Instantons and Cosmologies in String Theory

Andrés Collinucci

Para mamá
Per papà

The work presented in this thesis was performed at the Centre for Theoretical Physics of the Rijks*universiteit* Groningen, in The Netherlands.

Printed by Universal Press - Science Publishers / Veenendaal, The Netherlands.

Rijks*universiteit* Groningen

Instantons and Cosmologies in String Theory

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijks*universiteit* Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
vrijdag 7 oktober 2005
om 13.15 uur

door

Giulio Andrea Collinucci

geboren op 10 juli 1977
te Genève

Promotor: Prof. dr. E. A. Bergshoeff

Beoordelingscommissie: Prof. dr. T. Ortín
Prof. dr. A. Van Proeyen
Prof. dr. J. G. Russo

ISBN: 90-367-2360-4

Contents

Introduction

If we knew what we were doing, it wouldn't be called research, would it?
A. Einstein

Unification is one of the driving principles of modern theoretical physics. Maxwell showed us around the 1860's that electricity and magnetism were not to be thought of as separate forces, but as two different manifestations of the same entity, *electromagnetism*. Later on, in the beginning of the twentieth century, Einstein asked the famous question: "What do you see if you chase a ray of light? Can you see it in its rest frame?" Although the question was not motivated by any unexplained experimental results, the answer led to a great revolution in physics. To answer his question, Einstein had to find a way to combine the framework Galilean mechanics with Maxwell's electromagnetism. This led to *special relativity*, and eventually *general relativity*, which completely modified the way we view space, time, and gravity.

In the early twentieth century, quantum mechanics was well-established as *the* theory to describe the electron in the atom. As experiments at the subatomic level became more sophisticated, being able to collide subatomic particles at higher and higher energies, it was noticed that the number of particles is not conserved in a collision process. A new theory was needed that could at the same time deal with the fact that particles are small, and hence quantum mechanical, and highly relativistic, traveling at speeds comparable to the speed of light. This led to the development of *relativistic quantum mechanics* and eventually *quantum field theory*, between the late 1920's and the 1950's. The latter is built to incorporate special relativity and quantum mechanics in one framework. Because it is relativistic, quantum field theory treats mass and energy as a single entity. Consequently, it no longer requires the conservation of the particle number in a process, as long as the total energy/mass at the end of a process is the same as that at the beginning. Because it is quantum mechanical, it also allows for a temporary violation of energy/mass conservation, leading to the off-shell intermediate states that one sees in Feynman diagrams.

The ultimate success of quantum field theory came from its concrete application to the *standard model* of particles. This model, which was developed in the 1970's, describes three of the four fundamental forces of nature:

- **Electromagnetism**: this force is responsible for most of the phenomena we observe in our lives besides gravity, such as the electric repulsion that keeps solid objects from simply merging into each other, and the fact that we can chat on cellular phones.

- **The strong nuclear force**: this is what keeps nuclei from flying apart due to their electric repulsion.
- **The weak nuclear force**: this force is responsible for radioactivity.

The standard model describes these three forces and the particles that are *charged* under them in a single gauge theory with $SU(3) \times SU(2) \times U(1)$ as its symmetry group. This theory has been experimentally confirmed within its regime of validity beyond a shadow of doubt.

We could now ask a question that is not *yet* motivated by unexplained observational data, but is in the spirit of unification: "What takes place inside a black hole?" The first notion of a black hole was discovered by Schwarzschild, as the first solution to the Einstein equations ever written down. A lot of efforts have been made, and are still being made, in order to understand the real physical meaning of this mathematical solution. What is interesting about black holes, is that they provide us with a Gedankenexperiment that forces general relativity and quantum mechanics together. The former is necessary because it is the framework for strong gravitational fields, whereas the latter is necessary because black holes are made of matter that is compressed to a very small space. This is where we notice the shortcomings of quantum field theory, and general relativity. They are seemingly incompatible. Although a theory of quantum gravity does not yet exist, there are two candidate theories: string theory, and loop quantum gravity. In this thesis, we will work with string theory.

String theory is an attempt to describe very high energy densities such as the inside of a black hole. However, it is more ambitious than that. It also has the potential to unify gravity with the other aforementioned forces of nature into one single framework, which would be valid in all possible regimes of energy and size. String theorists hope to formulate a *theory of everything*.

The theory is derived from the very simple idea that fundamental particles, which were always thought of as points (i.e. objects of zero size), are actually tiny vibrating strings of Planck length size (i.e. $\sim 10^{-35}$ m). The strings do not have fixed length, but a fixed tension, or energy density. This means that the mass of any given string is determined by its vibrational state. For instance, if it spins really fast, it will tend to stretch by centrifugal force, and will have a higher mass than a string that does not spin. Whereas a particle cannot have angular momentum, but only intrinsic spin, a string does have angular momentum. So string theory regards all different kinds of particles as being made out of the same 'fabric', and properties such as spin and mass are no longer intrinsic[1], but simply labels of the states in which the strings are. Trying to formulate a quantum theory of a relativistic strings (special or general relativistic), creates a world of mathematical structure that is both beautiful and complicated.

Needless to say, the path toward such an ambitious goal as formulating the theory of everything is plagued with obstacles. Although the theory has been around for several decades, as of this writing, it is still in its infancy. One might even say that string theory has so far made bigger contributions to mathematics than to physics. A major drawback of string theory is that it is only defined *perturbatively*. This means one has to assume that strings interact *weakly* with each other, in order to even define the theory. In order to be able to perform calculations, however, one often has to make one more approximation: the low energy approximation. This approximation requires that one only consider the massless states of the string. It also requires

[1] However, the difference between fermionic and bosonic strings is in some sense still intrinsic.

that spacetime curvature be weak. Once those criteria are met, one can treat string theory as a field theory. To be specific, the field theories used to approximate string theory are called *supergravities*. Throughout this thesis, we will be working with this approximation.

In chapter 3, we will discuss D-instantons. These are objects that arise in the supergravity approximation of string theory, yet they can actually provide us with *non-perturbative* information about string theory, i.e. they show effects that cannot be found by means of naïve perturbation theory. They are analogous to instantons in ordinary field theory in that they can only be found in the Euclidean formulation of the path integral. The D-instanton can be interpreted as a quantum field theoretic tunneling amplitude between two states of the spacetime metric, and the *axion-dilaton* scalar of type IIB supergravity. It yields a non-perturbative contribution to the calculation of the path integral. We will be studying a non-supersymmetric kind of D-instanton. We will show its relation to the better known supersymmetric D-instanton in terms of the $SL(2,\mathbb{R})$ duality symmetry of type IIB supergravity. We will also show how the general D-instanton can be viewed as a spatial section of a charged black hole, one dimension higher.

Another challenge of string theory is that it manifests itself in different forms. Until the mid 1990's, there were actually five different consistent formulations of string theory, which was very unsettling for those who believed it to be a theory of everything. However, in the mid 1990's, Edward Witten and other physicists showed that these five theories, together with eleven-dimensional supergravity (a bonus theory, so to speak), were actually different limits of one unique theory now known as *M-theory*. Unfortunately, not much is known about M-theory itself. Even the origin of its name is a mystery. One often illustrates this novel understanding of string theory by drawing a hexagon, where the corners represent all six limiting cases of M-theory, the latter being represented by the content of the polygon. The six theories are related to each other via so-called *dualities*. A duality can be thought of as the abstract generalization of the Fourier transform. Fourier transforming a differential equation means writing down the same problem in different variables, according to a certain map. A problem that seems impossible in one set of variables, can be a one-line calculation in the new variables. String theory dualities relate different theories in their opposite regimes, or sometimes they relate a theory to itself. For instance, type IIB string theory is *S-dual* to itself. This S-duality manifests itself via the action of the group $SL(2,\mathbb{Z})$ on the degrees of freedom of the theory, and it sometimes maps weakly coupled string theory to its strongly coupled counterpart.

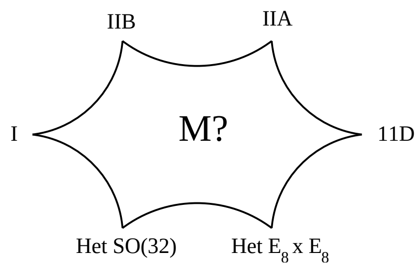


Figure 1: *The five string theories, 11-D supergravity, and M-theory.*

Although this picture shows us what we know about string and M-theory, it mainly shows us what we do not yet know. We know only the six corners of the hexagon, i.e. the extreme regimes of these theories. Everything between the corners is uncharted territory.

Another interesting discovery of the 1990's is Maldacena's AdS/CFT conjecture [1]. The latter asserts that type IIB string theory in a certain background geometry, namely $AdS_5 \times S^5$, is fully equivalent to supersymmetric Yang-Mills theory in four dimensions ($\mathcal{N} = 4, d = 4$ SYM). This is another example of a strong/weak duality. It relates the two theories in their opposite regimes of coupling strength. Therefore, this is useful for exploring the weakly and strongly coupled phases of both theories, but not the intermediate phases.

