}^s W_0^{(j)} + \sum_{k=1}^s \left(W_{\rm dyn}^{(k)} \prod_{j\neq k}^s W_0^{(j)}\right) + \sum_{l>k}^s \left(W_{\rm dyn}^{(k)} W_{\rm dyn}^{(l)} \prod_{j\neq k,l}^s W_0^{(j)}\right) + \dots \right]\,. \end{aligned} \tag{3.95}$$

We want the last and all further terms to scale away as $M_{\rm pl}^{-r}$ and higher with $r > 0$, while the second term(s) should be constant. As a consequence the first term will scale as $M_{\rm pl}^r$. Assuming a scaling that is homogeneous across sectors, this implies

$$W_0^{(j)} \sim M_{\rm pl}^{\frac{3(s-1)+r}{s}}\,, \qquad W_{\rm dyn}^{(j)} \sim M_{\rm pl}^{\frac{(3-r)(s-1)}{s}}\,, \tag{3.96}$$

for each of the $j \in \{1, \ldots, s\}$. With this scaling, a general term consisting of l dynamical superpotentials and $s - l$ constant parts, scales as

$$\frac{W_{\rm dyn}^{l} W_0^{s-l}}{M_{\rm pl}^{3(s-1)}} \sim M_{\rm pl}^{r(1-l)} \,, \tag{3.97}$$

and as constructed any term containing dynamical interactions between sectors, $l > 2$, is Planck-suppressed. To ensure a vanishing constant term as in eq. (3.94), r is again limited to the range $0 < r < 1$.

Let us conclude with a comment on the physical meaning behind the scaling (3.92). It may appear that we have changed the canonical RG-scaling of the theory. This is not quite true. For the interacting terms in the potential, it is the coefficients in the product $W_0^{(2)} W_{\rm dyn}^{(1)} = W_{\rm susy}^{(1)}$ that ought to obey canonical RG-scaling. This precisely corresponds to holding $W_{\rm susy}^{(j)}$ fixed as $M_{\rm pl} \to \infty$ (see footnote *11*). On the other hand, the scaling of the constant term in the potential has changed from its canonical value. However, this is very natural in a supersymmetric theory. The constant term, $\prod_j W_0^{(j)}$, equals the ground state energy. Precisely supersymmetric theories can "naturally" explain non-canonical scaling of the cosmological constant (at the loop level; the scaling of the bare ground state energy can be different in every model). A non-integer power is strange but $r = 1$ is certainly a viable option in a supersymmetry-breaking ground state: it is the natural scaling in theories with higher supersymmetry [262] when combined with a subleading $\log(M_{\rm pl}/m_{\rm susy})$ breaking. Our engineering analysis only focuses on power-law scaling and these can always have subleading logarithms. ($r = 2$ would correspond to the cosmological constant for a spontaneously broken $\mathcal{N} = 1$ theory due to mass splitting).

3.10 Conclusions

In this chapter we have studied the effect of hidden sectors on the fine-tuning of F-term inflation in supergravity, identifying a number of issues in the current methodology of fine-tuning inflation in supergravity. Fine-tuning inflationary models is only valid when the neglected physics does not affect this fine-tuning, in which case the inflationary physics can be studied independently. As shown in figures 3.8 and 3.9 this assumption holds only under very special circumstances. The reason is that the everpresent gravitational couplings will always lead to a mixing of the hidden sectors with the inflationary sector.

First, we have argued in which way the action represents the two sectors as minimally coupled as possible. Rather than adding the superpotentials, the correct action is obtained by adding the Kähler functions, which preserves Kähler invariance in each sector independently. This leads to the action (3.31), where the superpotentials are multiplied.

Although this ansatz is extremely useful in the context of cosmology, it demands due dilligence in a number of aspects. We have argued in section 3.9 that a (cosmological) constant term must be included to prevent the superpotentials from vanishing and rendering the Kähler function infinite on the solution. This constant and its cross-terms call for a second scale when taking the limit in which the Planck mass goes to infinity and gravity is turned off. This limit is delicate and restricts the scaling of the individual terms in the superpotential.

For a hidden sector vacuum that preserves supersymmetry, the sectors decouple consistently [198–202]. However, for a supersymmetry breaking vacuum the inflationary dynamics is generically altered, where the nature and the size of the change depends on the scale of supersymmetry breaking. For a hidden sector with a low scale of supersymmetry breaking, like the Standard Model, the cross coupling scales with the scale of supersymmetry breaking, and is therefore typically small. Yet, as shown in section 3.4.3, also the lightest mass of the hidden sector scales with the scale of supersymmetry breaking within that sector. This light mode is strongly affected by the inflationary physics and thus evolves during inflation. Therefore, any single field analysis is completely spoiled as discussed in section 3.6.3.

For massive hidden sectors, the problem is more traditional. For a small hidden sector supersymmetry breaking scale, one has a conventional decoupling as long as the lightest mass of the hidden sector is much larger than the inflaton mass. However, for large hidden sector supersymmetry breaking, this intuition fails. Then, the off-diagonal terms in the mass matrix (3.47) will lead to a large correction of the η-parameter.

To conclude, any theory that is working by only tuning the inflaton sector has made severe hidden assumptions about the hidden sector, which typically will not be easily met. Methodologically the only sensible approach is to search for inflation in a full theory, including knowledge of all hidden sectors.

Appendix D

Some supergravity relations

For easy reference to the reader, we use this appendix to state the relevant derivatives of the supergravity potential of a two-sector system coupled via

$$G(\phi^i, \overline{\phi}^{\bar{\imath}}, q^a, \overline{q}^{\overline{a}}) = G^{(1)}(\phi^i, \overline{\phi}^{\bar{\imath}}) + G^{(2)}(q^a, \overline{q}^{\overline{a}}) \ . \tag{D.1}$$

We use middle-alphabet Latin indices $\{i, \bar{\imath}\}$ to denote the fields in the inflationary sector, beginning-alphabet Latin indices $\{a, \overline{a}\}$ to denote the fields in the hidden sector and Greek indices $\{\alpha, \overline{\alpha}\}$ to denote the full system. Derivatives with respect to these fields are denoted by subscripts, e.g. $\partial_i G = G_i$ and $\partial_i \partial_j G = G_{ij}$. The Hessian $G_{\alpha\overline{\beta}}$ describes the metric of the (product-) manifold parametrized by the fields. This is a Kähler manifold and hence $\nabla_\alpha G_{\overline{\beta}} = G_{\alpha\overline{\beta}}$.

The supergravity potential is

$$V = e^G (G_\alpha G^\alpha - 3) = e^G (G_{\overline{\alpha}} G^{\overline{\alpha}} - 3) = e^G (G_a G^a + G_i G^i - 3) \ . \tag{D.2}$$

Its covariant derivatives are denoted with subscripts (note that this is a different convention than the one used for the Kähler function G), e.g. $\nabla_i V = \partial_i V = V_i$ and $\nabla_i \nabla_j V = V_{ij}$. In terms of derivatives of G, the first derivatives of V are given by

$$V_i = G_i V + e^G \left((\nabla_i G_j) G^j + G_i \right) \ , \tag{D.3}$$

$$V_{\bar{\imath}} = G_{\bar{\imath}} V + e^G \left((\nabla_{\bar{\imath}} G_{\bar{\jmath}}) G^{\bar{\jmath}} + G_{\bar{\imath}} \right) \ , \tag{D.4}$$

and similar expressions for V_a and $V_{\overline{a}}$. The Hessian of covariant derivatives is

$$V_{ij} = \nabla_i G_j V + G_i V_j + G_j V_i - G_i G_j V + e^G \left[(\nabla_i \nabla_j G_k) G^k + 2\nabla_i G_j \right] \ , \tag{D.5}$$

$$V_{i\overline{\jmath}} = G_{i\overline{\jmath}} V + G_i V_{\overline{\jmath}} + G_{\overline{\jmath}} V_i - G_i G_{\overline{\jmath}} V + e^G \left[R_{i\overline{\jmath}k\overline{l}} G^k G^{\overline{l}} + G^{k\overline{l}} \nabla_i G_k \nabla_{\overline{\jmath}} G_{\overline{l}} + G_{i\overline{\jmath}} \right] \ , \tag{D.6}$$

$$\begin{aligned} V_{ia} &= \nabla_a G_i V + G_i V_a + G_a V_i - G_i G_a V + e^G \left[(\nabla_a \nabla_i G_\alpha) G^\alpha + \nabla_i G_a + \nabla_a G_i \right] \\ &= G_i V_a + G_a V_i - G_i G_a V \ , \end{aligned} \tag{D.7}$$

$$\begin{aligned} V_{i\overline{a}} &= G_{i\overline{a}} V + G_i V_{\overline{a}} + G_{\overline{a}} V_i - G_i G_{\overline{a}} V + e^G \left[R_{\alpha\overline{\beta}i\overline{a}} G^\alpha G^{\overline{\beta}} + G^{\alpha\overline{\beta}} \nabla_i G_\alpha \nabla_{\overline{a}} G_{\overline{\beta}} + G_{i\overline{a}} \right] \\ &= G_i V_{\overline{a}} + G_{\overline{a}} V_i - G_i G_{\overline{a}} V \ , \end{aligned} \tag{D.8}$$

and similar expressions for the other $V_{\alpha\beta}$. The equalities in (D.7) and (D.8) are a result of the specific form of the Kähler function (D.1).

Appendix E

Mass eigenmodes in a stabilized sector

In this appendix we provide some intermediate results in the calculation of (3.38–3.39). Using the expressions as stated in appendix D, to first order in $|G_q|$, the second derivatives of the potential are given by

$$V_{qq} = e^G \left[(2 + e^{-G}V)\nabla_q G_q + (\nabla_q \nabla_q G_q)G^q\right] + \mathcal{O}(|G_q|^2) , \tag{E.1}$$

$$V_{q\overline{q}} = e^G \left[G_{q\overline{q}}(1 + e^{-G}V) + G^{q\overline{q}}(\nabla_q G_q)(\nabla_{\overline{q}} G_{\overline{q}})\right] + \mathcal{O}(|G_q|^2) . \tag{E.2}$$

Using the supersymmetry breaking restriction (3.35) in (E.1) and (E.2), we find

$$V_{qq} = -e^G G_{q\overline{q}} \left[(2 + e^{-G}V)(1 + e^{-G}V)\widehat{G^q}^{-2} - G^{q\overline{q}}(\nabla_q \nabla_q G_q)G^q\right] + \mathcal{O}(|G_q|^2) , \tag{E.3}$$

$$\begin{aligned} V_{q\overline{q}} &= e^G \left[G_{q\overline{q}}(1 + e^{-G}V) + (1 + e^{-G}V)^2 G^{q\overline{q}} G_{q\overline{q}} G_{q\overline{q}}\right] + \mathcal{O}(|G_q|^2) \\ &= e^G G_{q\overline{q}}(2 + e^{-G}V)(1 + e^{-G}V) + \mathcal{O}(|G_q|^2) , \end{aligned} \tag{E.4}$$

and hence

$$\begin{aligned} |V_{qq}| &= e^G G_{q\overline{q}}(2 + e^{-G}V)(1 + e^{-G}V) \times \\ &\quad \times \sqrt{1 - \frac{2G^{q\overline{q}}\mathrm{Re}\left\{(\nabla_q \nabla_q G_q)G^q \widehat{G^q}^{-2}\right\}}{(2 + e^{-G}V)(1 + e^{-G}V)} + \frac{|G^{q\overline{q}}(\nabla_q \nabla_q G_q)G^q|^2}{(2 + e^{-G}V)^2(1 + e^{-G}V)^2}} + \mathcal{O}(|G_q|^2) \\ &= e^G G_{q\overline{q}} \left[(2 + e^{-G}V)(1 + e^{-G}V) - G^{q\overline{q}}\mathrm{Re}\left\{(\nabla_q \nabla_q G_q)\widehat{G^q}^3\right\}|G^q|\right] + \mathcal{O}(|G_q|^2) . \end{aligned} \tag{E.5}$$

Then (3.37) is evaluated to be

$$m_q^- = e^G G^{q\overline{q}}\mathrm{Re}\left\{(\nabla_q \nabla_q G_q)\widehat{G^q}^3\right\}|G^q| + \mathcal{O}(|G_q|^2) , \tag{E.6}$$

$$m_q^+ = e^G \left[2(2 + e^{-G}V)(1 + e^{-G}V) - G^{q\overline{q}}\mathrm{Re}\left\{(\nabla_q \nabla_q G_q)\widehat{G^q}^3\right\}|G^q|\right] + \mathcal{O}(|G_q|^2) . \tag{E.7}$$

Bibliography

[1] S. Hardeman, J. M. Oberreuter, G. A. Palma, K. Schalm, and T. van der Aalst, *The everpresent eta-problem: knowledge of all hidden sectors required*, *JHEP* **1104** (2011) 009, [arXiv:1012.5966].

[2] A. Achucarro, S. Hardeman, J. M. Oberreuter, K. Schalm, and T. van der Aalst, *Decoupling limits in multi-sector supergravities*, *JCAP* **1303** (2013) 038, [arXiv:1108.2278].

[3] J. M. Oberreuter, K. Schalm, and J. P. van der Schaar, *Finite N effects in holographic singularity resolution*, *to appear* (2013).

[4] G. Lopes Cardoso, A. Ceresole, G. Dall'Agata, J. M. Oberreuter, and J. Perz, *First-order flow equations for extremal black holes in very special geometry*, *JHEP* **10** (2007) 063, [arXiv:0706.3373].

[5] G. L. Cardoso, J. M. Oberreuter, and J. Perz, *Entropy function for rotating extremal black holes in very special geometry*, *JHEP* **05** (2007) 025, [hep-th/0701176].

[6] R. Kunze, *Von der inspiration,* in *eines jeden einziges leben: gedichte*. S. Fischer Verlag, Frankfurt/Main, 1986. reprinted with friendly permission of the publisher.

[7] R. Bluhm, V. Kostelecky, and N. Russell, *CPT and Lorentz tests in hydrogen and anti-hydrogen*, *Phys.Rev.Lett.* **82** (1999) 2254–2257, [hep-ph/9810269].

[8] D. Kapner, T. Cook, E. Adelberger, J. Gundlach, B. R. Heckel, *et. al.*, *Tests of the gravitational inverse-square law below the dark-energy length scale*, *Phys.Rev.Lett.* **98** (2007) 021101, [hep-ph/0611184].

[9] K. Wilson and J. B. Kogut, *The Renormalization group and the epsilon expansion*, *Phys.Rept.* **12** (1974) 75–200.