The lack of a viable framework for M-theory prevents us from deriving the laws of physics that govern strings. String theory, as it is currently formulated, does not tell us in what spacetime manifold we actually live. It treats the spacetime metric as a non-dynamical background, and imposes the Einstein equation on it as a consistency condition that restricts the kinds of allowed manifolds. In order to be able to pinpoint a unique background for string theory, one would need a theory with a *vacuum selection principle*. This is analogous to having a system with degenerate vacua and no potential that can lift the degeneracy. The best one can do is to look for backgrounds through trial and error, and see which ones are most consistent with the physical world we live in. A new emergent philosophy among string theorists, the so-called *landscape* scenario, suggests that there is no vacuum selection principle, but that all possible universes actually coexist as *bubbles* in a *megaverse*. According to this picture, we happen to live in one of the few universes where the constants of nature are such that life is possible, but other universes where it is not possible also exist. However, these are causally disconnected from ours.

So far, no verifiable or falsifiable prediction has been made by string theory. This is due to two reasons: first, technological limitations make it impossible to measure any string effects in a particle accelerator. Second, even if particle accelerators were capable of making measurements at an arbitrary energy level, string theory has not told us yet what we would see, due to its complicated nature.

Recently, however, hopes of getting string theory to make contact with reality have been revived by cosmology. First of all, cosmological processes such as supernovae are the ultimate particle accelerators, reaching energies far higher than CERN could ever dream of. Secondly, recent measurements have confirmed that our universe is undergoing a period of *accelerated expansion*. This provides string theorists with the challenge/opportunity to derive a scenario from string theory that produces accelerated expansion that is consistent with observations.

In chapters 5 and 6, we will be studying a certain class of cosmological models containing Einstein gravity and scalar fields, some of which are derivable from string theory, and some with yet unknown fundamental origins. We will specifically see when these models lead to accelerating universes, be it *eternal* or *transient* acceleration.

The Big Bang scenario, which is a widely accepted account for the early history of our universe, states that the latter was once very dense and hot, emitting perfect *blackbody radiation*. The microwave spectrum of this radiation, the famous *Cosmic Microwave Background Radiation*, has been observed and thoroughly studied, and is consistent with the Big Bang theory. Earlier in this introduction, I mentioned that black holes provided us with a 'theoretical laboratory' in which to study quantum gravity. The Big Bang is actually a real life laboratory for quantum gravity, as it describes a very dense, and hence highly curved spacetime, where short-distance physics is dominant. If string theory is a theory of everything, it must ultimately explain and 'smooth out' the Big Bang singularity.

This thesis is organized as follows: chapter 1 is a basic introduction to the bosonic string, and its quantization. There, I will also briefly explain the conformal field theory approach to string theory, and the physical interpretation of spacetime backgrounds. Finally, a brief summary of superstring and supergravity theories will be provided.

In chapter 2, I will introduce instantons in quantum mechanics and field theory, thereby explaining the semiclassical approximation in a Euclidean signature. This will be illustrated with examples, including the Yang-Mills instanton. Then, I will present a brief introduction to solitons. Finally, I will explain the correspondence between solitons and instantons.

Chapter 3, which is based on a publication, concerns type IIB non-extremal D-instantons. First, I will review the $SL(2,\mathbb{R})$ symmetry of type IIB supergravity and generalize to arbitrary dimensions and dilaton coupling. Later, this theory will also be generalized to theories with multiple scalars. Then, the solutions will be presented, as well as their $SL(2,\mathbb{R})$ properties. After a brief introduction to Euclidean wormholes, we will see that one class of solutions gives rise to such geometries. In analogy with the soliton-instanton correspondence explained in chapter 2, a correspondence between D-instantons and charged black holes; and D-instantons and p-branes will be established. The calculation of the action for these instanton solutions will be presented, alongside with a discussion about the potential quantum mechanical role of non-extremal D-instantons in string theory. Finally, I will comment on some work in progress, where these D-instantons are put to work in the AdS/CFT context.

In chapter 4, I will give a basic introduction to modern cosmology and its issues. I will begin by introducing the *Friedmann-Lemaître-Robertson-Walker* metric and the standard terminology for the matter and energy content of the universe. Then, I will review three main problems in cosmology: the *horizon, flatness*, and *relics* problems, and we will see how these are solved by inflation. I will then discuss present day acceleration, mentioning some of the current methods being used by string theorists to *derive* it.

The goal of chapter 5, which is based on a publication, will be to describe gravity-scalar models for cosmology with single-exponential potentials. We will see that these systems can be formulated as autonomous systems, and that power-law and de Sitter solutions can be regarded as critical points. We will then analyze the solutions that interpolate between critical points, paying attention to trajectories that have periods of acceleration.

In chapter 6, we will generalize on the previous chapter by analyzing multiple-exponential potentials. This chapter is based on a publication, in which the critical points are given for the most general case for the first time. This analysis is novel in that it includes cases that are even more general than what is known as 'generalized assisted inflation'. The analysis will be illustrated by some examples of potentials with higher-dimensional origins via compactifications of gravity over three-dimensional *group manifolds*.

Just as instantons and solitons have similar mathematical structures, D-instantons and FLRW cosmologies are also mathematically similar. They are both gravity-scalar configurations that depend on only one coordinate (be it time-like or space-like). They both asymptote to 'trivial' configurations, but have non-trivial interpolating behavior, much like kink solutions. They can probably be viewed as sections of non-trivial bundles over the circle. In chapter 7, this parallelism will be pursued in two ways: first, we will see that some D-instantons can be related to cosmologies via Wick rotation. Then, we will see that D-instantons and scalar cosmologies can be viewed as the trajectories of particles in a fictitious *scalar manifold* or *target space*. This interpretation not only puts these solutions on equal footing, it even patches them as two portions of the same trajectory. We will see how this suggests a possible resolution of the Big Bang singularity, by means of smooth Big Crunch to Big Bang transitions that have an intermediate Euclidean period.

Chapter 1

String theory in a nutshell

1.1 Introduction

In this chapter, the basic definitions and foundations of string theory will be laid. We will start by reviewing the relativistic point-particle in the formalism of the variational principle. Then, we will repeat this for the relativistic bosonic string. After a brief introduction into the canonical quantization of the string and the resulting spectrum, we will study the string in the path integral formalism. The notions of vertex operators, and the genus expansion of string Feynman diagrams will be introduced. This will allow us to understand how non-trivial spacetime backgrounds, on which the string can propagate, can be interpreted as coherent states of strings. Then, we will briefly see that requiring classical symmetries to hold quantum mechanically imposes constraints on spacetime backgrounds by means of β-functions. In the low energy approximation, these constraints can be interpreted as spacetime field theories. Field theories obtained as low energy approximations to string theory will be the main framework of this thesis. Finally, a brief summary of supersymmetric string theories and their low energy limits will be provided.

In the following, I will be borrowing heavily (and sometimes verbatim) from Polchinski's textbooks [2, 3]. However, the philosophy behind this chapter is *not* to provide the reader with yet another carbon copy of the standard textbooks, and certainly not to improve on the latter. The main goal of this chapter is to show a minimal selection from the standard textbooks in order to schematically explain how the low energy limit of the quantized theory of relativistic strings (which is a QFT in the two world-sheet dimensions) is a classical field theory in spacetime.

1.1.1 The relativistic point-particle

Before we begin our journey into the theory of strings, let us review our knowledge of relativistic point-particles through the action principle.

To describe the motion of a particle moving in a D-dimensional Minkowski spacetime we can define $D-1$ functions of time $X^1(X^0), \ldots, X^{D-1}(X^0)$, which give the particle's position in space at any given time X^0. We can also make this description covariant by parametrizing the particle's *world-line* with a variable τ, such that we now have D functions $X^0(\tau), \ldots, X^{D-1}(\tau)$ on equal footing. One can derive the equations of motion from the variational principle through the following action:

$$S = -m \int d\tau \, (-\dot{X}^\mu(\tau)\, \dot{X}_\mu(\tau))^{1/2}, \tag{1.1}$$

where m is the particle's mass, and the dot represents a τ-derivative. This action measures the relativistically invariant arc-length (or proper time) of the world-line, and the classical particle will move along the trajectory that extremizes this quantity. The Euler-Lagrange equations for the X^μ's are then,

$$\partial_\tau \left(\frac{m\, \dot{X}^\mu}{(-\dot{X}^\mu\, \dot{X}_\mu)^{1/2}} \right) = 0\,. \tag{1.2}$$

The conjugate momenta to the particle's spacetime coordinates are the following:

$$P^\mu = \frac{m\, \dot{X}^\mu}{(-\dot{X}^\mu\, \dot{X}_\mu)^{1/2}}\,, \tag{1.3}$$

from which we easily derive the on-mass-shell constraint

$$P^2 + m^2 = 0. \tag{1.4}$$

Although this action allows for an easy derivation of the classical equations of motion and on-shell condition, it does not accommodate the case of the massless particle. Moreover, the square root of the integrand makes this action awkward to work with in a path integral calculation. Fortunately, there is a more convenient form which eliminates these two features by introducing an auxiliary field:

$$S' = \tfrac{1}{2} \int d\tau \, (e^{-1}(\tau)\, \dot{X}^\mu(\tau)\, \dot{X}_\mu(\tau) - e(\tau)\, m^2)\,. \tag{1.5}$$

The auxiliary field $e(\tau)$ is the world-line *einbein*. In other words, it is the square root of (minus) the metric $g_{\tau\tau}(\tau) = -e(\tau)^2$ that lives on the one-dimensional τ-space. This metric is an independent field and is therefore *not* induced by the spacetime Minkowski metric $g_{\mu\nu}$. The first property we should establish about this action is that it is equivalent to the previous one (1.1) (except for the massless case). To show this we compute the equations of motion of the *einbein*:

$$m^2\, e^2 + \dot{X}^\mu\, \dot{X}_\mu = 0\,. \tag{1.6}$$

Substituting this back into (1.5) we find that $S = S'$. Notice also that the conjugate momenta are now given by

$$P^\mu = \frac{\dot{X}^\mu}{e}\,, \tag{1.7}$$

which, combined with (1.6) gives the on-mass-shell constraint as an equation of motion.