[10] J. Callan, Curtis G., *Broken scale invariance in scalar field theory*, *Phys.Rev.* **D2** (1970) 1541–1547.

[11] K. Symanzik, *Small distance behavior in field theory and power counting*, *Commun.Math.Phys.* **18** (1970) 227–246.

[12] A. Shomer, *A Pedagogical explanation for the non-renormalizability of gravity*, arXiv:0709.3555.

[13] S. Weinberg and E. Witten, *Limits on Massless Particles*, *Phys.Lett.* **B96** (1980) 59.

[14] J. D. Bekenstein, *Black holes and entropy*, *Phys.Rev.* **D7** (1973) 2333–2346.

[15] S. Hawking, *Particle Creation by Black Holes*, *Commun.Math.Phys.* **43** (1975) 199–220.

[16] S. Hawking, *Black hole explosions*, *Nature* **248** (1974) 30–31.

[17] G. 't Hooft, *A Planar Diagram Theory for Strong Interactions*, *Nucl.Phys.* **B72** (1974) 461.

[18] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, *Large N field theories, string theory and gravity*, *Phys.Rept.* **323** (2000) 183–386, [hep-th/9905111].

[19] B. R. Greene, *The elegant universe: Superstrings, hidden dimensions, and the quest of the ultimate theory*. Norton, New York (USA), 1999.

[20] J. M. Maldacena, *The large N limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231–252, [hep-th/9711200].

[21] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253–291, [hep-th/9802150].

[22] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Gauge theory correlators from non-critical string theory*, *Phys. Lett.* **B428** (1998) 105–114, [hep-th/9802109].

[23] E. T. Akhmedov, *A Remark on the AdS / CFT correspondence and the renormalization group flow*, *Phys.Lett.* **B442** (1998) 152–158, [hep-th/9806217].

[24] E. Alvarez and C. Gomez, *Geometric holography, the renormalization group and the c theorem*, *Nucl.Phys.* **B541** (1999) 441–460, [hep-th/9807226].

[25] J. de Boer, E. P. Verlinde, and H. L. Verlinde, *On the holographic renormalization group*, *JHEP* **0008** (2000) 003, [hep-th/9912012].

[26] E. P. Verlinde and H. L. Verlinde, *RG flow, gravity and the cosmological constant*, *JHEP* **0005** (2000) 034, [hep-th/9912018].

[27] T. Faulkner, N. Iqbal, H. Liu, J. McGreevy, and D. Vegh, *Holographic non-Fermi liquid fixed points*, arXiv:1101.0597.

[28] I. Heemskerk and J. Polchinski, *Holographic and Wilsonian Renormalization Groups*, *JHEP* **1106** (2011) 031, [arXiv:1010.1264].

[29] A. Einstein, *The Formal Foundation of the General Theory of Relativity*, *Sitzungsber.Preuss.Akad.Wiss.Berlin (Math.Phys.)* **1914** (1914) 1030–1085.

[30] A. Einstein, *The Field Equations of Gravitation*, *Sitzungsber.Preuss.Akad.Wiss.Berlin (Math.Phys.)* **1915** (1915) 844–847.

[31] E. Hubble, *A relation between distance and radial velocity among extra-galactic nebulae*, *Proc.Nat.Acad.Sci.* **15** (1929) 168–173.

[32] E. Hubble and M. L. Humason, *The Velocity-Distance Relation among Extra-Galactic Nebulae*, *Astrophys.J.* **74** (1931) 43–80.

[33] G. Lemaître, *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, *Annales de la Societe Scietifique de Bruxelles* **47** (1927) 49–59.

[34] G. Lemaitre, *Republication of: The beginning of the world from the point of view of quantum theory*, *Nature* **127** (1931) 706.

[35] A. Einstein, *Cosmological Considerations in the General Theory of Relativity*, *Sitzungsber.Preuss.Akad.Wiss.Berlin (Math.Phys.)* **1917** (1917) 142–152.

[36] W. de Sitter, *Einstein's theory of gravitation and its astronomical consequences, First Paper*, *Mon.Not.Roy.Astron.Soc.* **76** (1916) 699–728.

[37] W. de Sitter, *Einstein's theory of gravitation and its astronomical consequences, Second Paper*, *Mon.Not.Roy.Astron.Soc.* **77** (1916) 155–184.

[38] W. de Sitter, *Einstein's theory of gravitation and its astronomical consequences, Third Paper*, *Mon.Not.Roy.Astron.Soc.* **78** (1917) 3–28.

[39] Friedman, A., *Über die Krümmung des Raumes*, *Zeitschrift für Physik A Hadrons and Nuclei* **10** (1922) 377–386. 10.1007/BF01332580.

[40] H. P. Robertson, *On the Foundations of Relativistic Cosmology*, *Proceedings of the National Academy of Science* **15** (Nov., 1929) 822–829.

[41] H. Robertson, *Relativistic Cosmology*, *Rev.Mod.Phys.* **5** (1933) 62–90.

[42] A. G. Walker, *Distance in an expanding universe*, *Monthly Notices of the Royal Astronomical Society* **94** (Dec., 1933) 159.

[43] A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, and R. Chornock, *A 3Determination of the hubble constant with the hubble space telescope and wide field camera 3*, *The Astrophysical Journal* **730** (2011), no. 2 119.

[44] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson, *The 6df galaxy survey: baryon acoustic oscillations and the local hubble constant*, *Monthly Notices of the Royal Astronomical Society* **416** (2011), no. 4 3017–3032.

[45] A. A. Penzias and R. W. Wilson, *A Measurement of excess antenna temperature at 4080-Mc/s*, *Astrophys.J.* **142** (1965) 419–421.

[46] A. A. Penzias and R. W. Wilson, *Measurement of the Flux Density of CAS a at 4080 Mc/s.*, *Astrophysical Journal* **142** (Oct., 1965) 1149.

[47] R. Dicke, P. Peebles, P. Roll, and D. Wilkinson, *Cosmic Black-Body Radiation*, *Astrophys.J.* **142** (1965) 414–419.

[48] G. Gamow, *The Evolution of the Universe*, *Nature* **162** (Oct., 1948) 680–682.

[49] G. Gamow, *The Origin of Elements and the Separation of Galaxies*, *Physical Review* **74** (Aug., 1948) 505–506.

[50] R. Alpher, H. Bethe, and G. Gamow, *The origin of chemical elements*, *Phys.Rev.* **73** (1948) 803–804.

[51] R. A. Alpher and R. C. Herman, *On the relative abundance of the elements*, *Phys. Rev.* **74** (Dec, 1948) 1737–1742.

[52] D. J. Fixsen, *The Temperature of the Cosmic Microwave Background*, *APJ* **707** (Dec., 2009) 916–920, [arXiv:0911.1955].

[53] R. Penrose, *Singularities and time-asymmetry*, in *General relativity. An Einstein centenary survey* (S. Hawking and W. Israel, eds.). Cambridge University Press, 1979.

[54] S. Carlip, *Horizons, constraints, and black hole entropy*, *Int.J.Theor.Phys.* **46** (2007) 2192–2203, [gr-qc/0601041].

[55] S. Carlip, *Black Hole Entropy and the Problem of Universality*, *J.Phys.Conf.Ser.* **67** (2007) 012022, [gr-qc/0702094].

[56] S. Carlip, *Black Hole Entropy and the Problem of Universality*, arXiv:0807.4192.

[57] J. L. Cardy, *Operator Content of Two-Dimensional Conformally Invariant Theories*, *Nucl.Phys.* **B270** (1986) 186–204.

[58] H. Bloete, J. L. Cardy, and M. Nightingale, *Conformal Invariance, the Central Charge, and Universal Finite Size Amplitudes at Criticality*, *Phys.Rev.Lett.* **56** (1986) 742–745.

[59] E. P. Verlinde, *Fusion Rules and Modular Transformations in 2D Conformal Field Theory*, *Nucl.Phys.* **B300** (1988) 360.

[60] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys.Lett.* **B379** (1996) 99–104, [hep-th/9601029].

[61] G. 't Hooft, *Dimensional reduction in quantum gravity*, gr-qc/9310026. Essay dedicated to Abdus Salam.

[62] L. Susskind, *The World as a hologram*, *J.Math.Phys.* **36** (1995) 6377–6396, [hep-th/9409089].

[63] L. Bombelli, R. K. Koul, J. Lee, and R. D. Sorkin, *A Quantum Source of Entropy for Black Holes*, *Phys.Rev.* **D34** (1986) 373–383.

[64] M. Srednicki, *Entropy and area*, *Phys.Rev.Lett.* **71** (1993) 666–669, [hep-th/9303048].

[65] W. Israel, *Thermo field dynamics of black holes*, *Phys.Lett.* **A57** (1976) 107–110.

[66] R. Brustein, M. B. Einhorn, and A. Yarom, *Entanglement interpretation of black hole entropy in string theory*, *JHEP* **0601** (2006) 098, [hep-th/0508217].

[67] O. Aharony, O. Bergman, D. L. Jafferis, and J. Maldacena, *N=6 superconformal Chern-Simons-matter theories, M2-branes and their gravity duals*, *JHEP* **0810** (2008) 091, [arXiv:0806.1218].

[68] A. Strominger, *The dS / CFT correspondence*, *JHEP* **0110** (2001) 034, [hep-th/0106113].

[69] S. Hawking and G. Ellis, *The Large scale structure of space-time*. Cambridge University Press, Cambridge (UK), 1973.

[70] J. L. Petersen, *Introduction to the Maldacena conjecture on AdS / CFT*, *Int.J.Mod.Phys.* **A14** (1999) 3597–3672, [hep-th/9902131].

[71] A. Zaffaroni, *Introduction to the AdS-CFT correspondence*, *Class.Quant.Grav.* **17** (2000) 3571–3597.

[72] E. D'Hoker and D. Z. Freedman, *Supersymmetric gauge theories and the AdS / CFT correspondence*, in *Strings, Branes and Extra Dimensions: TASI 2001 Proceedings* (Steven S. Gubser and Joseph D. Lykken, ed.), (River Edge (USA)), pp. 3–158, World Scientific, 2002. hep-th/0201253.

[73] J. M. Maldacena, *TASI 2003 lectures on AdS / CFT*, in *Progress in string theory. Proceedings, Summer School, TASI 2003, Boulder, USA, June 2-27, 2003* (J. M. Maldacena, ed.), (Hackensack (USA)), pp. 155–203, World Scientific, 2004. hep-th/0309246.

[74] J. Maldacena, *The Gauge/gravity duality*, in *Black holes in higher dimensions* (Gary Horowitz, ed.). Cambridge University Press, Cambridge (UK), 2011. arXiv:1106.6073.

[75] C. W. Misner, K. Thorne, and J. Wheeler, *Gravitation*. W. H. Freeman, San Francisco, 1974.

[76] R. M. Wald, *General Relativity*. Chicago University Press, Chicago (USA), 1984.

[77] S. M. Carroll, *Lecture notes on general relativity*, gr-qc/9712019.

[78] V. Balasubramanian, P. Kraus, and A. E. Lawrence, *Bulk versus boundary dynamics in anti-de Sitter space-time*, *Phys.Rev.* **D59** (1999) 046003, [hep-th/9805171].

[79] R. Grimm, M. Sohnius, and J. Wess, *Extended Supersymmetry and Gauge Theories*, *Nucl.Phys.* **B133** (1978) 275.

[80] M. B. Green, J. H. Schwarz, and L. Brink, *N=4 Yang-Mills and N=8 Supergravity as Limits of String Theories*, *Nucl.Phys.* **B198** (1982) 474–492.

[81] J. H. Schwarz, *Covariant Field Equations of Chiral N=2 D=10 Supergravity*, *Nucl.Phys.* **B226** (1983) 269.

[82] E. Witten, *Multi-trace operators, boundary conditions, and AdS/CFT correspondence*, hep-th/0112258.

[83] A. Sever and A. Shomer, *A Note on multitrace deformations and AdS/CFT*, *JHEP* **0207** (2002) 027, [hep-th/0203168].

[84] P. Breitenlohner and D. Z. Freedman, *Stability in Gauged Extended Supergravity*, *Annals Phys.* **144** (1982) 249.

[85] A. Bernamonti and B. Craps, *D-Brane Potentials from Multi-Trace Deformations in AdS/CFT*, *JHEP* **0908** (2009) 112, [arXiv:0907.0889].

[86] J. L. Karczmarek and A. Strominger, *Matrix cosmology*, *JHEP* **0404** (2004) 055, [hep-th/0309138].

[87] S. R. Das and J. L. Karczmarek, *Spacelike boundaries from the c=1 matrix model*, *Phys.Rev.* **D71** (2005) 086006, [hep-th/0412093].

[88] T. Banks, W. Fischler, S. Shenker, and L. Susskind, *M theory as a matrix model: A Conjecture*, *Phys.Rev.* **D55** (1997) 5112–5128, [hep-th/9610043].

[89] B. Craps, S. Sethi, and E. P. Verlinde, *A Matrix big bang*, *JHEP* **0510** (2005) 005, [hep-th/0506180].

[90] S. R. Das and J. Michelson, *Matrix membrane big bangs and D-brane production*, *Phys.Rev.* **D73** (2006) 126006, [hep-th/0602099].

[91] B. Craps, A. Rajaraman, and S. Sethi, *Effective dynamics of the matrix big bang*, *Phys.Rev.* **D73** (2006) 106005, [hep-th/0601062].

[92] E. J. Martinec, D. Robbins, and S. Sethi, *Toward the end of time*, *JHEP* **0608** (2006) 025, [hep-th/0603104].

[93] P. Kraus, H. Ooguri, and S. Shenker, *Inside the horizon with AdS / CFT*, *Phys.Rev.* **D67** (2003) 124022, [hep-th/0212277].

[94] L. Fidkowski, V. Hubeny, M. Kleban, and S. Shenker, *The Black hole singularity in AdS / CFT*, *JHEP* **0402** (2004) 014, [hep-th/0306170].

[95] G. Festuccia and H. Liu, *Excursions beyond the horizon: Black hole singularities in Yang-Mills theories. I.*, *JHEP* **0604** (2006) 044, [hep-th/0506202].

[96] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, *Local bulk operators in AdS/CFT: A Boundary view of horizons and locality*, *Phys.Rev.* **D73** (2006) 086003, [hep-th/0506118].

[97] M. Cvetic, S. Nojiri, and S. D. Odintsov, *Cosmological anti-de Sitter space-times and time dependent AdS / CFT correspondence*, *Phys.Rev.* **D69** (2004) 023513, [hep-th/0306031].