Let us list the symmetries of the action (1.5):

- D-dimensional spacetime Poincaré transformations:

$$X^\mu \to X'^\mu = \Lambda^\mu{}_\nu\, X^\nu + A^\mu\,, \tag{1.8}$$

where $\Lambda^\mu{}_\nu$ is an SO$(1, D-1)$ matrix and A^μ is an arbitrary D-vector.

- World-line reparametrizations:

$$\begin{aligned} \tau &\to \tau' \\ e(\tau) &\to e'(\tau') = e(\tau)\,\frac{d\tau}{d\tau'} \\ X^\mu(\tau) &\to X^\mu(\tau')\,. \end{aligned} \tag{1.9}$$

This action is by construction Poincaré invariant. The second symmetry merely confirms the fact that the physics of a particle should be independent of how one chooses to parametrize its world-line. Notice that we could make a paradigm shift and regard this system (1.5) as a one-dimensional theory of D scalar fields $X^\mu(\tau)$ and a metric $g_{\tau\tau}(\tau) = -e(\tau)^2$. In that case, the D-dimensional Poincaré symmetry would be interpreted as an internal symmetry of the scalar fields, and world-line reparametrization invariance would be seen as invariance under general coordinate transformations in one dimension. Although this interpretation may appear strange in this case, this point of view will prove to be a very powerful tool in string theory.

1.1.2 The relativistic bosonic string

Now we are ready to deal with the bosonic string. We will proceed by analogy with the case of the particle. A particle sweeps out a world-*line* in spacetime, which means that we can describe it as an embedding of a one-dimensional manifold into a D-dimensional Minkowski spacetime. A string sweeps out a two dimensional world-*sheet*, this requires an embedding of a two dimensional manifold into D-dimensional Minkowski spacetime. The string coordinates will then be functions of two parameters $X^\mu(\tau, \sigma)$. We can derive equations of motion for the string by requiring that the world-sheet extremize its invariant surface. In order to measure that surface we define the *induced* metric on the world-sheet h_{ab}, where a, b run over the world-sheet indices:

$$h_{ab} = \partial_a X^\mu\, \partial_b X^\nu\, \eta_{\mu\nu}\,. \tag{1.10}$$

Then, the string will extremize the so-called Nambu-Goto action:

$$S_{NG} = -\frac{1}{2\pi\,\alpha'}\int d\tau\, d\sigma\, (-\det h_{ab})^{1/2}\,. \tag{1.11}$$

In the case of the point particle we needed a constant with units of energy to make the action dimensionless (i.e. the mass), in this case we need energy per unit length. Hence, the constant $1/(2\pi\,\alpha')$ will play the role of the string tension.

Once again, we can derive equations of motion from this action; however, if we expect to use it in a path integral formalism we should find an action without a square root. In order to achieve this we must again introduce an auxiliary world-sheet metric γ_{ab}. The action we are after is called the Brink-Di Vecchia-Howe-Deser-Zumino action [4, 5] or Polyakov action [6, 7]:

$$S_P = -\frac{1}{4\pi\,\alpha'}\int d\tau\, d\sigma\, (-\gamma)^{1/2}\, \gamma^{ab}\, \partial_a X^\mu\, \partial_b X_\mu\,, \tag{1.12}$$

where $\gamma = \det \gamma_{ab}$. This action has a more familiar kinetic term for the X^μ, which makes the path integral easy to evaluate. If we eliminate the auxiliary metric γ_{ab} from this action by using its equations of motion, we will find that the Polyakov action is equivalent to the Nambu-Goto action (1.11).

Let us list the symmetries of the Polyakov action:

- Poincaré transformations in D-dimensional spacetime:

$$X^\mu \rightarrow X'^\mu = \Lambda^\mu{}_\nu X^\nu + A^\mu \,, \tag{1.13}$$

  where $\Lambda^\mu{}_\nu$ is an $SO(1, D-1)$ matrix and A^μ is an arbitrary D-vector.

- World-sheet reparametrizations:
  Defining a generalized world-sheet coordinate $\sigma^a = (\tau, \sigma)$ for $a = 0, 1$ we have,

$$\begin{aligned} \sigma^a &\rightarrow \sigma'^a(\tau, \sigma), \\ X^\mu(\tau, \sigma) &\rightarrow X^\mu(\tau', \sigma'), \\ \gamma_{ab} &\rightarrow \gamma'_{cd}(\tau', \sigma') \frac{\partial \sigma^c}{\partial \sigma'^a} \frac{\partial \sigma^d}{\partial \sigma'^b} \,. \end{aligned} \tag{1.14}$$

- World-sheet Weyl rescalings:

$$\gamma_{ab} \rightarrow \gamma'_{ab} = e^{2\,\omega(\tau,\sigma)} \gamma_{ab} \,. \tag{1.15}$$

The first two symmetries, (1.13) and (1.14) are analogous to the point-particle symmetries, (1.8) and (1.9). The last one (1.15), however, is specifically due to the fact that we are dealing with a two dimensional extended object. This symmetry tells us that we should regard all Weyl-equivalent metrics on the world-sheet as physically equivalent. From the two dimensional point of view, we have a scalar field theory with an internal D-dimensional Poincaré invariance, Weyl-rescaling invariance, and invariance under two-dimensional general coordinate transformation. This field theory falls under the category of *conformal field theories*. Two dimensional CFT's are a very special kind of CFT when it comes to doing both classical *and* quantum computations, due to techniques that exist only for two dimensions. This is analogous to the fact that there are much more powerful techniques to do analysis on the complex plane than there are for higher dimensional complex spaces.

Let us now write down and solve the equations of motion for the Polyakov action (1.12). Varying the string coordinates X^μ, we get the following equation:

$$\begin{aligned} \delta S_P = \frac{1}{2\pi\alpha'} \int d\tau\, d\sigma\, \partial_a \{ (-\gamma)^{1/2} \gamma^{ab} \partial_b X^\mu \} \,\delta X^\mu \\ - \frac{1}{2\pi\alpha'} \int d\tau\, (-\gamma)^{1/2} \partial_\sigma X_\mu \,\delta X^\mu \Big|_{\sigma=0}^{\sigma=l} \,. \end{aligned} \tag{1.16}$$

To make this variation zero both terms must vanish independently. The first term requires the two-dimensional Laplacian of the X^μ's to vanish. The second term requires a choice of boundary conditions, for which there are three possibilities:

- Open string Neumann b.c.s:

$$\partial_\sigma X^\mu(\tau, 0) = \partial_\sigma X^\mu(\tau, l) = 0 \,. \tag{1.17}$$

These conditions imply that no momentum flows in or out through the string endpoints, and, hence, that these move freely.

- Open string Dirichlet b.c.s:

$$\delta X^\mu(\tau, 0) = \delta X^\mu(\tau, l) = 0\,. \tag{1.18}$$

These conditions mean that we are fixing the string endpoints and no longer consider them as dynamical.

- Closed string (periodic b.c.s):

$$X^\mu(\tau, 0) = X^\mu(\tau, l)\,. \tag{1.19}$$

This is the requirement that the string be closed, i.e. that it have no endpoints.

For open strings, the Neumann boundary conditions (1.17) are the only conditions that are consistent with spacetime Poincaré invariance, whereas the Dirichlet b.c.'s (1.18) explicitly break it. For instance, if one imposes Neumann b.c.s on $D-p-1$ string coordinates and Dirichlet b.c.s on $p+1$ of the coordinates, this means that the string endpoints are stuck to a $p+1$-dimensional hypersurface of spacetime called Dp-brane, where the 'D' stands for Dirichlet. That's why the latter were discarded for a long time as unphysical until Polchinski discovered in 1995 [8] that D-branes are an integral part of string theory.