[98] C.-S. Chu and P.-M. Ho, *Time-dependent AdS/CFT duality and null singularity*, *JHEP* **0604** (2006) 013, [hep-th/0602054].

[99] C.-S. Chu and P.-M. Ho, *Time-dependent AdS/CFT duality. II. Holographic reconstruction of bulk metric and possible resolution of singularity*, *JHEP* **0802** (2008) 058, [arXiv:0710.2640].

[100] S. R. Das, J. Michelson, K. Narayan, and S. P. Trivedi, *Time dependent cosmologies and their duals*, *Phys.Rev.* **D74** (2006) 026002, [hep-th/0602107].

[101] F.-L. Lin and W.-Y. Wen, *Supersymmetric null-like holographic cosmologies*, *JHEP* **0605** (2006) 013, [hep-th/0602124].

[102] A. Awad, S. R. Das, K. Narayan, and S. P. Trivedi, *Gauge theory duals of cosmological backgrounds and their energy momentum tensors*, *Phys.Rev.* **D77** (2008) 046008, [arXiv:0711.2994].

[103] A. Awad, S. R. Das, S. Nampuri, K. Narayan, and S. P. Trivedi, *Gauge Theories with Time Dependent Couplings and their Cosmological Duals*, *Phys.Rev.* **D79** (2009) 046004, [arXiv:0807.1517].

[104] T. Hertog and G. T. Horowitz, *Towards a big crunch dual*, *JHEP* **07** (2004) 073, [hep-th/0406134].

[105] T. Hertog and G. T. Horowitz, *Holographic description of AdS cosmologies*, *JHEP* **04** (2005) 005, [hep-th/0503071].

[106] B. Craps, T. Hertog, and N. Turok, *On the Quantum Resolution of Cosmological Singularities using AdS/CFT*, *Phys.Rev.* **D86** (2012) 043513, [arXiv:0712.4180].

[107] N. Turok, B. Craps, and T. Hertog, *From Big Crunch to Big Bang with AdS/CFT*, arXiv:0711.1824.

[108] J. L. Barbon and E. Rabinovici, *Holography of AdS vacuum bubbles*, *JHEP* **1004** (2010) 123, [arXiv:1003.4966].

[109] L. Battarra and T. Hertog, *Particle Production near an AdS Crunch*, *JHEP* **1012** (2010) 017, [arXiv:1009.0992].

[110] M. Gunaydin, L. Romans, and N. Warner, *Gauged N=8 Supergravity in Five-Dimensions*, *Phys.Lett.* **B154** (1985) 268.

[111] M. Gunaydin, L. Romans, and N. Warner, *Compact and Noncompact Gauged Supergravity Theories in Five-Dimensions*, *Nucl.Phys.* **B272** (1986) 598.

[112] M. Pernici, K. Pilch, and P. van Nieuwenhuizen, *Gauged N=8 D=5 Supergravity*, *Nucl.Phys.* **B259** (1985) 460.

[113] D. Freedman, S. Gubser, K. Pilch, and N. Warner, *Continuous distributions of D3-branes and gauged supergravity*, *JHEP* **0007** (2000) 038, [hep-th/9906194].

[114] J. Maldacena, *Vacuum decay into Anti de Sitter space*, arXiv:1012.0274.

[115] D. Harlow and L. Susskind, *Crunches, Hats, and a Conjecture*, arXiv:1012.5302.

[116] S. R. Coleman and E. J. Weinberg, *Radiative Corrections as the Origin of Spontaneous Symmetry Breaking*, *Phys. Rev.* **D7** (1973) 1888–1910.

[117] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. 2. Fourier Analysis, Selfadjointness*. Elsevier Science Publishing, New York (USA), 1975.

[118] M. Carreau, E. Farhi, S. Gutmann, and P. F. Mende, *The Functional Integral for Quantum Systems with Hamiltonians Unbounded from below*, *Annals Phys.* **204** (1990) 186–207.

[119] J. Honerkamp, *Chiral multiloops*, *Nucl.Phys.* **B36** (1972) 130–140.

[120] J. Honerkamp, *The Question of invariant renormalizability of the massless Yang-Mills theory in a manifest covariant approach*, *Nucl.Phys.* **B48** (1972) 269–287.

[121] R. Kallosh, *The Renormalization in Nonabelian Gauge Theories*, *Nucl.Phys.* **B78** (1974) 293.

[122] L. F. Abbott, *The Background Field Method Beyond One Loop*, *Nucl. Phys.* **B185** (1981) 189.

[123] N. Nielsen, *On the Gauge Dependence of Spontaneous Symmetry Breaking in Gauge Theories*, *Nucl.Phys.* **B101** (1975) 173.

[124] R. Fukuda and T. Kugo, *Gauge Invariance in the Effective Action and Potential*, *Phys.Rev.* **D13** (1976) 3469.

[125] S. P. Martin, *Two-loop effective potential for a general renormalizable theory and softly broken supersymmetry*, *Phys. Rev.* **D65** (2002) 116003, [hep-ph/0111209].

[126] R. Ball, *Chiral Gauge Theory*, *Phys.Rept.* **182** (1989) 1.

[127] C. Ford, I. Jack, and D. Jones, *The Standard model effective potential at two loops*, *Nucl.Phys.* **B387** (1992) 373–390, [hep-ph/0111190].

[128] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading (USA), 1995.

[129] E. J. Weinberg and A.-q. Wu, *Understanding Complex Perturbative Effective Potentials*, *Phys. Rev.* **D36** (1987) 2474.

[130] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et. al.*, *Discovery of a supernova explosion at half the age of the Universe and its cosmological implications*, *Nature* **391** (1998) 51–54, [astro-ph/9712212].

[131] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et. al.*, *Measurements of Omega and Lambda from 42 high redshift supernovae*, *Astrophys.J.* **517** (1999) 565–586, [astro-ph/9812133].

[132] **Supernova Search Team** Collaboration, A. G. Riess *et. al.*, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron.J.* **116** (1998) 1009–1038, [astro-ph/9805201].

[133] **WMAP** Collaboration, D. Spergel *et. al.*, *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters*, *Astrophys.J.Suppl.* **148** (2003) 175–194, [astro-ph/0302209].

[134] **WMAP** Collaboration, E. Komatsu *et. al.*, *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys.J.Suppl.* **192** (2011) 18, [arXiv:1001.4538].

[135] J. Islam, *An Introduction to mathematical cosmology*. Cambridge University Press, Cambridge (UK), 1992.

[136] S. Dodelson, *Modern cosmology*. Academic Press, Amsterdam (The Netherlands), 2003.

[137] V. Mukhanov, *Physical foundations of cosmology*. Cambridge University Press, Cambridge (UK), 2005.

[138] A. Liddle, *An introduction to modern cosmology*. Wiley, Chichester (UK), 2009.

[139] H. Georgi and S. Glashow, *Unity of All Elementary Particle Forces*, *Phys.Rev.Lett.* **32** (1974) 438–441.

[140] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*, *Phys.Rev.* **D10** (1974) 275–289.

[141] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961) 579–588.

[142] S. Weinberg, *A Model of Leptons*, *Phys.Rev.Lett.* **19** (1967) 1264–1266.

[143] A. Salam, *Weak and Electromagnetic Interactions*, *Conf.Proc.* **C680519** (1968) 367–377.

[144] G. 't Hooft, *Magnetic Monopoles in Unified Gauge Theories*, *Nucl.Phys.* **B79** (1974) 276–284.

[145] Y. Zeldovich and M. Khlopov, *On the Concentration of Relic Magnetic Monopoles in the Universe*, *Phys.Lett.* **B79** (1978) 239–241.

[146] J. Preskill, *Cosmological Production of Superheavy Magnetic Monopoles*, *Phys.Rev.Lett.* **43** (1979) 1365.

[147] A. H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, *Phys.Rev.* **D23** (1981) 347–356.

[148] A. D. Linde, *A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems*, *Phys.Lett.* **B108** (1982) 389–393.

[149] A. Albrecht and P. J. Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, *Phys.Rev.Lett.* **48** (1982) 1220–1223.

[150] R. H. Brandenberger and C. Vafa, *Superstrings in the Early Universe*, *Nucl.Phys.* **B316** (1989) 391.

[151] R. Brandenberger, D. A. Easson, and A. Mazumdar, *Inflation and brane gases*, *Phys.Rev.* **D69** (2004) 083502, [hep-th/0307043].

[152] P. J. Steinhardt and N. Turok, *A Cyclic model of the universe*, hep-th/0111030.

[153] P. J. Steinhardt and N. Turok, *Cosmic evolution in a cyclic universe*, *Phys.Rev.* **D65** (2002) 126003, [hep-th/0111098].

[154] J. Khoury, B. A. Ovrut, P. J. Steinhardt, and N. Turok, *Density perturbations in the ekpyrotic scenario*, *Phys.Rev.* **D66** (2002) 046005, [hep-th/0109050].

[155] P. McFadden and K. Skenderis, *Holography for Cosmology*, *Phys.Rev.* **D81** (2010) 021301, [arXiv:0907.5542].

[156] P. McFadden and K. Skenderis, *The Holographic Universe*, *J.Phys.Conf.Ser.* **222** (2010) 012007, [arXiv:1001.2007].

[157] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, *Gravity and global symmetries*, *Phys.Rev.* **D52** (1995) 912–935, [hep-th/9502069].

[158] M. B. Green, J. Schwarz, and E. Witten, *Superstring Theory. Vol. 1: Introduction*. Cambridge University Press, Cambridge (UK), 1987.

[159] M. B. Green, J. Schwarz, and E. Witten, *Superstring Theory. Vol. 2: Loop Amplitudes, Anomalies and Phenomenology*. Cambridge University Press, Cambridge (UK), 1987.

[160] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*. Cambridge University Press, Cambridge (UK), 1998.

[161] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*. Cambridge University Press, Cambridge (UK), 1998.

[162] P. S. Aspinwall, B. R. Greene, and D. R. Morrison, *Calabi-Yau moduli space, mirror manifolds and space-time topology change in string theory*, *Nucl.Phys.* **B416** (1994) 414–480, [hep-th/9309097].

[163] B. R. Greene, *String theory on Calabi-Yau manifolds*, in *Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996* (C. Efthimiou and B. Greene, eds.), pp. 543–726. World Scientific, Singapore, 1997. hep-th/9702155.

[164] A. Font and S. Theisen, *Introduction to string compactification*, *Lect.Notes Phys.* **668** (2005) 101–181.

[165] M. Grana, *Flux compactifications in string theory: A Comprehensive review*, *Phys.Rept.* **423** (2006) 91–158, [hep-th/0509003].

[166] M. R. Douglas and S. Kachru, *Flux compactification*, *Rev. Mod. Phys.* **79** (2007) 733–796, [hep-th/0610102].

[167] R. Blumenhagen, B. Kors, D. Lust, and S. Stieberger, *Four-dimensional String Compactifications with D-Branes, Orientifolds and Fluxes*, *Phys.Rept.* **445** (2007) 1–193, [hep-th/0610327].

[168] P. Candelas and X. de la Ossa, *Moduli space of calabi-yau manifolds*, *Nucl. Phys.* **B355** (1991) 455–481.

[169] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, *De Sitter vacua in string theory*, *Phys. Rev.* **D68** (2003) 046005, [hep-th/0301240].

[170] S. Gukov, C. Vafa, and E. Witten, *CFT's from Calabi-Yau four-folds*, *Nucl. Phys.* **B584** (2000) 69–108, [hep-th/9906070].

[171] K. Dasgupta, G. Rajesh, and S. Sethi, *M theory, orientifolds and G - flux*, *JHEP* **9908** (1999) 023, [hep-th/9908088].

[172] B. R. Greene, K. Schalm, and G. Shiu, *Warped compactifications in M and F theory*, *Nucl.Phys.* **B584** (2000) 480–508, [hep-th/0004103].

[173] S. B. Giddings, S. Kachru, and J. Polchinski, *Hierarchies from fluxes in string compactifications*, *Phys. Rev.* **D66** (2002) 106006, [hep-th/0105097].

[174] V. Balasubramanian, P. Berglund, J. P. Conlon, and F. Quevedo, *Systematics of Moduli Stabilisation in Calabi-Yau Flux Compactifications*, *JHEP* **03** (2005) 007, [hep-th/0502058].

[175] J. P. Conlon, F. Quevedo, and K. Suruliz, *Large-volume flux compactifications: Moduli spectrum and D3/D7 soft supersymmetry breaking*, *JHEP* **08** (2005) 007, [hep-th/0505076].

[176] S. Kachru *et. al.*, *Towards inflation in string theory*, *JCAP* **0310** (2003) 013, [hep-th/0308055].

[177] D. H. Lyth, *What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?*, *Phys.Rev.Lett.* **78** (1997) 1861–1863, [hep-ph/9606387].

[178] E. Silverstein and A. Westphal, *Monodromy in the CMB: Gravity Waves and String Inflation*, *Phys.Rev.* **D78** (2008) 106003, [arXiv:0803.3085].

[179] L. McAllister, E. Silverstein, and A. Westphal, *Gravity Waves and Linear Inflation from Axion Monodromy*, *Phys.Rev.* **D82** (2010) 046003, [arXiv:0808.0706].

[180] X. Dong, B. Horn, E. Silverstein, and A. Westphal, *Simple exercises to flatten your potential*, *Phys.Rev.* **D84** (2011) 026011, [arXiv:1011.4521].

[181] S. Groot Nibbelink and B. J. W. van Tent, *Density perturbations arising from multiple field slow- roll inflation*, hep-ph/0011325.

[182] S. Groot Nibbelink and B. J. W. van Tent, *Scalar perturbations during multiple field slow-roll inflation*, *Class. Quant. Grav.* **19** (2002) 613–640, [hep-ph/0107272].

[183] C. M. Peterson and M. Tegmark, *Testing Two-Field Inflation*, *Phys.Rev.* **D83** (2011) 023522, [arXiv:1005.4056].

[184] C. M. Peterson and M. Tegmark, *Testing Multi-Field Inflation: A Geometric Approach*, arXiv:1111.0927.

[185] S. Cremonini, Z. Lalak, and K. Turzynski, *On Non-Canonical Kinetic Terms and the Tilt of the Power Spectrum*, *Phys.Rev.* **D82** (2010) 047301, [arXiv:1005.4347].