Let us now focus on the open string with Neumann b.c.s and solve the equations of motion from the first term in (1.16):

$$\partial_a((-\gamma)^{1/2}\,\gamma^{ab}\,\partial_b X^\mu) = 0\,. \tag{1.20}$$

For general γ_{ab} this can be a non-trivial equation to solve. However, we are in luck. In two dimensions there are enough symmetries to make this equation trivial. The first symmetry we make use of is invariance under general coordinate transformations (1.14). One can show that, in two dimensions, it is *locally* possible to bring *any* metric to a *conformally flat* form through an appropriate coordinate transformation:

$$\sigma^a \to \sigma'^a \tag{1.21}$$

$$\gamma^{ab} \to \gamma'^{ab} = e^\phi\,\eta^{ab} = e^\phi \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.22}$$

where ϕ is some function of τ and σ. Now we are only a Weyl transformation (1.15) away from a flat metric. However, by inspecting of (1.20), we see that the conformal factor simply drops out. The solution for X in (1.20) is the following:

$$X^\mu(\tau, \sigma) = x^\mu + 2\,\alpha'\,p^\mu\,\tau + i\,(2\,\alpha')^{1/2} \sum_{n\neq 0} \frac{1}{n}\,\alpha_n^\mu\,e^{-i\,n\,\tau}\,\cos n\sigma\,, \tag{1.23}$$

where we require $\alpha_{-n}^\mu = (\alpha_n^\mu)^*$ to ensure reality. The parameter x^μ can be thought of as the string's initial center-of-mass position, p^μ as its center-of-mass momentum, and the α_n^μ as the oscillation modes of the string. Note that in the string action we did not fix the mass but rather the tension or energy per unit length of the string. Since length of the string depends on its oscillatory state, so will its mass. This makes sense relativistically, exciting the string's oscillatory modes means putting energy into it, and energy is the same as mass. In fact, by using Hamiltonian dynamics, one can show that the string's mass is given by the following relation:

$$M^2 = \frac{1}{\alpha'} \sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n \tag{1.24}$$

For the closed string one follows an analogous procedure to that for the open string case. One discovers, however, that the closed string has two sets of oscillators α^μ and $\tilde{\alpha}^\mu$, the so-called *right-* and *left-movers*, which can be viewed as non-stationary waves on the world-sheet traveling to the right and to the left respectively.

1.1.3 The bosonic string spectrum

We will now schematically study the quantum spectrum of the bosonic string. For a detailed account of what we are about to do, the reader is referred to any standard textbook on String theory such as [2] and [9].

Let us begin with the canonical quantization of the open string. Just as in the case of the point particle, the string is quantized by replacing Poisson bracket into commutators:

$$\begin{aligned} \{X^\mu(\tau,\sigma), \Pi^\nu(\tau,\sigma')\} &\to [X^\mu(\tau,\sigma), \Pi^\nu(\tau,\sigma')] = i\,\eta^{\mu\nu}\,\delta(\sigma-\sigma')\,, \\ \text{and} \qquad \{x^\mu, p^\nu\} &\to [x^\mu, p^\nu] = i\,\eta^{\mu\nu}\,, \end{aligned} \tag{1.25}$$

where $\Pi^\mu = (1/2\pi\,\alpha')\,\dot{X}^\mu$. Promoting the string coordinates to operators implies that the string oscillators α_n^μ are themselves promoted to operators. In fact, they acquire the following commutation relations:

$$[\alpha_m^\mu, \alpha_n^\nu] = i\,m\,\delta_{m+n}\eta^{\mu\nu}\,, \tag{1.26}$$

which we recognize as the commutation relations of the harmonic oscillator, where the α_{-n} and α_n are the creation and annihilation operators, respectively. So the string can be thought of as an eigenstate of the momentum operator p^μ with an infinite number of harmonic oscillators, each at a different excitation level. To create a state, define a "vacuum" state with some definite momentum $|\,p;0,0,...\,\rangle$, and then act on it with α_{-n}^μ operators. This will generate a string with definite momentum and oscillatory modes. The mass of the string will be given by a modified version of the classical formula (1.24). The quantum formula will count the number of harmonic oscillators and add a zero-point energy:

$$M^2 = \frac{1}{\alpha'} \left(\sum_{n=1}^{\infty} \alpha_{-n} \cdot \alpha_n - 1 \right). \tag{1.27}$$

Note that every oscillator α^μ carries a spacetime Lorentz index. It can be shown that the state created by acting with an oscillator on the vacuum, $\alpha_{-n}^\mu\,|\,p;0\,\rangle$, behaves as a vector boson. More generally, it can be shown that any string state will behave as a particle with mass and spin determined by the number of its oscillators and their indices. The closed string spectrum is also generated by harmonic oscillators. Its spectrum, however, is different from that of the open string. One key difference is that the closed string spectrum contains a massless spin-two particle which behaves like a *graviton*, whereas the open string does not.

Note also that the mass of a state is inversely proportional to α'. This means that in the low energy approximation to string theory (low α'), the massive states will become very massive and will be difficult to excite. That's why one can focus on the massless states when doing this approximation.

The goal of these first three subsections was to introduce the classical string and its quantum mechanical spectrum in a fair amount of detail. In the next two subsections, I will explain the Feynmann path integral quantization of the string and show that, in the low energy approximation (i.e. $\alpha' \to 0$), string theory can be effectively described by a spacetime field theory containing gravity, an antisymmetric tensor, and a scalar. This is a rather ambitious goal and a detailed treatment of this subject would require a lot of formalism and space, and would divert us from the main topic of this thesis: to study particular field theory configurations with gravity and scalar fields. I will, therefore, not show any detailed calculations; however, I will try to give an overview that is self-contained in that it does not require any new concepts beyond those of basic quantum field theory and path integrals. For an account that really does justice to the subject of the path integral quantization of the string, the reader is referred to [2, 9, 10].

1.1.4 The string path integral

Now that we know how the string spectrum comes about, let us turn to the path integral formalism to see how string amplitudes are defined.

When we want to compute quantum mechanical amplitudes for a point particle, the Feynmann path integral procedure instructs us to sum over all possible histories (world-lines) $x(t)$ that the particle can take, given some initial and final positions x_i and x_f, and to weight each with the phase $\exp(iS/\hbar)$, where $S = S[x(t)]$ is the action evaluated on the path. The partition function is then the following:

$$Z = \int d[x]\, e^{-iS[x]} \,. \tag{1.28}$$

This is sometimes referred to as first quantization in old fashioned language. It is nothing other than *quantum mechanics*. When we want to compute a quantum field theory amplitude using path integrals, we have to sum over all possible configurations a field ϕ can take given some spacetime boundary conditions, each weighted again by a phase. This yields the following partition function:

$$Z = \int d[\phi]\, e^{-iS[\phi]} \,, \tag{1.29}$$

where Z stands for *Zustandssumme* (sum of states). In the old fashioned language, this is second quantization. However, many physicists regard this as a misnomer because the procedure quantizes the field only once. This should just be called *quantum field theory*

In string theory we will be summing over all trajectories the string can take, i.e. over all possible world-sheets, and weight each with the Polyakov action (1.12):

$$Z = \int d[X]\, d[\gamma]\, e^{-iS[X,g]} \,, \tag{1.30}$$

where X represents the spacetime coordinates of the string and γ the world-sheet metric. This is the analog of (1.28). In other words, this path integral describes the quantum mechanics of the string.

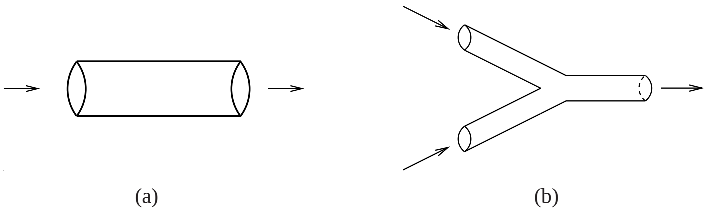


Figure 1.1: *Feynman diagram of a closed string: (a) propagator; (b) three-point function.*

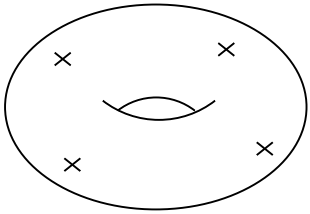


Figure 1.2: *Feynman diagram of the one-loop four-point diagram.*

We can take, however, a radically different point of view. If we view the Polyakov action as a two-dimensional action of fields, the path integral (1.30) becomes the analog of (1.29), summing over all configurations the fields $X^\mu(\sigma,\tau)$ and $\gamma_{ab}(\sigma,\tau)$ can take: This means that we have to sum over all scalar field configurations and all world-sheet geometries with given boundary conditions. For instance, the open and closed string propagators and string 3-point functions will contain diagrams[1] such as those in figures 1.1(a) and (b). By working in the Euclidean (Wick rotated) path integral formalism, and thus summing over Euclidean[2] two-dimensional metrics, one can use the conformal symmetry of the theory to map all world-sheets to compact *Riemann* surfaces. All external legs, which are infinitely long, are brought to a finite distance from each other. For instance, the closed string propagator diagram in figure 1.1(a), which was a cylinder, gets mapped to a sphere and the external legs get mapped to two points on the sphere. The "one-loop" four-point function diagram gets mapped to a torus with four points as external legs, see figure 1.2. The general rule is that all diagrams are mapped to compact closed or open surfaces and their external legs are mapped to points on the surfaces. However, it seems strange to map the external legs to points. After all, these external legs are supposed to represent initial and final states of strings, so mapping these to points seems to lose all the stringy information of these states. It turns out that the proper way to do this is to include what are called *vertex operators* on the compact surfaces. These operators $V(\sigma,\tau)$, which are inserted in the path integral, will supplement the latter with all the stringy information about initial and final states. For example, the state with no oscillators excited (the tachyon), but with some momentum p is

[1] It is also possible to draw diagrams representing the process where two open strings join at their endpoints, thereby forming a closed string. This implies that open string theory must include closed string modes.

[2] It is not always possible to perform the Wick rotation. When dealing with cosmological models, i.e. time-dependent spacetimes, Wick rotations can make the metric complex, see [11]

translated into the following vertex operator:

$$|\,0; p\,\rangle \qquad \Rightarrow \qquad \int d^2z \; : e^{i\,p\cdot X} : , \tag{1.31}$$

where z is a complex coordinate representing τ and σ, and : : represents normal ordering. Then, the two-point function for a tachyon with momentum p is computed as follows:

$$\langle\, 0; p\,|\, e^{iHT}\,|\, 0; p\,\rangle = \langle\, 0\,| \left(\int d^2z \; : e^{i\,p\cdot X} : \right)^{\dagger} \left(\int d^2z' \; : e^{i\,p\cdot X} : \right) |\, 0\,\rangle \tag{1.32}$$

$$= \int d[X]\, d[\gamma] \left(\int d^2z \; : e^{i\,p\cdot X} : \right)^{\dagger} \left(\int d^2z' \; : e^{i\,p\cdot X} : \right) e^{-i\,S[X,g]} . \tag{1.33}$$

The state that corresponds to the closed string graviton looks as follows:

$$\zeta_{\mu\nu}\, \alpha^{\mu}_{-1}\, \alpha^{\nu}_{-1}\; |\, 0; p\,\rangle \qquad \Rightarrow \qquad \int d^2z \; : \zeta_{\mu\nu}\, \partial_z X^{\mu}\, \partial_{\bar{z}} X^{\nu}\, e^{i\,p\cdot X} : , \tag{1.34}$$

where $\zeta_{\mu\nu}$ is a symmetric tensor. This is actually not all that strange and new. In ordinary QFT one must also use operator insertions in the path integral in order to "prepare" the initial and final states of an amplitude. For instance:

$$\langle\, x_{i_1} \cdots x_{i_n}\,|\, e^{iHT}\,|\, x_{f_1} \cdots x_{f_n}\,\rangle = \langle\, 0\,|\, \phi(x_{i_1}) \cdots \phi(x_{i_n})\, \phi(x_{f_1}) \cdots \phi(x_{f_n})\,|\, 0\,\rangle \tag{1.35}$$

$$= \frac{1}{Z} \int d[\phi]\, \phi(x_{i_1}) \cdots \phi(x_{i_n})\, \phi(x_{f_1}) \cdots \phi(x_{f_n})\, e^{-S} , \tag{1.36}$$

where $|\, 0\,\rangle$ is the Fock space vacuum.

It now seems like we have a rule for computing amplitudes, represent all external legs with operator insertions in the path integral, and sum over all two-dimensional compact surfaces. Summing over all surfaces means summing over all metrics and topologies of surfaces. The topology of a two dimensional compact surface is completely specified by the number of its boundaries, crosscaps, and handles (genus). But this procedure, being so similar to what we usually do in QFT, raises a very important question. The genus of a diagram is pictorially very reminiscent of the number of loops of a quantum field theory diagram. For instance, take the torus diagram with four vertex operators, shrink the string to a point particle and you will recover a one-loop diagram for a 4-point function in ϕ^4 theory. In a weakly coupled field theory, loop diagrams are usually suppressed by the coupling constant. So the big question is: where is the analog of this in string theory? Is there such a thing as a string coupling constant that keeps track of the loop order? Well, it turns out that when we wrote down the Polyakov action (1.12), we didn't write the most general action consistent with all the symmetries we found so far ((1.13), (1.14), (1.15)). There is one more piece we could have added, the two-dimensional gravity action [3]:

$$\chi = \frac{1}{4\pi} \int_{\mathcal{M}} d\tau\, d\sigma (\gamma)^{1/2}\, R + \frac{1}{2\pi} \int_{\partial M} ds\, K\, , \tag{1.37}$$

[3] Note that we are now working in the Euclidean formalism, so there is no minus sign under the square root in $(\gamma)^{1/2}$

where the first term is the Ricci scalar and the second term is the extrinsic curvature for a manifold with a boundary (an open string world-sheet). Although very geometric in nature, this action is a topological invariant for two-dimensional manifolds. It basically counts the genus and the number of boundaries and crosscaps of a surface.

$$\chi = 2 - 2g - b - c\,, \tag{1.38}$$

where g is the genus, b the number of boundaries, and c the number of crosscaps. Therefore, if we write the string action as follows:

$$S = S_P + \lambda\chi\,, \tag{1.39}$$

then diagrams will be weighted by a factor $e^{-\lambda\chi}$. If λ is small, we will say that string theory is *weakly coupled* and hence defined *perturbatively*. In this case, diagrams will be suppressed as their genus grows, just like QFT diagrams are suppressed as their loop number grows. If it is large, then we are in the *strongly coupled regime* of string theory, where most of the known techniques from field theory break down and very little is known. In the next section, we will see where this string coupling constant λ comes from; the answer will be quite surprising.

1.1.5 Strings in background fields

So far, we have been studying the theory of a string propagating in a D-dimensional flat spacetime. An obvious generalization at this point would be to start all over again with a Polyakov-like action that has a curved spacetime metric:

$$S_\sigma = \frac{1}{4\pi\alpha'}\int d\tau\, d\sigma\, (\gamma)^{1/2}\,\gamma^{ab}\,G_{\mu\nu}(X)\,\partial_a X^\mu\,\partial_b X^\nu\,. \tag{1.40}$$

This action is called a *non-linear sigma model*. From the two-dimensional perspective, this nontrivial spacetime metric $G_{\mu\nu}(X)$ plays the role of a *field-dependent coupling*, where the fields in question are the D scalar fields X^μ.

The attentive reader should be skeptical about this operation. Although it seems natural to replace the flat spacetime metric with a curved one, we should ask ourselves the following question: if the string is supposed to be the fundamental object which generates all particles and forces including gravity, are we allowed to simply put in by hand a curved metric in the action from which we will derive the string spectrum? In other words, if the graviton is a state of the string, how can we include gravity into the action that we must quantize in order to *find* the graviton? This seems like a vicious circle. However, there is a way out of it. The following explanation is borrowed from Polchinski's textbook [2].

Let us first consider a background spacetime metric that is nearly flat:

$$G_{\mu\nu}(X) = \eta_{\mu\nu} + h_{\mu\nu}(X)\,, \tag{1.41}$$

where $h_{\mu\nu}(X)$ is small. If we expand the integrand of the path integral we obtain the following:

$$e^{-S_\sigma} = e^{-S_P}\left(1 - \frac{1}{4\pi\alpha'}\int d\tau\, d\sigma\,\gamma^{1/2}\,\gamma^{ab}\,h_{\mu\nu}\,\partial_a X^\mu\,\partial_b X^\nu + \cdots\right), \tag{1.42}$$

where S_σ is the sigma model action (1.40), and S_P is the Polyakov action (1.12). The second term in the parenthesis is of the form of a vertex operator for a closed string graviton state (1.34), with $h_{\mu\nu} \propto \zeta_{\mu\nu}\, e^{i\,p\,X}$. So this perturbation of the background metric (1.41) can be viewed as the emission or absorption of a graviton state. Furthermore, if we have a full $G_{\mu\nu}$ background metric, we can view it as an exponentiation of a graviton vertex operator; i.e. a coherent state of gravitons. This validates our naive replacement of the Minkowski spacetime metric for a general curved metric in the non-linear sigma model (1.40).

Let us look back on what we have done so far. We wrote down an action for a string that propagates in a flat spacetime. By quantizing it we found that the string generates particles of different spin, including the graviton. Then, we included gravity into our starting action and discovered that this operation was merely an insertion of a coherent state of gravitons. A natural question at this point would be: can we include other fields in our action that can be viewed as coherent superpositions of other string states? The answer is yes.

Focusing on the massless closed string modes we can write the following action:

$$S_\sigma = \frac{1}{4\pi\,\alpha'} \int \gamma^{1/2} \left[(\gamma^{ab}\, G_{\mu\nu}(X) + i\,\epsilon^{ab}\, B_{\mu\nu}(X))\, \partial_a X^\mu\, \partial_b X^\nu + \alpha'\, R\, \Phi(X) \right], \tag{1.43}$$

where $B_{\mu\nu}$ is the background antisymmetric tensor, Φ is the background scalar (called dilaton), and R is the two-dimensional Ricci scalar. This is the most general action consistent with Poincaré invariance, two-dimensional g.c.t. invariance, and Weyl invariance, containing all massless closed string modes as background fields. In order for this theory to be consistent from the two-dimensional point of view, one needs to make sure that the classical Weyl symmetry is also a symmetry of the quantum theory. This is accomplished by requiring that the expectation value of the trace of the stress-energy tensor of the CFT vanish. This is just the requirement that a current that is classically conserved also be quantum mechanically conserved. This calculation, which we will not contemplate here, is called *anomaly cancelation*. Canceling the Weyl anomaly implies requiring that certain functions called *beta*-functions vanish. Up to first order in α', they look as follows:

$$\begin{aligned}
\beta^G_{\mu\nu} &= \alpha' \left(R_{\mu\nu} + 2\,\nabla_\mu\,\nabla_\nu \Phi - \frac{1}{4}\, H_{\mu\kappa\sigma}\, H_\nu{}^{\kappa\sigma} \right) + O(\alpha'^2)\,,\\
\beta^B_{\mu\nu} &= \alpha' \left(-\frac{1}{2}\, \nabla^\kappa H_{\kappa\mu\nu} + \nabla^\kappa \Phi\, H_{\kappa\mu\nu} \right) + O(\alpha'^2)\,,\\
\beta^\Phi &= \alpha' \left(\frac{D-26}{6\,\alpha'} - \frac{1}{2}\, \nabla^2 \Phi + \nabla_\kappa \Phi\, \nabla^\kappa \Phi - \frac{1}{24}\, H_{\kappa\mu\nu}\, H^{\kappa\mu\nu} \right) + O(\alpha'^2)\,,
\end{aligned} \tag{1.44}$$

where $H_{\mu\nu\kappa} \equiv \partial_\mu B_{\nu\kappa} + \partial_\nu B_{\kappa\mu} + \partial_\kappa B_{\mu\nu}$. These three β-functions must vanish independently. By taking proper linear combinations of these equations we are left with something very peculiar. We are left with equations that look like equations of motion for the spacetime background fields. For instance, the equation for the background spacetime metric $G_{\mu\nu}(X)$ turns out to be the Einstein equation. So the quantum string imposes constraints on its field-dependent couplings that look like spacetime field equations! Another peculiarity about these equations is that they require[4] that $D = 26$. The quantum mechanical string is thus only consistent in 26 spacetime