[186] S. Cremonini, Z. Lalak, and K. Turzynski, *Strongly Coupled Perturbations in Two-Field Inflationary Models*, *JCAP* **1103** (2011) 016, [arXiv:1010.3021].

[187] L. McAllister and E. Silverstein, *String Cosmology: A Review*, *Gen. Rel. Grav.* **40** (2008) 565–605, [arXiv:0710.2951].

[188] D. Baumann and L. McAllister, *Advances in Inflation in String Theory*, *Ann. Rev. Nucl. Part. Sci.* **59** (2009) 67–94, [arXiv:0901.0265].

[189] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, *False vacuum inflation with Einstein gravity*, *Phys. Rev.* **D49** (1994) 6410–6433, [astro-ph/9401011].

[190] M. Berg, D. Marsh, L. McAllister, and E. Pajer, *Sequestering in String Compactifications*, *JHEP* **1106** (2011) 134, [arXiv:1012.1858].

[191] R. Kallosh, A. Linde, and T. Rube, *General inflaton potentials in supergravity*, *Phys. Rev.* **D83** (2011) 043507, [arXiv:1011.5945].

[192] D. H. Lyth and A. R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure*, p. 294. Cambridge University Press, Cambridge (UK), 2009.

[193] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, *Features of heavy physics in the CMB power spectrum*, *JCAP* **1101** (2011) 030, [arXiv:1010.3693].

[194] M. Gomez-Reino and C. A. Scrucca, *Locally stable non-supersymmetric minkowski vacua in supergravity*, *JHEP* **05** (2006) 015, [hep-th/0602246].

[195] E. Cremmer, P. Fayet, and L. Girardello, *Gravity Induced Supersymmetry Breaking and Low-Energy Mass Spectrum*, *Phys. Lett.* **B122** (1983) 41.

[196] P. Binetruy and M. K. Gaillard, *Temperature corrections, supersymmetric effective potentials ans inflation*, *Nucl. Phys.* **B254** (1985) 388.

[197] I. Ben-Dayan, R. Brustein, and S. P. de Alwis, *Models of Modular Inflation and Their Phenomenological Consequences*, *JCAP* **0807** (2008) 011, [arXiv:0802.3160].

[198] K. Choi, A. Falkowski, H. P. Nilles, M. Olechowski, and S. Pokorski, *Stability of flux compactifications and the pattern of supersymmetry breaking*, *JHEP* **11** (2004) 076, [hep-th/0411066].

[199] S. P. de Alwis, *Effective potentials for light moduli*, *Phys. Lett.* **B626** (2005) 223–229, [hep-th/0506266].

[200] S. P. de Alwis, *On integrating out heavy fields in susy theories*, *Phys. Lett.* **B628** (2005) 183–187, [hep-th/0506267].

[201] A. Achucarro and K. Sousa, *F-term uplifting and moduli stabilization consistent with Kahler invariance*, *JHEP* **0803** (2008) 002, [arXiv:0712.3460].

[202] A. Achúcarro, S. Hardeman, and K. Sousa, *Consistent Decoupling of Heavy Scalars and Moduli in N=1 Supergravity*, *Phys. Rev.* **D78** (2008) 101901, [arXiv:0806.4364].

[203] S. C. Davis and M. Postma, *Successfully combining SUGRA hybrid inflation and moduli stabilisation*, *JCAP* **0804** (2008) 022, [arXiv:0801.2116].

[204] T. W. B. Kibble, *Topology of Cosmic Domains and Strings*, *J. Phys.* **A9** (1976) 1387–1398.

[205] R. Jeannerot, J. Rocher, and M. Sakellariadou, *How generic is cosmic string formation in SUSY GUTs*, *Phys. Rev.* **D68** (2003) 103514, [hep-ph/0308134].

[206] F. L. Bezrukov and M. Shaposhnikov, *The Standard Model Higgs boson as the inflaton*, *Phys. Lett.* **B659** (2008) 703–706, [arXiv:0710.3755].

[207] C. Germani and A. Kehagias, *New Model of Inflation with Non-minimal Derivative Coupling of Standard Model Higgs Boson to Gravity*, *Phys. Rev. Lett.* **105** (2010) 011302, [arXiv:1003.2635].

[208] C. P. Burgess, J. M. Cline, H. Stoica, and F. Quevedo, *Inflation in realistic D-brane models*, *JHEP* **09** (2004) 033, [hep-th/0403119].

[209] L. Covi, M. Gomez-Reino, C. Gross, J. Louis, G. A. Palma, *et. al.*, *Constraints on modular inflation in supergravity and string theory*, *JHEP* **0808** (2008) 055, [arXiv:0805.3290].

[210] L. Covi, M. Gomez-Reino, J. Louis, G. A. Palma, and C. A. Srucca, *de Sitter vacua in no-scale supergravities and Calabi-Yau string models*, *JHEP* **06** (2008) 057, [arXiv:0804.1073].

[211] L. Covi, M. Gomez-Reino, C. Gross, G. A. Palma, and C. A. Scrucca, *Constructing de Sitter vacua in no-scale string models without uplifting*, *JHEP* **03** (2009) 146, [arXiv:0812.3864].

[212] D. Baumann, A. Dymarsky, I. R. Klebanov, L. McAllister, and P. J. Steinhardt, *A Delicate Universe*, *Phys. Rev. Lett.* **99** (2007) 141601, [arXiv:0705.3837].

[213] A. Achúcarro, S. Hardeman, and K. Sousa, *F-term uplifting and the supersymmetric integration of heavy moduli*, *JHEP* **11** (2008) 003, [arXiv:0809.1441].

[214] A. Achucarro, J.-O. Gong, S. Hardeman, G. A. Palma, and S. P. Patil, *Mass hierarchies and non-decoupling in multi-scalar field dynamics*, *Phys.Rev.* **D84** (2011) 043502, [arXiv:1005.3848].

[215] R. Kallosh and A. Linde, *New models of chaotic inflation in supergravity*, *JCAP* **1011** (2010) 011, [arXiv:1008.3375].

[216] S. Mollerach, *Isocurvature baryon perturbations and inflation*, *Phys. Rev.* **D42** (1990) 313–325.

[217] D. Polarski and A. A. Starobinsky, *Isocurvature perturbations in multiple inflationary models*, *Phys. Rev.* **D50** (1994) 6123–6129, [astro-ph/9404061].

[218] A. D. Linde and V. F. Mukhanov, *Nongaussian isocurvature perturbations from inflation*, *Phys. Rev.* **D56** (1997) 535–539, [astro-ph/9610219].

[219] J.-c. Hwang and H. Noh, *Cosmological perturbations with multiple fluids and fields*, *Class. Quant. Grav.* **19** (2002) 527–550, [astro-ph/0103244].

[220] D. Wands, N. Bartolo, S. Matarrese, and A. Riotto, *An observational test of two-field inflation*, *Phys. Rev.* **D66** (2002) 043520, [astro-ph/0205253].

[221] F. Di Marco, F. Finelli, and R. Brandenberger, *Adiabatic and Isocurvature Perturbations for Multifield Generalized Einstein Models*, *Phys. Rev.* **D67** (2003) 063512, [astro-ph/0211276].

[222] B. J. W. van Tent, *Multiple-field inflation and the CMB*, *Class. Quant. Grav.* **21** (2004) 349–370, [astro-ph/0307048].

[223] C. T. Byrnes and D. Wands, *Curvature and isocurvature perturbations from two-field inflation in a slow-roll expansion*, *Phys. Rev.* **D74** (2006) 043529, [astro-ph/0605679].

[224] D. Wands, *Multiple field inflation*, *Lect. Notes Phys.* **738** (2008) 275–304, [astro-ph/0702187].

[225] Z. Lalak, D. Langlois, S. Pokorski, and K. Turzyński, *Curvature and isocurvature perturbations in two-field inflation*, *JCAP* **0707** (2007) 014, [arXiv:0704.0212].

[226] K. A. Malik and D. Wands, *Cosmological perturbations*, *Phys. Rept.* **475** (2009) 1–51, [arXiv:0809.4944].

[227] D. Langlois, F. Vernizzi, and D. Wands, *Non-linear isocurvature perturbations and non-Gaussianities*, *JCAP* **0812** (2008) 004, [arXiv:0809.4646].

[228] D. Langlois and S. Renaux-Petel, *Perturbations in generalized multi-field inflation*, *JCAP* **0804** (2008) 017, [arXiv:0801.1085].

[229] D. Langlois, S. Renaux-Petel, D. A. Steer, and T. Tanaka, *Primordial fluctuations and non-Gaussianities in multi- field DBI inflation*, *Phys. Rev. Lett.* **101** (2008) 061301, [arXiv:0804.3139].

[230] D. Langlois, S. Renaux-Petel, D. A. Steer, and T. Tanaka, *Primordial perturbations and non-Gaussianities in DBI and general multi-field inflation*, *Phys. Rev.* **D78** (2008) 063523, [arXiv:0806.0336].

[231] G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, *Non-linear perturbations in multiple-field inflation*, *Phys. Rev.* **D73** (2006) 083521, [astro-ph/0504508].

[232] L. Verde, L.-M. Wang, A. Heavens, and M. Kamionkowski, *Large-scale structure, the cosmic microwave background, and primordial non-gaussianity*, *Mon. Not. Roy. Astron. Soc.* **313** (2000) L141–L147, [astro-ph/9906301].

[233] N. Bartolo, S. Matarrese, and A. Riotto, *Non-Gaussianity from inflation*, *Phys. Rev.* **D65** (2002) 103505, [hep-ph/0112261].

[234] F. Bernardeau and J.-P. Uzan, *Non-Gaussianity in multi-field inflation*, *Phys. Rev.* **D66** (2002) 103506, [hep-ph/0207295].

[235] D. Seery and J. E. Lidsey, *Primordial non-gaussianities from multiple-field inflation*, *JCAP* **0509** (2005) 011, [astro-ph/0506056].

[236] G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, *Large non-Gaussianity in multiple-field inflation*, *Phys. Rev.* **D73** (2006) 083522, [astro-ph/0506704].

[237] G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, *Quantitative bispectra from multifield inflation*, *Phys. Rev.* **D76** (2007) 083512, [astro-ph/0511041].

[238] F. Vernizzi and D. Wands, *Non-Gaussianities in two-field inflation*, *JCAP* **0605** (2006) 019, [astro-ph/0603799].

[239] A. J. Tolley and M. Wyman, *The Gelaton Scenario: Equilateral non-Gaussianity from multi-field dynamics*, *Phys. Rev.* **D81** (2010) 043502, [arXiv:0910.1853].

[240] X. Chen and Y. Wang, *Large non-Gaussianities with Intermediate Shapes from Quasi-Single Field Inflation*, *Phys. Rev.* **D81** (2010) 063511, [arXiv:0909.0496].

[241] X. Chen and Y. Wang, *Quasi-Single Field Inflation and Non-Gaussianities*, *JCAP* **1004** (2010) 027, [arXiv:0911.3380].

[242] **Planck** Collaboration, *The Scientific programme of Planck*, astro-ph/0604069.

[243] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, *Phys. Rept.* **110** (1984) 1–162.

[244] J. Wess and J. Bagger, *Supersymmetry and supergravity*. Princeton University Press, Princeton (USA), 1992.

[245] T. Banks, D. B. Kaplan, and A. E. Nelson, *Cosmological Implications of Dynamical Supersymmetry Breaking*, *Phys. Rev.* **D49** (1994) 779–787, [hep-ph/9308292].

[246] R. L. Arnowitt and P. Nath, *Supersymmetry and supergravity: Phenomenology and grand unification*, in *Particles and fields. Proceedings, 7th J.A. Swieca Summer School, Campos do Jordao, Sao Paulo, Brazil, January 10-23, 1993* (O. J. Eboli and V. O. Rivelles, eds.). World Scientific, Singapore, 1993. hep-ph/9309277.

[247] J. Bagger, E. Poppitz, and L. Randall, *The R axion from dynamical supersymmetry breaking*, *Nucl. Phys.* **B426** (1994) 3–18, [hep-ph/9405345].

[248] J. Bagger, *Dynamical supersymmetry breaking in supergravity theories*, in *Theory meets experiment. Proceedings, 18th Johns Hopkins Workshop on Current Problems in Particle Theory, Florence, Italy, August 31 - September 2, 1994* (R. Casalbuoni, G. Domokos, S. Kovesi-Domokos, and B. Monteleoni, eds.). World Scientific, Singapore, 1994. hep-ph/9503245.

[249] C. Muñoz, *Soft supersymmetry breaking terms and the mu problem*, in *Elementary particle physics: Present and future. Proceedings, International Workshop, Valencia, Spain, June 5-9, 1995* (J. Valle and A. Ferrer, eds.). World Scientific Publishing Company, Singapore, River Edge, NJ (USA), 1995. hep-th/9507108.

[250] M. Dine, *TASI lectures on M theory phenomenology*, in *Strings, branes and gravity. Proceedings, Theoretical Advanced Study Institute, TASI'99, Boulder, USA, May 31-June 25, 1999* (J. A. Harvey, S. Kachru, and E. Silverstein, eds.). World Scientific, River Edge, NJ (USA), 2000. hep-th/0003175.

[251] L. Alvarez-Gaume, C. Gomez, and R. Jimenez, *Minimal Inflation*, *Phys.Lett.* **B690** (2010) 68–72, [arXiv:1001.0010].

[252] M. Dine and L. Pack, *Studies in Small Field Inflation*, *JCAP* **1206** (2012) 033, [arXiv:1109.2079].

[253] A. Achucarro, S. Mooij, P. Ortiz, and M. Postma, *Sgoldstino inflation*, *JCAP* (2012) to appear, [arXiv:1203.1907].

[254] D. Gallego, *On the Effective Description of Large Volume Compactifications*, *JHEP* **06** (2011) 087, [arXiv:1103.5469].

[255] V. S. Kaplunovsky and J. Louis, *Model independent analysis of soft terms in effective supergravity and in string theory*, *Phys.Lett.* **B306** (1993) 269–275, [hep-th/9303040].

[256] A. D. Linde, *Hybrid inflation*, *Phys.Rev.* **D49** (1994) 748–754, [astro-ph/9307002].

[257] R. Barbieri, E. Cremmer, and S. Ferrara, *Flat and positive potentials in N=1 supergravity*, *Phys. Lett.* **B163** (1985) 143.

[258] J. P. Hsu, R. Kallosh, and S. Prokushkin, *On brane inflation with volume stabilization*, *JCAP* **0312** (2003) 009, [hep-th/0311077].

[259] P. Binetruy, G. Dvali, R. Kallosh, and A. Van Proeyen, *Fayet-Iliopoulos terms in supergravity and cosmology*, *Class. Quant. Grav.* **21** (2004) 3137–3170, [hep-th/0402046].

[260] G. Dvali, Q. Shafi, and R. K. Schaefer, *Large scale structure and supersymmetric inflation without fine tuning*, *Phys.Rev.Lett.* **73** (1994) 1886–1889, [hep-ph/9406319].

[261] A. D. Linde and A. Riotto, *Hybrid inflation in supergravity*, *Phys.Rev.* **D56** (1997) 1841–1844, [hep-ph/9703209].

[262] S. Ferrara, C. Kounnas, and F. Zwirner, *Mass formulae and natural hierarchy in string effective supergravities*, *Nucl. Phys.* **B429** (1994) 589–625, [hep-th/9405188].

Summary

In a way, physics is still recovering from the 1896 revolution started by Max Planck in Berlin, then being fostered in Göttingen and from there finally conquering the world: the conception of quantum mechanics. The fact that objects on a very small scale such as atoms, atomic nuclei and photons behave qualitatively very different from our every day experience is not only puzzling the layman, but still ignites a lot of discussions in physics. A very important one is, how this theory tallies with the other big theory developed around the beginning of the previous century, namely the theory of relativity.

Quantum Mechanics and Relativity

The two big theories, which we have in physics are in a way antagonists. To see this, it is important to realize that every theory in physics comes with a domain of applicability. For a specific question in mind, there is a suitable theory to answer it, and if not, it can be developed. The beauty is, that mostly, a theory does not only apply to one question but to a whole class of them. We even know, which questions belong to the same class and all of them thus need to be treated with the same theory, otherwise leaving us with a puzzling contradiction. The different classes are characterized by the value of some parameter e.g. some energy scale.

In the case of quantum mechanics, the parameter which occurs in all the expressions is Planck's constant $\hbar$, which is negligibly small compared to all the other quantities, that enter in the physics of everyday questions. However, it is of the same order of magnitude as quantities on the atomic scale. All questions for which $\hbar$ is a considerable number should be treated quantum mechanically. In the case of the theory of special relativity the distinguishing parameter ist the speed of light. The intriguing effects of special relativity are all supressed and invisible, if the velocities involved in the problem are small as compared to the speed of light, as in everyday situations, but they do becomes visible for situations involving cosmic rays, ultra-fast trains or warp-driven space-flight. There is one new insight, though, which is very substantial to special relativity and has changed the way we are thinking about physics very fundamentally since the Newtonian paradigm. This is the observation, that time and space are not independent coordinates, which provide labels for when and where a specific event has happened, they

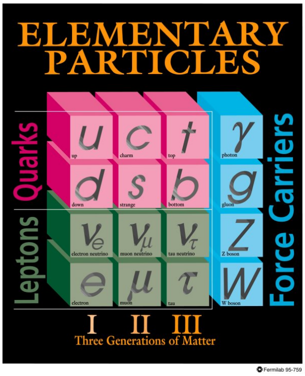


Figure 3.1: *The standard model of particle physics contains and describes all the elementary particles we know of today. Source: Fermilab*

are rather intricately related and cannot be considered or even exist separate from each other. Special relativity has uncovered that the notion of space must be extended, called space-time.

It is very informative to understand the motivation, which led Einstein to the discovery of special relativity. What puzzeled him was that the laws of classical mechanics according to Newton were incompatible with the laws of electrodynamics as uncovered by Maxwell. The motivation for developing special relativity was to unify two theories, which were completely correct in their own realm but led to a contradiction in a regime, which brought their characteristics to an overlap.

Many years later, the same problem arose with the theory of special relativity and the theory of quantum mechanics. Combining the symmetry of space-time with quantum physics it turned out that the "whole is more than the sum of its parts". This new thing took the shape as quantum field theory. It is the theory, which describes the creation, annihilation, the behaviour and interaction of elementary particles. The standard model of particle physics, in which all matter and forces are fundamentally described by such particles, is formulated in the language of quantum field theory with specific symmetries, which relate the couplings and masses of particles. Thanks to the symmetries those particles can be grouped in families and generations, as is depicted in figure 3.1.

The next revolution in relativity was the discovery of general relativity, which is our still valid theory of gravity. We use it to calculate the planetary motions, correct the signals of GPS-satellites and explain the development of the cosmos. At the basis of this extension of his earlier theory of special relativity was Einsteins insight, that acceleration, which you feel for instance in a rollercoaster, has the same effect, namely being pushed into the seat, as gravity, which accelerates freely falling objects or gives them a weight, if they are held fixed. Thinking this idea through, it turns out that the four dimensional continuum of space and time is not flat and static but rather bent, curved and not fixed at all. The motion of a body is percieved as its motion on the shortest path through space and time. Every mass in the continuum will cause some deformation, some bending, like a heavy sphere would do on a rubber sheet. A smaller object, a planet, say, passing a heavier object, the sun, say, feels this distortion and will alter its path to keep it being the shortest. Projected into the three dimensional space, this will then have the effect that it orbits the heavier object due to the gravitational force exerted by it.

GENESIS

Since space and time are not fixed but evolve and change according to the laws of general relativity, this theory gives us the fascinating prospect of studying the genesis of the universe itself. Depending on the inital conditions, on the amount and kind of matter in the universe, its evolution can be calculated and looks distinctly different. Ideally, those theoretical results can be compared with experimental findings and thus teach us about the real nature of nature.

A classic and very important such observation is that the light coming from other galaxies is red-shifted. This means its frequency is lower than we would naturally expect. This can be attributed to the same effect, which makes the siren of an approaching ambulance sound higher than when it recedes, the Doppler effect. The light from distant galaxies shows, that they are receding from us and we can conclude that the universe is expanding. The theory, indeed, has a solution, which describes an expanding universe, characterized by the Hubble parameter, the rate at which distances in the universe become larger.

If this observation is extrapolated backwards in time the cosmos shrinks and all galaxies get closer and closer together. At some point, they all merge and the density of mass increases coninuously. Thereby, the universe gets hotter and hotter. At some point, matter will change its shape and state, first atoms disintegrate, then nuclei until all the matter is transformed into a hot and dense plasma, which is even unpenetrable by light. If this extrapolation is done even further, at some point, we hit a singularity and the theory breaks down. That means, the concentration of matter curves space-time infinitely strong. General relativity cannot deal with that. This is also the point, where our ability to go even further back in time within the established paradigm stops. We percieve this point as the beginning of the universe and we picture this beginning as a big bang, a spontaneous, very energetic explosion, which provided enough energy to drive the cosmic evolution ever since.

SOME KEY COSMOLOGICAL PROBLEMS

We don’t like singularities. Rather than having any physical meaning, they signal that we have not understood the physics of the problem we are looking at. Also, our natural curiosity does not want to be forbidden to ask questions like what was before the big bang. In other words, we want to resolve the big bang singularity, we want to have a theory, which is well under control and which allows to describe the dynamics not only some time after, but also during and maybe even before the big bang.

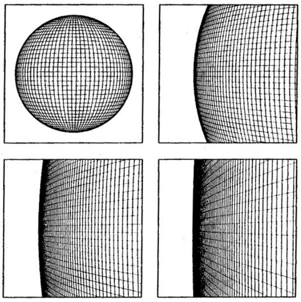
But even in the time after the big bang, where the theory of general relativity is applicable and does make predictions, those do not tally very well with observations. There are three classic problems of big bang cosmology, which make it necessary to adjust the cosmological scenario at early times in some crucial ways. Those problems are the so-called *flatness problem*, the *horizon problem* and the *monopole problem*.

Measurements indicate that the universe is very smooth and flat. Large parts of it are empty and the deformations of space-time caused by the matter scarsely distributed throughout it are very small as compared to the scale of the universe. But this is a very unlikely situation according to the standard cosmology as derived from general relativity. This lack of plausibility is perceived as the flatness problem.

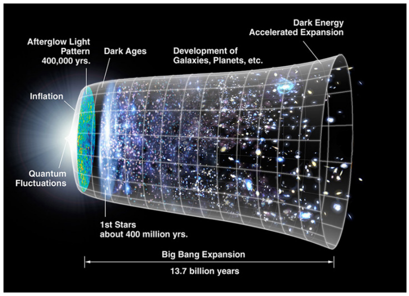
Furthermore, when we look at the very eldest photons which can be observed, the cosmic microwave background radiation, a very striking fact ist that their temperature is very uniform. We can assign the same temperature to it, no matter which direction in the sky the radiation is coming from. In general, physical systems need some time to attain a uniform temperature, to equilibriate. If you pour a cup of coffee, its temperature is higher than the temperature of the room, in which you are going to sit down and drink it. You can only enjoy it hot because it takes some time until the room temperature and the temperature of the coffee have equilibriated and you better make sure to finish it before that happens. The cup of coffee will only equilibriate with its surrounding. It will not equilibriate with the air in the room of your neighbour's house. The regions from which the background radiation coming from different directions emenates, however, never were in touch. This could only happen, to remain in the picture, if you and your neighbours have made an agreement to keep your houses at exactly the same temperature. It seems very unlikely that this sort of conspiracy has happened in the early universe, which is referred to as the horizon problem.

Finally, monopoles are something like electrons, however, they do not carry electric but magnetic charge. Usually, magnets always come with a north- and a south pole. If you split those, you do not get two particles with either a north- and a south "charge", corresponding to an electron and a positron, but you get two magnets with both north- and south poles. This asymmetry between electric and magnetic force is supposed to disappear at energies much higher than attainable at particle accelerators but much lower than those, which occured in the early universe. This means, that magnetic monopoles should have been just as abundant as electrons back than. Up to date, we have not found a single such monopole. The question, where they have gone is known as the monopole problem.

It is essential to understand the nature of these problems. They do not prove anything or render the theory of general relativity invalid. They are problems with how "natural", how likely it is, that the universe has come into the shape it is today under the assumption of the theory. The *naturalness* of the observed world within a model is perceived as a measure of the amount of understanding which a certain model comprises.



(a) Cosmic inflation is supposed to have blown up the cosmos in a very short time to a multitude of its original size. The curvature of the universe is only visible, if its radius is not too big as compared to the scale of our expeiment. Thus, the universe looks flat after inflation. Source: Griffith Observatory, Caltech.



(b) A cartoon of the expansion of the cosmos from the big bang until today. This thesis is concerned with the part on the very left until the formation of the cosmic microwave background. Source: LAMBDA archives WMAP.

Figure 3.2: *Cosmic inflation*

Inflation as the New Cosmological Paradigm

The most popular mechanism to resolve these problems, which has been proposed to date, is cosmic inflation. The idea behind it is that if only the universe had expanded very rapidly at the beginning, the above problems would be solved or rather eliminated. During such a rapid expansion, the universe would have been blown up to 10^{28} times its size within an instant. This would be large enough, that all the observable cosmos originated from the same piece of the primordial soup (or coffee). Then, it is not surprising, that it has the same temperature everywhere and the horizon problem is solved. It would also have reduced every initial curvature of the universe to being unperceivable, just as the curvature of the earth is unperceivable on an everyday scale, because the radius of our planet is too big (see figure 3.2(a)). In this picture, there is no flatness problem. Also, magnetic monopoles which might well have been there at the beginning of the evolution are being homeopathically diluted during this expansion and it is no surprise that we do not observe them, which finally removes the monopole problem.

It turns out that the equations of general relativity admit the possibility of such a rapid expansion. Such a *cosmic inflation* would be caused by a particle, the *inflaton*, which has a very high potential energy at the very beginning of the universe. It would blow up the universe until at some point, it converts its energy to kinetic energy and leaves the universe alone.

Let us carefully note again, what kind of a solution this is. Several *fine-tuning problems* of tweaking the initial conditions of our universe have been replaced by the dynamics of one physical field, the inflaton. The period of inflation makes the cosmic evolution independent of

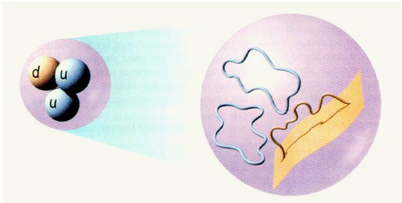


Figure 3.3: *Elementary particles are interpreted as vibrations of tiny strings. Source: University of Oregon*

the precise form of the initial conditions and increases the naturalness of our universe. Rather than tweaking some parameters into convenient shape, the task of the cosmologist is now to study the physical dynamics of a newly postulated field and come up with some idea to prove its existence.[1]

I should mention, that inflation has nothing to say about how to deal with the big bang singularity. The inflationary period will definitely remove the big bang from our observability. Since the evolution does not depend crucially on the initial conditions any longer, which would have to be imposed at the time of the big bang, it seems less crucial from an observational point of view to understand the big bang. It remains, though, an inconsistency of the theory, which should be dealt with.

The Early Universe as a Laboratory for Quantum Gravity

Traditionally, cosmological models use only general relativity as their underlying theory. This is a good theory to study the late evolution of the universe. When examining the very early universe, however, the typical length scales are small such that the effects of quantum mechanics are important and gravity cannot be applied without taking them into account. The early universe is an era, where gravity was strong and length scales were small. The laws of quantum mechanics should then be applied to gravity.

This is a big challenge. The difference between gravity and the other forces is that gravity couples to everything, to every energy, to every mass, even to itself. Where for the other forces, we only have to measure a small number of coupling constants, for gravity, there are infinitely many. Thus, it is difficult to quantize gravity. The approach for doing it nevertheless, which is considered in this thesis is *string theory*. The idea behind it is to replace elementary, point-like particles by extended objects as it is suggested in figure 3.3. These objects can vibrate just as the strings of a violin. Different particles are just an interpretation of different harmonics of the vibration of these strings, just like different tunes. Those strings have a very high tension, so that they look very much like point particles at everyday energies, where they are completely

[1] As a note added in proof let me point out that very recently the spectacular results of the PLANCK satellite mission have confirmed the inflationary paradigm to very high and unprecedented accuracy.

contracted, but they look extended at the high energies of the early universe. This theory has a very rich structure and contains a graviton, the quantum particle of the gravitational force.