[4]They only require $D = 26$ if the background dilaton Φ is constant. Solutions such as the so-called *linear dilaton theory* with $D < 26$ do exist, however, they are not phenomenologically attractive.

dimensions! There is one more peculiar thing we should notice. The string coupling constant, which we called λ in (1.39) is actually the background value of the dilaton Φ, as can be seen from (1.43). So the string coupling is not a free parameter of the theory, it is determined by a background field of the string itself!

The constraints for the background fields not only look like equations of motion for spacetime fields, they can also be derived from a spacetime action:

$$S = \frac{1}{2\kappa_0^2} \int d^D X \, (-G)^{1/2} \, e^{-2\Phi} \Big[R + 4 \, \nabla_\mu \Phi \, \nabla^\mu \Phi - \frac{1}{12} H_{\mu\nu\lambda} H^{\mu\nu\lambda} - \frac{2\,(D-26)}{3\,\alpha'} + O(\alpha') \Big] , \tag{1.45}$$

where κ_0 is some physically meaningless constant. Whenever we are working in the low energy approximation of string perturbation theory, we can simply regard string theory as a spacetime field theory defined by the action above. We focus on the massless modes of the string because, for small α', the massive modes become very heavy and they decouple from the theory.

The search for solutions to (1.45) is also a search for a string theory background to quantize the string on. Such a background is often called a string theory *vacuum* because, after quantizing the string around it, it acquires the interpretation of a local minimum of some ‘potential’ for the string to oscillate about. The question however remains: what potential, and of what theory? The β-functions provide us with consistency conditions that dictate in what spacetime backgrounds strings are allowed to propagate. However, because string theory must ultimately be a theory of spacetime, and not just a two-dimensional CFT, one would like to be able to treat these backgrounds as vacuum states of some quantum theory. Such a theory does not yet exist. Because of that, there are a myriad of backgrounds to choose from and no principle that allows us to distinguish them. There is, in some sense, a vacuum *degeneracy*, because we do not have something like a potential that can help us distinguish the different ‘states’ of spacetime. One of the great challenges in string theory is finding what is called a *vacuum selection principle* that will actually pinpoint what background is *the* background for strings.

In recent years, however, the debate has shifted from the question: "what is the vacuum selection principle?", to the question: "should there be a vacuum selection principle?" L. Susskind has proposed a scenario, in which all possible allowed vacua actually exist [12]. This *landscape* scenario consists of stating that our universe is just one constituent of a *megaverse*, in which all kinds of universes (corresponding to all kinds of string theory backgrounds) exist, but are causally disconnected. In this approach, there is no room for a vacuum selection principle.

1.2 Superstrings and supergravities

1.2.1 Superstring theories

So far we have been studying the bosonic string, which is a fine toy model, but not a realistic description of particle physics for two reasons: first, the spectrum of the bosonic string contains a tachyon (i.e. a particle with negative mass), which indicates an instability of the string background. Second, it doesn’t contain any fermions since the oscillators only generate integer spin particles. To overcome these problems, we need to generalize the string to a supersymmetric string. By upgrading the two dimensional CFT to a supersymmetric conformal field theory, or *superconformal* field theory, and imposing consistency conditions on the quantized theory, the string will turn out to have spacetime supersymmetry, the tachyon will be projected out of the spectrum, and, as a bonus, the number of required spacetime dimensions will be reduced from 26 to 10. I will now give an intuitive overview of how the superstring is developed, the theories it leads to, and what its low energy approximations are (i.e. supergravities).

The basic form of the supersymmetric world-sheet action is as follows:

$$S = \frac{1}{4\pi} \int d\tau\, d\sigma \left(\frac{2}{\alpha'}\, \partial X^\mu\, \bar{\partial} X_\mu + \psi^\mu\, \bar{\partial}\psi_\mu + \tilde{\psi}^\mu\, \partial\tilde{\psi}_\mu \right), \tag{1.46}$$

where the ψ^μ are D two-dimensional fermions. This theory is a superconformal field theory. By analyzing its spectrum in analogy with the bosonic string, one will find that the world-sheet fermions also have oscillators, which act as raising and lowering operators on the vacuum. This will give rise to not only spacetime bosonic states, but also spacetime fermionic states. In fact, by properly counting the bosonic and fermionic states that are generated, one finds that this theory has spacetime supersymmetry. This means that the number of bosonic degrees of freedom matches the number of fermionic degrees of freedom. This theory turns to be anomaly-free only in ten spacetime dimensions.

A more detailed study of the superstring will show that it is actually possible to define *five* different consistent supersymmetric string theories:

- **Type I**: This is a theory of unoriented open strings.
- **Type II**: There are two theories in this category, **Type IIA** and **Type IIB**. These are theories of closed strings, and they differ in the boundary conditions applied on the world-sheet fermions.
- **Heterotic**: There are two heterotic string theories. These theories are constructed in very peculiar ways, and they naturally have non-Abelian spacetime gauge symmetries. Their groups are indicated by their names: **Het** $E_8 \times E_8$, and **Het** SO(32).

Type I and the Type II theories are *a priori* not free of tachyons. However, a certain projection must be performed on the spectrum for consistency conditions, after which all tachyons are gone. This projection is called the *GSO* projection after Gliozzi, Scherk and Olive [13]. All five of these string theories live in ten spacetime dimensions.

1.2.2 Supergravities

We will now write down the low energy effective actions for these five supersymmetric string theories. But before we do so, let us look back to the case of the bosonic string for some moral guidance. When we quantized the string, we required that the classical symmetries (spacetime Poincaré, and 2-D Weyl invariance) be respected at the quantum level. This lead to the vanishing of the β-functions of the field-dependent couplings $G_{\mu\nu}(X)$, $B_{\mu\nu}(X)$, $\Phi(X)$, which carved out for us a procedure to write down a unique spacetime field theory that describes the massless modes of the string at low energy.

The five supersymmetric string theories not only have the symmetries of the bosonic string, but they each come with a different form of spacetime supersymmetry. It turns out that supersymmetry is a stringent enough constraint that, given the dimensionality and field content of a theory, there is only one possible spacetime action one can write down. This means that all we need to know to construct the low energy effective actions for the massless modes for these five string theories is their spectrum and the kind of supersymmetry they have. The resulting actions are called *supergravities*. Their symmetries naturally combine general coordinate transformation invariance and local supersymmetry as the name suggests. The bosonic parts of the actions for the five supergravities are the following:

- **Type IIA**

$$\begin{aligned} S_{\mathrm{IIA}} = \frac{1}{2\kappa_0^2} \int d^{10} (-G)^{1/2} \Big(e^{-2\Phi} \Big[R + 4 (\nabla \Phi)^2 - \frac{1}{12} (H^{(3)})^2 \Big] \\ - \frac{1}{4} (G^{(2)})^2 - \frac{1}{48} (G^{(4)})^2 \Big) - \frac{1}{4\kappa_0} \int B^{(2)} \, dC^{(3)} \, dC^{(3)} , \end{aligned} \tag{1.47}$$

where G is the 10-dimensional metric, Φ the dilaton, $H^{(3)} = dB^{(2)}$ the field strength of a two-form, $G^{(2)} = dC^{(1)}$ the field strength of a one-form, and $G^{(4)} = dC^{(3)} + H^{(3)} \wedge C^{(1)}$ can be seen as the modified field strength of a three-form.