However, the idea also has some problems. Mathematical consistency requires the theory to be formulated in ten dimensions rather than the four dimensions, to which our cosmos has evolved. To comply with the world around us, we need to get rid of the extra six dimensions. The way to do is is to curl them up to a very small size, much smaller than the accuracy of our experiments. At high energies, though, these *extra dimensions* would be visible. Research over the past decennia have shown that there is an enormous number of ways to go from ten to four dimensions and a lot of research in both physcis and mathematics has been examining the structures that can arise. We are not concerned with such questions here. For us, the important observation is that by compactifying extra dimensions, a lot of new parameters are introduced into the theory. In principle, the size, the shape and the exact form of the compactification is not specified. These parameters are new objects in the theory, so-called *moduli fields*. This means both good and bad luck for cosmology. On the one hand, these fields might be the inflaton, which we have conjectured above. On the other hand, it is rather unclear how these extra fields influence the dynamics of the cosmos.

Singularity Resolution and Inflation in this Thesis

Is string theory going to help us tame the big bang singularity? This is the first question which I am investigating in my thesis.

String theory has blessed us with a surprising insight: the physics of string theory in a certain space-time can be described by a quantum field theory on its boundary. This novel technique goes under the name of *AdS/CFT-correspondence* or *gauge/gravity duality*. The information about what is happening within a space-time, like our cosmos, can be recovered by studying a well understood theory on its boundary. Even better, the stronger gravity, the better behaved and understood is the corresponding field theory. Remember that the biggest problem with the big bang is that gravity becomes infinitely strong.

I use a very specific realisation, a toy model of a big bang singularity and examine, how it looks in the field theory, which belongs to it. It turns out that the field theory replicates the singularity of gravity. The big bang desguises itself as an instability in the field theory. However, this is only true in the limit, where the strings shrink to point particles. I have performed a calculation, in which the coupling between strings, which is supposed to give an important contribution to the physics in the early universe, is taken into account. It turns out, that including these string theory effects regularizes the field theory and makes the big bang a reasonable concept. It even makes sense now to ask what happened before the big bang, albeit we are very far from answering such questions.

In the second part of the thesis, I turn my attention to cosmic inflation. Its effectiveness very much depends on the specific form of the potential of the inflaton field. For it to support inflation long enough to solve the cosmological problems, it needs to be very flat. To have a really natural explanation of this phase of cosmology, this potential should be derived from a fundamental theory like string theory. Only then, Inflation would really solve the fine-tuning problem, because only then is the phenomenon explained in a natural way from a deeper understanding.

The proliferation of fields, which we have in string theory upon compactification, makes it very difficult to examine a model in all detail. Therefore, the usual procedure is to truncate ones model to only a few interesting fields and keep the rest nicely tugged away under the assumption that this can be done without invalidating ones model.

The question which I have investigated together with my colleagues is if those fields take revenge. We show that the effect these degrees of freedom have on the model are usually greately underestimated. Only under very specific circumstances is it admissible to neglect these fields. In the physicsally relevant cases, these conditions amount to choose conditions – again – in a very specific manner. Thus, fine-tuning, which inflation was supposed to remove, seems to have just been swept under some rug, which our more thorough analysis has lifted.

CONCLUSION

Examining the cosmology of the very early universe within a fundamental theory like string theory is necessary and exciting but dangerous. On the one hand, string theory has new features and techniques, which allow us to study the cosmological problems in a qualitatively different fashion. My results indicate that some of the gravest problems might be solved by string theory. On the other hand, string theory is beyon human control. One must be careful that one has really taken all the effects into account, which might well silently been reintroduce the problems one has set out to solve.

Meanwhile, advances in cosmology such as the ones reported in this thesis are well capable of satisfying the humen curiosity and the frontier of exploration is yet again pushed ahead a bit.

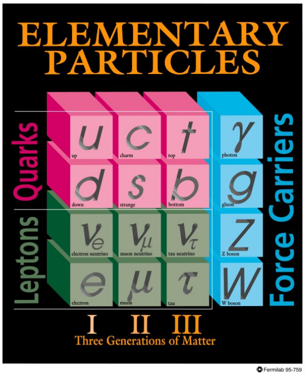
Samenvatting

In zekere zin probeert de fysica nog steeds te herstellen van de revolutie die Max Planck 1896 in Berlijn begonnen is en die, na wat versterking in Göttingen, de wereld veroverd heeft: de uitvondst van de kwantummechanica. Het feit dat objecten op hele kleine schaal zoals atomen, atoomkernen en fotonen zich kwalitatief heel anders gedragen dan onze dagelijkse ervaring verbaast niet alleen de leek maar leidt ook onder fysici nog steeds tot hevige discussies. Een hele belangrijke vraag is hoe deze theorie past bij de andere grote vondst van het begin van de vorige eeuw, namelijk de relativiteitstheorie.

Kwantummechanica en relativiteit

De twee grote theorie en die we in de natuurkunde kennen zijn in zekere zin antagonisten. Om dit te begrijpen is het belangrijk erbij stil te staan, dat elke theorie van de natuurkunde een domein van geldigheid heeft. Voor een specifieke vraag is er een passende theorie of, zo niet, kan die bedacht worden. Het mooie is dat een theorie meestal niet alleen voor één specifieke vraag van toepassing is, maar voor een hele klasse ervan. We weten ook welke vraagstukken bij dezelfde klasse horen en dus op dezelfde manier onderzocht moeten worden. Anders hebben we een raadselachtige tegenspraak. De verschillende klasses zijn meestal gekarakteriseerd door de een of andere parameter, bijvoorbeeld een energieschaal.

In het geval van de kwantummechanica is deze parameter de constante van Planck $\hbar$, die verschrikkelijk klein is ten opzichte van alle andere grootheden die we in alledaagse situaties tegenkomen. Maar hij is van dezelfde orde van grote als grootheden op een atomaire schaal. Alle vraagstukken, waarvoor $\hbar$ een aanzienlijk getal is, moeten kwantummechanisch beschouwd worden. In het geval van de speciale relativiteitstheorie is de karakteristieke parameter de lichtsnelheid. De verbijsterende effecten van de speciale relativiteitstheorie zijn allemaal onderdrukt en onzichtbaar als de snelheden die in een vraagstuk voorkomen klein zijn ten opzichte van de lichtsnelheid, zoals in alledaagse situaties. Ze worden wel zichtbaar bij kosmische straling, ultra-snelle treinen of de warpaandrijving. Er is echter één nieuw inzicht, die voor de speciale relativiteitstheorie heel substantieel is en de manier waarop we sinds de tijd van Newton over de natuurkunde nadenken voorgoed veranderd heeft. Dat is



Figuur 3.4: *Het standaardmodel van de deeltjesfysica bevat en beschrijft alle elementaire deeltjes die we vandaag de dag kennen. Bron: Fermilab*

de waarneming dat ruimte en tijd, waarmee vastgelegd wordt waar en wanneer iets gebeurd is, niet onafhankelijk van elkaar zijn, maar dat zij veeleer heel nauw verweven zijn en niet onafhankelijk van elkaar beschouwd worden, en zelfs niet eens bestaan kunnen. Speciale relativiteitstheorie houdt in dat het idee van de ruimte zelf verruimd moet worden en tot ruimtetijd wordt.

Het is heel interessant om de motivatie te begrijpen, die Einstein naar de ontdekking van de speciale relativiteitstheorie geleid heeft. Wat hem dwars zat is dat de wetten van de klassieke mechanica, zoals door Newton vastgesteld, niet strookten met de wetten van de elektrodynamica volgens Maxwell. De motivatie om de speciale relativiteitstheorie te ontwikkelen was om twee theorie en te verenigen, die op hun eigen domein volstrekt klopten maar elkaar tegenspraken als ze allebei van toepassing waren.

Velen jaren later bleek er een soortgelijk probleem tussen de speciale relativiteitstheorie en de theorie van de kwantummechanica te zijn. Toen men probeerde de symmetrie en van de ruimtetijd met de fysica van de atomen te combineren bleek het geheel meer te zijn dan de som van zijn delen. Dit nieuwe ding nam de vorm aan van een kwantumveldentheorie. Dit is een theorie die het ontstaan en vergaan, het gedrag en de wisselwerkingen van elementaire deeltjes beschrijft. Het standaardmodel van de deeltjesfysica, waarin alle materie en alle krachten op een fundamentele manier door deeltjes beschreven worden, is in de taal van de kwantumveldentheorie opgeschreven. Bepaalde symmetrie en relateren de eigenschappen van verschillende deeltjes aan elkaar en staan toe om ze in families en generaties in te delen zoals dit in figuur 3.4 afgebeeld is.

De volgende revolutie was de ontdekking van de algemene relativiteitstheorie, die we nog steeds als de beste beschrijving van de zwaartekracht beschouwen. We gebruiken hem om de banen van planeten te berekenen, de signalen van GPS satellieten te corrigeren en om de ontwikkeling van de kosmos te beschrijven. Deze uitbreiding van de speciale relativiteitstheorie is op Einstein's inzicht gegrond dat versnelling, die men bijvoorbeeld in een achtbaan voelt door in je stoel gedrukt te worden, hetzelfde effect heeft als de zwaartekracht, die vrij vallende objecten versnelt of hen een gewicht geeft als ze door iemand vastgehouden worden. Als men hierop doorgaat, blijkt het vierdimensionale continu um van ruimte en tijd niet vlak en statisch is, maar gebogen, gekromd en helemaal niet vast. De beweging van een lichaam wordt beschouwd als een beweging op het kortste pad door de ruimtetijd. Elke massa in dit continu um zorgt voor enige deformatie, voor een kromming, net zoals een zware bol een rubber plaat zou verbuigen. Een lichter object zoals een planeet die langs een zwaarder object heengaat, bijvoorbeeld de zon, kan deze verbuiging voelen en zal zijn pad aanpassen zodat deze de kortst mogelijke blijft. Als men dit op de driedimensionale ruimte projecteert lijkt het net alsof hij om het zwaardere object draait vanwege de zwaartekracht die ervan uitgaat.

GENESIS

Omdat ruimte en tijd niet meer vastliggen, maar volgens de wetten van de algemene relativiteitstheorie kunnen ontwikkelen en veranderen, hebben we nu de fascinerende mogelijkheid om de genesis van het universum zelf te bestuderen. Afhankelijk van de beginvoorwaarden, de hoeveelheid en soort materie in het universum, kan men zijn ontwikkeling berekenen, die heel verschillend kan uitpakken. Als het goed is, kan men deze theoretische resultaten dan met experimentele vondsten vergelijken en zo de echte natuur van de natuur leren.

Een klassiek en heel belangrijk voorbeeld van zo'n waarneming is dat het licht, dat van andere sterrenstelsels komt, rood verschoven is. Dat betekent dat zijn frequentie lager is dan wij normaal gesproken zouden verwachten. Dit komt door hetzelfde effect dat ervoor zorgt dat de sirene van een naderende ambulance hoger klinkt dan die van een vertrekkende, namelijk het Doppler effect. Het licht van verre sterrenstelsels toont aan dat deze van ons wegvliegen, en we kunnen dus concluderen dat het universum uitdijt. De theorie heeft daadwerkelijk een oplossing, die een uitdijend universum beschrijft. Deze wordt door de Hubbleparameter gekarakteriseerd, die het percentage aangeeft waarmee de afstanden in het universum groter worden.

Als men deze waarneming terug in de tijd extrapoleert wordt de kosmos steeds kleiner en alle sterrenstelsels komen steeds dichter en dichter op elkaar te zitten. Op een gegeven moment komen ze samen en de dichtheid neemt continu toe. Ondertussen word het universum steeds heter. Op den duur gaat de materie zijn gestalte en toestand veranderen, eerst gaan de atomen desintegreren, dan de atoomkernen totdat alle materie een heet en dicht plasma wordt, dat zelfs voor licht ondoordringbaar is. Als men in deze extrapolatie nog verder gaat komt men een singulariteit tegen, waar de theorie zijn geldigheid verliest. Dit betekent dat de concentratie van de materie de ruimtetijd oneindig sterk kromt. Dit kan de algemene relativiteitstheorie niet aan. Dit is ook het punt, waar onze mogelijkheid stopt om binnen hetzelfde paradigma nog verder terug te gaan in de tijd. We beschouwen dit punt als het begin van het universum en stellen hem ons voor als de oerknal, een spontane, heel energieke explosie, die genoeg energie ter beschikking stelt om de kosmische expansie sindsdien aan te drijven.

Sommige sleutelvragen van de cosmologie

We vinden singulariteiten niet leuk. In plaats van dat ze fysisch iets betekenen signaleren ze dat we de fysica van het probleem nog niet begrepen hebben. Ook wil zich onze natuurlijke nieuwgierigheid niet laten verbieden om vragen te stellen, zoals wat er dan voor de oerknal was. in anderen woorden, we willen de singulariteit van de oerknal oplossen. We willen een theorie die op een gecontroleerde manier de ontwikkeling van de kosmos niet alleen na, maar ook tijdens en misschien zelfs voor de oerknal beschrijft.

Maar ook in de tijd na de oerknal, waar de algemene relativiteitstheorie van toepassing is en voorspellingen doet, passen deze niet heel goed bij de waarnemingen. Er zijn drie klassieke problemen met de oerknalkosmologie, die het noodzakelijk maken het kosmologische scenario op een cruciale manier aan te passen. Dit zijn het *probleem van de vlakheid*, het *probleem van de horizon* en het *probleem van de monopolen*.

Metingen tonen aan dat het universum heel vlak en plat is. Grote delen ervan zijn leeg en de deformaties van de ruimtetijd door materie zijn ver verspreid en heel klein ten opzichte van de schaal van het universum. Maar deze uitkomst is volgens de standaard kosmologie, afgeleid van de algemene relativiteitstheorie, heel onwaarschijnlijk. Dit gebrek aan plausibiliteit wordt het probleem van de vlakheid genoemd.

Bovendien, als we de oudste fotonen die men kan waarnemen bekijken, de kosmische achtergrond straling, blijkt opvallend genoeg dat hun temperatuur uniform is. We kunnen hun dezelfde temperatuur toewijzen ongeacht uit welke richting van de hemel de straling vandaan komt. In het algemeen hebben fysische systemen wat tijd nodig om dezelfde temperatuur aan te nemen, met andere woorden, om in evenwicht te komen. Als je een kopje koffie inschenkt is zijn temperatuur aanzienlijk hoger dan de temperatuur van de kamer waarin je je bevindt. Je kan alleen van lekkere warme koffie genieten, omdat het een tijdje duurt voordat de temperatuur van de koffie en de temperatuur van de kamer in evenwicht komen, en je kunt dus maar beter zorgen dat je voor die tijd je koffie opgedronken hebt. Verder raakt die kop koffie alleen in evenwicht met zijn omgeving maar niet met de lucht in het huis van je buren. De verschillende regio's, vanwaar de kosmische achtergrondstraling uit verschillende richtingen vandaan komt, hadden echter nooit contact. Dit zou, om in het plaatje te blijven, alleen kunnen, als je een afspraak met je buren hebt gemaakt om de verwarming precies even hoog te zetten. Zo'n samenzwering in het vroege universum lijkt heel onwaarschijnlijk, wat men dan ook het probleem van de horizon noemt.

Monopolen, tenslotte, zijn zoiets als elektronen maar in plaats van elektrische hebben ze magnetische lading. Doorgaans hebben magneten altijd een noord- en een zuidpool. Als men deze uit elkaar knipt krijgt men niet één deeltje met een noord en één met een zuid-"lading", maar twee magneten met elk een noord- én een zuidpool. Deze asymmetrie tussen elektrische en magnetische kracht wordt geacht te verdwijnen bij energie en die veel hoger zijn dan wat we vandaag de dag met deeltjesversnellers kunnen bereiken, maar veel lager dan die in het vroege universum voorkwamen. Dat betekent, dat er toen net zoveel magnetische monopolen als elektronen hadden moeten zijn. Tot nu toe hebben we er echter nog geen gevonden. De vraag waar ze naartoe zijn staat bekend als het probleem van de monopolen.

Het is essentieel om het karakter van deze problemen te begrijpen. Ze bewijzen niets en ze maken de algemene relativiteitstheorie ook niet ongeldig. Het zijn problemen, hoe "natuurlijk", of hoe waarschijnlijk, het is dat onder de aanname van de theorie, een universum zoals die van ons is ontstaan. De *natuurlijkheid* van de waargenomen wereld binnen een model wordt beschouwd als een maat van het begrip dat een bepaald model heeft.

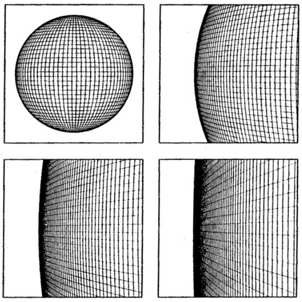
Inflatie als het nieuwe kosmologische paradigma

Tegenwoordig is het meest populaire mechanisme om deze problemen op te lossen kosmische inflatie. Het idee erachter is dat als er aan het begin een fase van snelle expansie (uitdijing) was, de voorgenoemde problemen opgelost of ge elimineerd zouden zijn. Tijdens zo'n snelle expansie zou het universum van het ene moment op het ander tot zo'n 10^{28} keer van zijn oorspronkelijke grootte gegroeid zijn. Dit zou groot genoeg zijn dat de hele waarneembare kosmos uit hetzelfde stukje van de oersoep (of -koffie) is ontstaan. Dan is het niet meer verrassend, dat de temperatuur overal hetzelfde is en het probleem van de horizon is opgelost. Dit zou ook elke oorspronkelijke kromming van het universum gereduceerd hebben en onwaarneembaar hebben gemaakt, net zoals de kromming van de aarde op alledaagse schaal niet wordt waargenomen, omdat de diameter van onze planeet te groot is (zie ook figuur 3.5(a)). In dit plaatje is het probleem van de vlakheid opgelost. Ook zouden magnetische monopolen, die in het begin van de evolutie er wel geweest zouden zijn, door de uitbreiding hom eopatisch verdunt zijn, zodat het geen verrassing is dat we ze niet waarnemen, waarmee tenslotte ook het probleem van de monopolen er niet meer is.

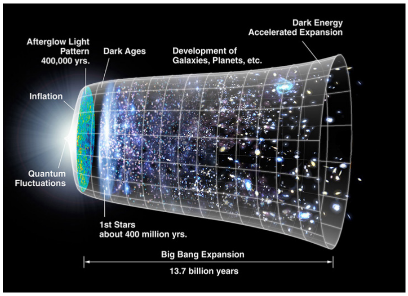
Het blijkt dat de vergelijkingen van de algemene relativiteitstheorie de mogelijkheid van zo'n snelle uitdijing toestaan. Deze *kosmische inflatie* zou door een deeltje veroorzaakt worden, het *inflaton*, dat in het begin van het universum een hele hoge potenti ele energie heeft. Deze zou het universum opblazen tot hij op een punt zijn potenti ele energie in kinetische energie omgezet heeft en het universum verder met rust laat.

Laten we nog een goed ernaar kijken wat voor een soort oplossing kosmische inflatie is. Meerdere problemen van nauwe, handmatige afstemming van de beginvoorwaarden van ons universum worden vervangen door de dynamica van één fysisch veld, het inflaton. De periode van inflatie maakt de kosmische evolutie onafhankelijk van de precieze vorm van de beginvoorwaarden en verhoogd de natuurlijkheid van ons universum. In plaats van het afstemmen van enkele parameters tot een gepaste waarde is de taak van de kosmoloog nu om de fysische dynamica van een nieuw gepostuleerd veld te bestuderen en met een voorstel te komen om zijn bestaan experimenteel aan te tonen.[2]

[2] Als een noot toegevoegd in de proefdruk merken we op dat heel recentelijk de spectaculaire resultaten van de PLANCK-satellietmissie het inflationaire paradigma op een tot nu toe ongekende nauwkeurigheid heeft bevestigd.



(a) Kosmische inflatie wordt geacht het universum in een hele korte tijd op een veelvoud van zijn oorspronkelijke grootte opgeblazen te hebben. De kromming van het universum is alleen waarneembaar als zijn radius niet te groot is ten opzichte van de schaal van ons experiment. Dus ziet het universum na inflatie er vlak uit. Bron: Griffith Observatory, Caltech.



(b) Een schets van de uitdijing van de kosmos van de oerknal tot vandaag. Dit proefschrift gaat over het deel helemaal links tot de formatie van de kosmische microgolf achtergrond. Bron: LAMBDA archives WMAP.

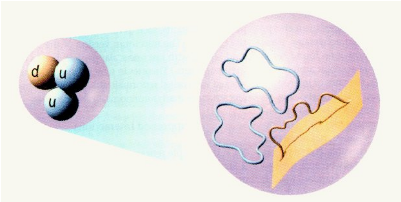
Figuur 3.5: *Kosmische inflatie*

Ik wil opmerken dat inflatie niets zegt over wat we met de oerknalsingulariteit moeten doen. De periode van inflatie zal de oerknal zeker aan onze waarneming onttrekken. Omdat de evolutie nu ook niet meer sterk van de beginvoorwaarden afhangt, die op het tijdstip van de oerknal hadden moeten worden vastgelegd, lijkt het uit het oogpunt van de waarnemingen minder belangrijk om de oerknal te begrijpen. Het blijft echter een inconsistentie van de theorie waar men een oplossing voor dient te vinden.

HET VROEGE UNIVERSUM ALS EEN LABORATORIUM VOOR KWANTUM ZWAARTEKRACHT

Traditioneel gebruiken kosmische modellen alleen algemene relativiteitstheorie als grondslag. Dit is een goede theorie om de late ontwikkeling van het universum te bestuderen. Als men het vroege universum onderzoekt zijn de typische lengteschalen echter klein dat de effecten van de kwantummechanica belangrijk worden en zwaartekracht niet toegepast kan worden zonder met hun rekening te houden. Het vroege universum is een tijdperk, waar gravitatie sterk en lengteschalen klein zijn. De wetten van de kwantummechanica moeten dan op de zwaartekracht toegepast worden.

Dat is een grote uitdaging. Het verschil tussen zwaartekracht en de andere krachten is dat gravitatie overal aan koppelt, aan elk energie, aan elke massa en zelfs aan zichzelf. Terwijl we



Figuur 3.6: *Elementaire deeltjes worden geinterprete erd als trillingen van minuscule snaren. Bron: University of Oregon*

voor de andere krachten maar een klein aantal koppelingsconstanten dienden te meten zijn er oneindig veel voor de zwaartekracht. Daarom is het moeilijk om zwaartekracht te kwantiseren. De manier om het alsnog te doen, die in dit proefschrift gebruikt wordt, is *snaartheorie*. Het idee erachter is om de elementaire puntdeeltjes door uitgerekte objecten te vervangen, zoals het in figuur 3.6 gesuggereerd wordt. Deze objecten kunnen net zoals de snaren van een viool trillen. Verschillende deeltjes zijn dan interpretaties van verschillende boventonen van de vibratie van deze snaren. Deze snaren hebben een hele hoge spanning zodat ze bij alledaagse energie en net op punt deeltjes lijken, doordat ze volledig opgerold zijn. Maar bij de hoge energie en van het vroege universum zijn ze uitgerekt. Snaartheorie heeft een hele rijke structuur en bevat een graviton, het kwantumdeeltje van de zwaartekracht.

Dit idee heeft echter ook zijn problemen. Wiskundige consistentie vergt dat de theorie in tien dimensies wordt geformuleerd in plaats van de vier dimensies waartoe onze kosmos zich ontwikkeld heeft. Om met de wereld om ons heen te stroken moeten we de extra zes dimensie kwijt zien te raken. De manier om dit te doen is door ze heel klein op te rollen, veel kleiner dan de nauwkeurigheid van onze experimenten. Pas bij hoge energie en zouden deze *extra dimensies* waarneembaar kunnen zijn. Onderzoek heeft tijdens de afgelopen decennia aangetoond dat er een enorme hoeveelheid van mogelijkheden is om van tien naar vier dimensies te gaan en heel veel onderzoek in zowel natuur- als ook wiskunde houdt zich bezig met de structuren die erdoor kunnen ontstaan. Deze vraagstukken zijn hier niet van belang. Voor ons is de belangrijke waarneming dat door het compactificie eren van extra dimensies een heleboel nieuwe parameters in de theorie ge introduceerd worden. In principe liggen de grootheid, de gedaante en de exacte vorm van de compactificatie niet vast. Deze parameters zijn nieuwe objecten in de theorie, zogeheten *modulivelden*. Dit is zowel goed als ook slecht nieuws voor de kosmologie. Aan de ene kant zouden deze velden het inflaton kunnen zijn, waarvan we eerder vereist hebben dat het bestaat. Aan de andere kant is het helemaal niet duidelijk hoe deze extra velden de dynamica van de kosmos be invloeden.

Oplossing van de singulariteit en inflatie in dit proefschrift

Gaat snaartheorie ons helpen om de oerknal te beteugelen? Dit is het eerste vraagstuk waarover ik me in mijn proefschrift buig.

Snaartheorie heeft ons met een verbijsterend inzicht gezegend: De fysica van de snaartheorie in een bepaalde ruimtetijd kan door een kwantumveldentheorie op diens rand beschreven worden. Deze nieuwe techniek word de *AdS/CFT-correspondentie* of *ijk-gravitatie dualiteit* genoemd. De informatie over wat er in een ruimtetijd zoals onze kosmos gebeurd, kan herkregen worden door een goed begrepen theorie op diens rand te bestuderen. Gelukkig gedraagt zich deze theorie des te beter en is beter begrepen hoe sterker de zwaartekracht gekoppeld is. Merk op dat het grootste probleem met de oerknal was dat de zwaartekracht daar oneindig sterk wordt.

Ik bekijk een hele specifieke realisatie, een gesimplificeerd model van de oerknalsingulariteit en onderzoek hoe deze in de veldentheorie eruit ziet die erbij hoort. Het blijkt dat de veldentheorie de singulariteit van de gravitatie repliceert. De oerknal verschuilt zich achter een instabiliteit in de veldentheorie. Dit is echter alleen het geval in het limiet, waar de snaren door puntdeeltjes benaderd worden. Ik heb een berekening gedaan waarin met de koppeling tussen twee snaren rekening wordt gehouden, waarvan we verwachten, dat hij in het vroege universum een belangrijke bijdrage geeft. Het blijkt dat het meenemen van deze snaartheorie-effecten de veldentheorie regularizeert en de oerknal een fatsoenlijk concept maakt. Het heeft dan zelfs zin om te vragen wat er voor de oerknal was, ook al zijn we nog ver van een antwoord op dit soort vragen.

Het tweede deel van mijn proefschrift gaat over kosmische inflatie. De effectiviteit hiervan hangt heel sterk van de specifieke vorm van de potentiaal van het inflatonveld af. Het moet erg vlak zijn om de kosmische inflatie lang genoeg te drijven, totdat de kosmologische problemen opgelost zijn. Om een daadwerkelijk natuurlijke verklaring van deze fase van de kosmologie te hebben zou men de potentiaal het liefst van een fundamentele theorie zoals snaartheorie afleiden. Alleen dan zou inflatie ook echt de problemen omtrent de afstemming van parameters oplossen omdat alleen dan het fenomeen vanuit een dieper begrip natuurlijk verklaard wordt.

De proliferatie van velden, die snaartheorie door de compactificatie vertoont, maakt het heel moeilijk om een model in alle details te onderzoeken. Daarom worden meestal alle velden, die niet interessant lijken weggegooid met de veronderstelling dat men dit kan doen zonder de conclusies van het model in twijfel te trekken.

De vraag die ik samen met mijn collega's bekeken heb is of deze velden wraak nemen. We laten zien dat het effect dat deze vrijheidsgraden op het model hebben meestal ver onderschat worden. Alleen onder hele specifieke omstandigheden is het toegestaan om deze velden te verwaarlozen. In de fysisch relevante gevallen komen deze voorwaarden – andermaal – er op neer om hele specifieke voorwaarden te kiezen. Dus lijkt de afstemming, die de kosmische inflatie overbodig zou maken, alleen onder het tapijt geveegd te zijn, en door onze grondigere analyse weer tevoorschijn gehaalt.

Conclusie

Onderzoek over de kosmologie van het hele vroege heelal vanuit een fundamentele theorie zoals snaartheorie is noodzakelijk en meeslepend maar ook gevaarlijk. Aan de ene kant heeft snaartheorie nieuwe eigenschappen en technieken, die het toestaan om kosmologische vraagstukken op een kwalitatief nieuwe manier te onderzoeken. Mijn resultaten duiden aan, dat snaartheorie iets kan bijdragen aan de oplossing van de meest zware problemen. Aan de andere kant is snaartheorie moeilijk te controleren. Men moet opletten dat men met alle relevante effecten rekening heeft gehouden, die anders stilletjes de problemen weer naar voren toveren die men eigenlijk dacht opgelost te hebben.

Ondertussen zijn de vooruitgangen in de kosmologie, zoals deze, waarover dit proefschrift gaat, er goed toe in staat om de menselijke nieuwsgierigheid te bevredigen en de grens van de verkenning is weer een klein stukje verlegd.