- **Type IIB**

$$\begin{aligned} S_{\mathrm{IIB}} = \frac{1}{2\kappa_0^2} \int d^{10} (-G)^{1/2} \Big(e^{-2\Phi} \Big[R + 4 (\nabla \Phi)^2 - \frac{1}{12} (H^{(3)})^2 \Big] \\ - \frac{1}{12} (G^{(3)} + C^{(0)} H^{(3)})^2 - \frac{1}{2} (dC^{(0)})^2 - \frac{1}{480} (G^{(5)})^2 \Big) \\ + \frac{1}{4\kappa_0^2} \int \Big(C^{(4)} + \frac{1}{2} B^{(2)} C^{(2)} \Big) G^{(3)} H^{(3)} , \end{aligned} \tag{1.48}$$

where $G^{(3)} = dC^{(2)}$, $G^{(5)} = dC^{(4)} + H^{(3)} \wedge C^{(2)}$, and $C^{(0)}$ is a scalar. To get the right number of degrees of freedom, one must impose that the field strength of the four-form $F^{(5)} = dC^{(4)}$ be self-dual: $F^{(5)} = *F^{(5)}$. However, this constraint can only be imposed at the level of the equations of motion.

- **Type I**

$$\begin{aligned} S_{\mathrm{I}} = \frac{1}{2\kappa_0^2} \int d^{10} (-G)^{1/2} \Big(e^{-2\Phi} \Big[R + 4 (\nabla \Phi)^2 \Big] \\ - \frac{1}{12} \Big(\tilde{G}^{(3)} \Big)^2 - \frac{\alpha'}{8} e^{-\Phi} \, Tr (F^{(2)})^2 \Big), \end{aligned} \tag{1.49}$$

where $\tilde{G}^{(3)} = dC^{(2)} - \frac{\alpha'}{4} \left[\frac{1}{30} Tr \left(A \wedge dA + \frac{1}{3} A \wedge A \wedge A \right) \right]$. The trace '$Tr$' runs over Yang-Mills group indices.

- **Heterotic**

$$S_{\text{Het}} = \frac{1}{2\kappa_0^2} \int d^{10} (-G)^{1/2} e^{-2\Phi} \Bigg(R + 4 (\nabla \Phi)^2 - \frac{1}{12} \left(\tilde{H}^{(3)}\right)^2 - \frac{\alpha'}{8} e^{-\Phi} Tr(F^{(2)})^2 \Bigg), \tag{1.50}$$

where $\tilde{H}^{(3)} = dB^{(2)} - \frac{\alpha'}{4} \left[\frac{1}{30} Tr \left(A \wedge dA + \frac{1}{3} A \wedge A \wedge A \right) \right]$.

This concludes the introduction to string theory. The main goal of this chapter was to explain how a quantum theory of relativistic strings can, in a certain approximation, lead to a spacetime gravitational field theory. Actually, two approximations were made. The first one is the assumption that strings interact weakly, i.e. that the string coupling constant given by the constant part of the dilaton is small. This allows us to define a CFT on the world-sheet perturbatively. The second assumption is the low energy approximation. At low energies only the massless states of the string are excited. In the β-functions this is manifested by a truncation of α' corrections. This is what allows us to write down a spacetime classical field theory, such as a supergravity, as an effective description of string theory.

Throughout this thesis we will be working with these approximations. In the next part, which consists of two chapters, we will study instantons and their role in string theory. In the second part, chapters 4, 5 and 6, we will study cosmology in the context of scalar-gravity theories. These theories are often supergravity Lagrangians that have been dimensionally reduced and truncated to contain only the metric and scalar fields. In the final part of this thesis, chapter 7, we will see how the first two parts come together in two different ways: first, we will see how Wick rotations can relate supergravity instantons to cosmological solutions. Then, we will make a paradigm shift and treat those two kinds of solutions on equal footing, by regarding them as trajectories of a particle in a fictitious *target space* parametrized by the scalar fields of the theory.

Chapter 2

Instantons

In this chapter we will study the basics of instantons, heavily borrowing material from the classic textbooks by S. Coleman [14] and R. Rajaraman [15]. First, we will see their application to quantum mechanics, which is conceptually and technically the simplest framework to introduce the topic. Then, we will move on to quantum field theory, where the example of the Yang-Mills instanton will give us all the tools to understand these objects in generality. Finally, solitons will be briefly introduced, and we will see how sometimes an instanton in D Euclidean dimensions can correspond to a soliton in $D+1$ Lorentzian dimensions.

2.1 Introduction

2.1.1 An alternative to WKB

In quantum mechanics it is possible for a particle to penetrate a region of potential energy that is higher than the particle's own energy. This classically forbidden motion is known as *quantum tunneling* and, for a general potential barrier, one can compute the tunneling amplitude of a particle by means of the WKB approximation. The latter is a so-called *semiclassical* approximation, which means that it requires small $\hbar$. Let us see what happens in the case of a particle of unit mass in $1+1$ dimensions, subject to some potential $V(x)$.

The Schrödinger equation reads:

$$\frac{d^2\psi}{d\,x^2} = \frac{2\,(V(x)-E)}{\hbar^2}\,\psi\,. \tag{2.1}$$

If $V(x) = \textit{constant}$, then the solution would be a plain wave:

$$\psi \propto e^{-i\,k\,x}, \quad \text{where} \quad k \equiv \frac{\sqrt{2\,(E-V)}}{\hbar}\,. \tag{2.2}$$

In the case of quantum tunneling $V > E$, so the momentum becomes imaginary, and instead of a plain wave, we obtain an exponentially decreasing function:

$$\psi \propto e^{-\kappa\,x}, \quad \text{where} \quad \kappa \equiv \frac{\sqrt{2\,(V-E)}}{\hbar}\,. \tag{2.3}$$

Let us now take a non-constant potential but make the approximation that $V(x)$ varies slowly compared to the rate of decay κ of the wave function. Then, we can rewrite the Schrödinger equation as follows:

$$\frac{d\psi}{dx} = \pm \frac{\sqrt{2\,(V(x)-E)}}{\hbar}\,\psi\,, \tag{2.4}$$

Differentiating this equation yields the original Schrödinger (2.1) upon dropping a term proportional to $V'/\hbar^2\,\kappa$. The solution for a particle tunneling to the right is then:

$$\psi \propto \exp\left(-\frac{1}{\hbar}\int \sqrt{2\,(V(x)-E)}\,dx\right). \tag{2.5}$$

The amplitude for the particle to tunnel is then:

$$\exp\left(-\frac{1}{\hbar}\int_a^b \sqrt{2\,(V(x)-E)}\,dx\right), \tag{2.6}$$

where a and b are the beginning and endpoint of the tunneling trajectory.

The approximation we made is a semiclassical one in the sense that it requires that $\hbar$ be 'small'. To see this, recall that differentiating the equation we actually solved (2.4) yielded the true Schrödinger (2.1) equation if we dropped a V' term. Comparing this term to the term that we did keep shows that the dimensionless quantity we are neglecting is $\hbar\,V'/(2\,(V-E))^{3/2}$, which is small in the semiclassical limit $\hbar \to 0$.

Now that we have obtained this result by using the WKB approximation, we will rederive it through a completely different method, which will be the subject of this chapter: the method of instantons.

Let us begin by rewriting (2.5) in a different way. First, we set the energy of the particle to zero (which can always be done via a suitable shift in the potential), $E = 0$. Then, we have:

$$\int_a^b \sqrt{2\,(V(x)-E)}\,dx = \int_a^b i\,p\,dx = \int_a^b i\,\frac{dx}{dt}\,dx\,, \tag{2.7}$$

where p is the momentum of the particle, and in the second equation we used the fact that the mass has been set to 1. If we perform a Wick rotation $t \to i\,\tau$ we can write this as follows:

$$\int_{\tau_a}^{\tau_b} p\,\dot{x}\,d\tau = \int_{\tau_a}^{\tau_b} \mathcal{L}_E\,d\tau = \mathcal{S}_E\,, \tag{2.8}$$

where $\mathcal{S}_E$ is the action of the classical particle in Euclidean spacetime with zero energy. This teaches us a new way to compute tunneling amplitudes. Simply compute the Euclidean action of the tunneling trajectory. To see where this comes from, let us turn to the language of path integrals.

Let us compute the tunneling amplitude for the same $(1+1)$-dimensional problem using path integrals. The amplitude is given by the following:

$$K(a,b;T) \equiv \langle x=a|e^{iHT/\hbar}|x=b\rangle = \int d[x(t)]\,e^{iS[x(t)]/\hbar} \tag{2.9}$$

$$\text{with} \qquad S \equiv \int_{t_a}^{t_b} \left(\tfrac{1}{2}\,(dx/dt)^2 - V(x)\right)\,dt\,, \tag{2.10}$$

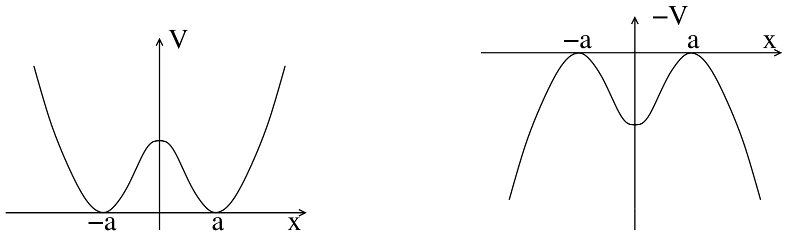


Figure 2.1: *Figure (a) depicts a double-well potential, while figure (b) depicts the inverted potential.*

where the path integral sums over all paths from $x = a$ to $x = b$, and $T \equiv t_b - t_a$. If we now analytically continue this to Euclidean spacetime (i.e. $t \to i\tau$), this becomes:

$$K_E(a, b; T) \equiv \langle x = a|e^{-HT/\hbar}|x = b\rangle = \int d[x(\tau)]\, e^{-S_E[x(\tau)]/\hbar} \tag{2.11}$$

$$\text{with} \qquad S_E \equiv \int_{\tau_a}^{\tau_b} \left(\tfrac{1}{2}\,(dx/d\tau)^2 + V(x)\right) d\tau\,. \tag{2.12}$$

There are basically two motivations to perform this Wick rotation: firstly, the Minkowskian path integral is rigorously speaking not well-defined. It is difficult to prove that the phases of trajectories that greatly differ from the classical path actually cancel out, in order to make the path integral convergent. However, since the partition function is an analytic function of time, one can properly define the path integral by Wick rotating into Euclidean signature, which yields a well-defined convergent object, and then Wick rotating physical results back to Minkowskian signature.