ACKNOWLEDGEMENTS

The completion of this thesis marks not only the end of an academic examination but it is also the landmark of a very distinct and intense period of my life. The decision to do my PhD in Amsterdam was a decision for adventure – an adventure, indeed, in which a lot of people took part.

The most important ones are of course my supervisors, Koenraad Schalm and Jan Pieter van der Schaar. Thank you for trusting me with challenging projects and a lot of independence. Koenraad, I admire your ability to find the very fundamental questions and your skill to answer them. I very much appreciate to have been adopted to the Leiden group for the second life of my PhD. Thank you, Jan Pieter for your enthousiasm and in particular for the discussions on a very fun last project. It is a pity it didn't make it into the thesis, but working on it with you was enjoyable and instructive. After all *non tractato quaerimus sed quaestioni tractamus*.

Many thanks also to Erik Verlinde for having given me the opportunity to do my PhD at an engaging place and for being my promotor. Although it didn't work out to work together I very much enjoyed the discussions we had and I have learned from your intuitive way of understanding a problem. Ik neem me ook een voorbeeld aan je evenwichtige en vriendelijke persoonlijkheid, die je kenmerkt.

Not only from the financial side, FOM was an awesome employer. I doubt that there is another organisation that professional and supportive in science. Not only did they pay my bills and travels, they also provided stimulating, useful and fun courses as well as personal support. I am particularly greatful to Joost Weber in the latter respect.

I am grateful to the Lorentz-institute for theoretical phyiscs at Leiden University for hospitality during part of my research. I have always felt welcome in the almost familiar atmosphere and I have enjoyed the opportunity to collaborate and discuss, in particular with Ted van der Aalst, Sjoerd Hardeman, Gonzalo Palma and on the second project with Ana Achúcharro. Ted, het was altijd onwijs gezellig, vrolijk en rijk aan inzicht om met je samen te werken, te "babbelen" en de wereld te begrijpen. Uiteraard is carneval na Waalre nooit meer hetzelfde. Sjoerd, je bent onvergeten. Bedankt, als dat nog zou kunnen, voor het delen van je inzichten, je enthousiasme en je gezelligheid. Thanks Gonzalo for allowing us to finish the project you had begun but in particular for our discussions at the beginning of my PhD. Ana, thanks for endless discussions, your advise, your friendliness and your concern about more or less everything. You showed me how persistent one should be in pursueing a question and a deep devotion to science and humanity.

The institute for theoretical physics in Amsterdam was and is an infamous place, that attracts a lot of amazing people. I enjoyed some very illuminating discussions with Jan de Boer, Adam Bzowski, Borun Chowdhury, Sebas Eliens, Leo van den Horn, Ingmar Kanitscheider, Erik Laenen, Paul de Lange, Paul McFadden, Daan Meerburg, Bernard Nienhuis, Theo Nieuwenhuizen, Milosz Panfil, Debabrata Panja, Kyriakos Papadodimas, Balt van Rees, Masaki Shigemori, Sheer El-Showk, Jan Smit, Marika Taylor and Kostas Skenderis. Also personally, I enjoyed the company of all the fellow PhD students and postdocs. In academic generations, a PhD takes a long time and one has an amazing amount of colleagues. Thanks to the people that welcomed me so friendly to the institute, Amir, Ardian, Bahar, Balt, Ingmar, Jan, Joao, Kyriakos, Leo, Liza, Lotte, Masaki, Meindert, Paul, Ruth, Sheer, Susanne, Thomas, Xerxes, and those whom I welcomed, Adam, Benjamin, Benoit, Bram, Borun, Daniel, Francis, Geoffrey, Goffredo, magic György, Jacopo, Jelena, the industrious Juan, Keiichi, Keun-Young, Kris, Marco, Maria, Michal, Milena, Voditel Milosz, Natascha, Paul, Ricardo, Sebas, Shanna, Tessa, Yegor. Thanks for the time together.

I had the great opportunity to travel a lot and to meet intersting people all over the world. I gratefully acknowledge discussions with Niayesh Afshordi, Nima Arkani-Hamed, Erik Bergshoef, Raphael Bousso, Robert Brandenberger, Ben Craps, Sera Cremonini, Scott Dodelson, John Duncan, Richard Easther, Marc Henneaux, Thomas Hertog, Jan-Willem van Holten, Dan Hooper, Veronica Hubeny, Gary Horowitz, Mark Jackson, Hendrik Jennen, John McGreevy, Juan Maldacena, David Mateos, Gustavo Niz, Jin-Ouk Ong, Elisabeta Pallante, Jan Perz, Marieke Postma, Mukund Rangamani, Diederik Roest, Blaise Rollier, Graham Ross, Slava Rychkov, Subir Sachdev, Savdeep Sethi, Konstadinos Sfetos, Bert Schellekens, Gary Shiu, Eva Silverstein, Martin Sloth, Gianmassimo Tasinato, Rob Timmermans, Doug Spolyar, James Stokes, Kip Thorne, Hermann Verlinde, Robert Wald, David Wands, Scott Wattson, Bernard de Wit, Edward Witten, Amos Yarom and possibly many more along my path.

The Dutch physics community is a very close one. I very much appreciated all the local lectures, conferences and extra events during which we could share experiences and ideas. In particular, I had a lot of fun every time I went to visit the friends in Groningen, Roel and Giuseppe amongst others, just across the road at NIKHEF, Michele, Reinier, Lisa, the Cosmopolitans Sander and Jan and in Utrecht, Anne, Hugo, Jasper, Maarten, Marius.

Mocht er eens verder niemand op het instituut aanwezig zijn, bij hen kan je steeds terecht voor lekkere koffie en steun: onze secretaresses begonnen bij de moederlijke Paula en Lotty, Bianca die altijd advies in moeilijke kwesties heeft, Jonneke, Sandra, Joycelyne, en uiteindelijk Anne-Marieke, bijzonder bedankt voor het opknappen van alle klusjes aan (en na) het einde van mijn promotie. Ook zeg ik dank aan het management, vooral Yocklang, steeds om het welzijn van iedereen bezorgd, Luuk en Joost.

Since life goes on even after a PhD, I am very grateful to Sander Bais and Herbert Wagner for very valuable career advise, that encouraged me to explore new fields of research. I would also like to thank Jean-Sebastian Caux for encouragement to persue my chosen path.

Around the institute, there were quite a few wonderful people with whom it was a lot of fun to do valuable things. First of all, there were the people from the social club. I deeply believe that the only purpose of an institute for theoretical physics is to bring people together so that the exchange of their thoughts and discussions will foster their insights and improve their research. Occasional social events are essential to achieve this goal. It was good to have the supersocials Balt, Liza, Shanna, Jesper, Joost, Richard and Witteke to join the efforts for creating a pleasant atmosphere at the institute. I think that we had some very nice and fun ideas.

I am also very happy that I could organise the reading course on conformal field theory together with Liza and Jesper with the help of the DRSTP. A special thanks to the lecturers Erik Verlinde, Thomas Quella, Jan de Boer, Kareljan Schoutens, Gleb Arutyunov and Bert Schellekens, who all did a marvellous job and have deepened at least my understanding of CFT.

I think that outreach of all sorts is not only important for the survival of our trade but also a fun thing that scientists should engage in. Shanna, Lisa, Lucas, we have found a particularly special way of doing so by setting up the nerdnite. I am proud that we managed to engage such an interesting variety of amazing speakers. Dank jullie allemaal voor de plezierige tijd die we er samen mee hebben gehad. Ik vind het ontzettend leuk, dat de nerdnite nog steeds zo populair is en volgens mij alleen maar populairder wordt. Ook CREA en David Rozema van harte bedankt voor de steun van het project. Het is fijn om het bij jullie in goede en betrokkene handen te weten en ik vind het erg jammer, dat ik er niet meer aan mee kan doen.

Obviously, things don't always run smoothly during a PhD. In such times, support is a valuable thing to get. I am very grateful to Yocklang Chong, Jan de Boer, Manuala Kulaxizi, Joke Takoma, Theo Nieuwenhuizen and Kareljan Schoutens, Jelena and Milena for this support. I am also thanking Marco Baggio and Christian Tan for bringing me to mindfullness meditation, which I think should be taught to every PhD student.

Housing in Amsterdam is a problem. Ik had ontzettend geluk met mijn reeks appartementen, die alleen maar beter werden. Mijn dank aan DUWO voor een heel spontaan, de Alliantie voor een heel knus en tenslotte Piet Sagel en Theo voor een bijna vorstelijk onderkomen.

"Es sungen drei Engel einen süßen Gesang, mit Freuden es selig in dem Himmel klang." Het leukste in Amsterdam was gewoon het CREAorkest. Het heeft een tijdje geduurd voordat ik erop kwam dat ik bij jullie moest zijn. Maar de drie jaren, die ik bij jullie mee mocht spelen vallen waarschijnlijk hooguit in muziek te vatten. Onze uitvoering van Mahler's derde symphonie in het Concertgebouw was zeker een hoogtepunt van mijn tijd in Amsterdam. Ik heb van elke repetitie, van elk concert en van elke borrel genoten. Wij – durf ik nog steeds te zeggen – zijn een vriendelijk orkest. Iedereen die ik daar heb leren kennen was op zijn manier waardevol en verrijkend voor me. Ik wil daarom het orkest als éen groot instrument beschouwen en jullie allemaal samen bedanken. Maar ik kan het niet laten om plaatsvervangend Bas Pollard, onze grote maestro op het voetstuk te zetten, waarop hij ons elke repetitie voorstond. Bas, ik heb onzettend veel van je geleerd, over de muziek, over enthousiasme, over uithoudingsvermogen, over leiderschap en over vriendschap. Het was altijd erg leuk. Ik kan jullie maar moeilijk missen, want “all’ Lust will Ewigkeit, will tiefe, tiefe Ewigkeit!”

Leaving Amsterdam, I also leave new friends. Joost en Ingrid, Ted, Thomas und Liu Na, Jelena and Milena, Richard and Kat, Ernst en Vincent, Nicole en Gideon, Mark, Nora, Ilaan, Ingmar, Jeroen, Jorien, Joukje en Remko. Thank you for having been there for me in the past years and please, continue to do so in the future. Nicht nur bei meinen neuen Freunden, sondern auch bei all jenen will ich mich bedanken, die mir mit ihrer Freundschaft in der Ferne und manchmal auch in der Stille treu geblieben sind.

Liebe Mami, lieber Papi, Ihr unterstützt mich seit jeher in all meinen Plänen und Ideen. Euer fester Glaube an mich und Eure Ermutigung war nie so wichtig, wie in den letzten fünf Jahren. Danke für alles, was Ihr mir bis jetzt gegeben habt. Liebe Heike, Martin, Tobias, Johanna, wir haben einander in den letzten Jahren viel zu wenig gesehen. Ich danke Euch für Eure Besuche und für Euer stetes Mitleben an meinen Entwicklungen. Wat ik met mijn eigen familie gemist heb, heb ik met jullie hartelijkheid weer terug gekregen. Bedankt, Rinske, Klaas, Annemiek, Maarten, Marjon en Mark voor de warmte en dat ik welkom ben in jullie midden.

Lieve Maaike, bijna de hele weg sinds het begin van mijn promotie ben je met mij gegaan. Je had met mij heel veel te verduren in deze tijd waarin je altijd mijn vaste steun was. We hebben een avontuur achter ons en nog zo velen voor ons. Ik weet dat ik het met jou allemaal aan kan. Je maakt mijn leven zo rijk. Dank je voor alles.

CURRICULUM VITAE

Johannes Michael Oberreuter was born in Gräfelfing near Munich and grew up in Passau in Lower Bavaria, Germany. He undertook his pre-univeristy education at the Gymnasium Leopoldinum in Passau, which he finished with the highest possible mark. He subsequently studied physics, mathematics, cemistry and philosophy at the Ludwig-Maximilians-University in Munich while he was a scholar of both the Studienstiftung des deutschen Volkes (National German Merit Foundation) and the Bavarian scholarship for highly gifted students as well as e-fellows. Already after two years at the university, he took the infamous part III of Mathematical Tripos at the University of Cambridge, where he is a member of Girton College. Restrospectively, he got awarded the degree of a Master of Advanced Study in Mathematics from the University of Cambridge in 2011. He went on to study physics and philosophy at the Ruprecht-Karls-University Heidelberg and finished his Diploma in Physics at the Ludwig-Maximilians-University in Munich under the supervision of Dieter Lüst and Gabriel Lopes Cardoso with a thesis titled *The black hole attractor mechanism*. He undertook research internships with Anton Zeilinger at the University of Vienna, Austria, Bengt Nielsson at Chalmers University in Göteborg, Sweden, and Ignacio Cirac and Michael Wolf at the Max-Planck-Institute for Quantum Optics, Garching, Germany.

After a short interlude as a business consultant at Roland Berger Strategy Consultants, he went on to do research at the University of Amsterdam under the supervision of Erik Verlinde, Koenraad Schalm and Jan Pieter van der Schaar. This research led to the present thesis.

He has reported his work at the Univerity of Groningen, the Institute of Advanced Study in Princeton, McGill University in Montreal, Fermilab in Chicago, the International Center for Theoretical Physics in Trieste, Imperial College London, Oxford University, The Institute of Cosmology and Gravitation at Portsmouth at the Iberian Strings Conference in Valencia, the Physics@FOM meeting in Veldhoven, the Aegean Summer school in Milos and the Santa Fe Institute. He has attended the Northeast Cosmology workshop at McGill university, the school Prospects of Theoretical Physics, Frontiers of Physics in Cosmology at the Institute of Advanced

Study in Princeton, the Iberian Strings in Valencia, The Amsterdam Summer workshop on string theory in 2008, 2010 and 2012, the Spring school on superstring theory and related topics at the ICTP, Trieste, the Postgraduate school in theoretical high energy physics at Driebergen, which he fondly remembers, the PhD school Brussels-Paris-Amsterdam, Integrability in gauge and string theory at the University of Utrecht.

During his PhD, he was involved in several teaching activities such as tutoring for statistical physics, electrodynamics, classical mechanics and from classical to quantum. He has been voted teaching assitant of the year in 2010. He has also supervised a bachelor's thesis (Magnetic monopoles in gauge theories). Furthermore, he was one of the organizers of the reading course on conformal field theory for the Dutch Research School for Theoretical Physics. He has been teaching a course on Paradoxa in Physics twice for the Deutsche Schülerakademie.

Since the completion of his PhD, he is a postdoc at the Georg-August-University of Göttingen in the field of non-equilibrium dynamics of quantum many-body systems in the group of Stefan Kehrein.