The second motivation is the fact that the partition function $\langle e^{-HT/\hbar}\rangle$, in the limit $T \to \infty$, projects the lowest energy eigenstates. This provides information about vacuum energy and the ground state wave function, as we will see later on. From this point of view, there is no need to think in terms of Euclidean time. The path integral for the partition function can be derived from first principles without use of the Wick rotation.

If we now take the limit $\hbar \to 0$, we see that the largest contribution to this path integral will come from a trajectory that minimizes the Euclidean action. If S_0 is the value of the action for such a trajectory, then, to leading order in $\hbar$, the Euclidean amplitude will go like $K_E \propto e^{-S_0/\hbar}$. The problem of extremizing the Euclidean action S_E is equivalent to that of extremizing the Minkowskian action of a particle subject to an inverted potential $-V(x)$. More explicitly, the variational equation of the Euclidean action (2.12),

$$\frac{d^2x}{d\tau^2} - \frac{dV}{dx} = 0\,, \tag{2.13}$$

looks just like the *classical* equation of motion of a particle in a potential $-V(x)$, as shown in figure 2.1(b). Solving this equation, we find that

$$\frac{dx}{d\tau} = \sqrt{2\,V}\,, \tag{2.14}$$

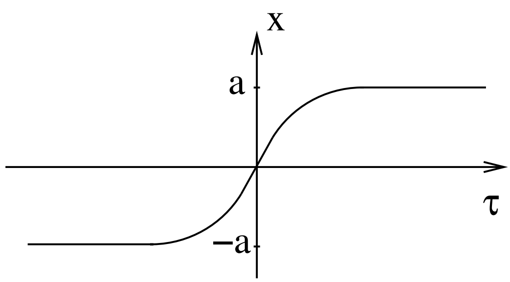


Figure 2.2: *The kink solution: a classically forbidden trajectory that interpolates between the two classical vacua of the double-well potential.*

and using this we can rewrite the action (2.12) as

$$S_0 = \int_{\tau_a}^{\tau_b} 2\,V\,d\tau = \int_{\tau_a}^{\tau_b} 2\,V\frac{d\tau}{dx}\,dx = \int_{\tau_a}^{\tau_b} \sqrt{2\,V}\,dx\,, \tag{2.15}$$

which matches our WKB calculation (for $E = 0$) (2.6). So, in order to compute a tunneling amplitude, instead of thinking of a classically forbidden trajectory where the particle goes through a potential barrier such as the one depicted in figure 2.1(a), we simply compute the action for a classically allowed trajectory where the particle rolls down from the top of the left-hand side hill and then up to the top of the right-hand side hill of the inverted potential in figure 2.1(b).

This classical trajectory $x_{cl}(\tau)$ will qualitatively have the shape depicted in fig 2.2. It is usually referred to as the *kink*. The precise shape of this trajectory is not important. What matters is that this function interpolates between the two constant functions, $x = -a$ and $x = a$, which are the two classical vacua of the double-well problem. It differs significantly from those two constant values only within a localized region in the range of τ, so the Lagrangian density is itself non-zero only in a finite region. It is because of this that the trajectory has finite action, giving rise to a non-zero contribution to the path integral.

2.1.2 A tool of the trade: The semiclassical approximation

Although the minimum of the Euclidean action gives the largest contribution to the path integral, it only constitutes a "point" of measure zero in the space of all trajectories we integrate over. This is emphatically stated and clearly explained in Coleman's work [14]. It is, therefore, a bit too brutal and incorrect to define the semiclassical approximation as a sum of contributions of the minimum (or minima) of the action. The semiclassical approximation consists in computing the path integral by approximating the *regions* around the local minima of the action with Gaussians. Although it is treated extensively in many standard QFT books such as [16], we will briefly go over it here. Let us start with by computing the following one-dimensional integral as a toy example:

$$I = \int_{-\infty}^{+\infty} \exp(-f(x)/\hbar)\,dx\,, \tag{2.16}$$

where we assume that $f(x)$ is bounded from below and has exactly one minimum at $x = x_0$. By expanding the function in its Taylor series around x_0, we can re-write the integral as follows:

$$I = \int_{-\infty}^{+\infty} dx \, \exp\left(-\frac{1}{\hbar}\,(f(x_0) + \tfrac{1}{2}\,(x - x_0)^2\, f''(x_0) + O\left((x - x_0)^3\right))\right), \tag{2.17}$$

$$= \exp(-\frac{1}{\hbar}\, f(x_0)) \int_{-\infty}^{+\infty} d\overline{x}\, \exp\left(\frac{1}{2\,\hbar}\,\overline{x}^2\, f''(x_0)\right) h(\overline{x})\,, \tag{2.18}$$

where $\overline{x} = x - x_0$ and $h(\overline{x})$ contains the higher order terms. If we take the limit $\hbar \to 0$, it can be easily shown that the Gaussian in the integrand becomes a δ-function of strength $\sqrt{2\,\pi\,\hbar/f''(0)}$. Since $h(x_0) = 1$, we have the following result[1] for small $\hbar$:

$$I \approx \exp\left(-\frac{1}{\hbar}\, f(x_0)\right) \sqrt{\frac{\pi\,\hbar}{f''(x_0)}}\, (1 + O(\hbar))\;. \tag{2.19}$$

Therefore, the semiclassical approximation does not only sum points of measure zero, it actually sums over the regions around minima. These regions have non-zero measure. This is reflected by the fact that the result (2.19) contains not only the value of the action minimum $f(x_0)$, but also the curvature around it $f''(x_0)$. In the case where $f(x)$ has many local minima one must approximate the calculation by summing over several Gaussian integrals, each centered at a local minimum.

In quantum mechanics, one performs an integral over the infinite dimensional space of paths $x(\tau)$, and the function f is replaced by the functional $S\,[x(\tau)]$, the action. If we discretize time, (i.e. $\tau = ..., \tau_{-i}, \tau_{-i+1}, ..., \tau_0, ...\tau_{i-1}, \tau_i, ...$), then the variables of the integral become the $x_i \equiv x(\tau_i)$. Let us rewrite our action as follows:

$$S\,[x(\tau)] = \int d\tau\,(-x\,\partial_\tau^2\, x + V(x))\,, \tag{2.20}$$

where we partially integrate the kinetic term. Notice that in a discrete time a derivative is simply a difference, i.e. $x'(\tau) \to x_{i+1} - x_i$; therefore, the kinetic term of the action can be represented by a matrix, $-x\,\partial^2\, x \to \sum_{i,j}\, x_i\, D_{ij}\, x_j$ for some symmetric D_{ij}. Hence, we can write the action as

$$S\,[x(\tau)] \to S\,(x_{0i}) = \sum_j \left(-\sum_k\, x_j\, D_{jk}\, x_k + V(x_j)\right) \tag{2.21}$$

for some proper choice of D_{jk}. Now, let us perform the semiclassical approximation by expanding the action around its minimum, x_{0i} (the classical path), and keeping only the quadratic terms:

$$S\,[x_i] = S\,[x_{0i}] + \sum_{jk}\, \overline{x}_j\, \frac{\partial^2 S\,[x_0]}{\partial x_j \partial x_k}\, \overline{x}_k\,, \tag{2.22}$$

$$= S_0 + \sum_{jk}\, \overline{x}_j \left(-D_{jk} + \frac{\partial^2 V(x_{0j})}{\partial x_j \partial x_k}\, \delta_{jk}\right) \overline{x}_k = S_0 + \sum_{jk}\, \overline{x}_j\, \left(A_{jk}\right)\, \overline{x}_k\,, \tag{2.23}$$

[1] Note that this requires $f'' \neq 0$.

where $S_0 \equiv S[x_{0i}]$, $\overline{x_i} \equiv x_i - x_{0i}$, and A_{jk} is some matrix. This form of the action now looks like the exponent of a multi-variable Gaussian. The result for a M-variable Gaussian integral with a generic matrix $\mathbf{A}$ is the following[2]:

$$\int_{\infty}^{+\infty} d\mathbf{x} \, \exp\left(-\frac{1}{2\hbar}\mathbf{x}^T \, \mathbf{A} \, \mathbf{x}\right) = \sqrt{\frac{(2\pi\hbar)^M}{\det \mathbf{A}}} \,, \tag{2.24}$$

where the determinant can be computed as a product of eigenvalues. In the continuum limit, the path integral defines determinants for operators. In the case at hand, it defines the following:

$$\int d[x(\tau)] \, \exp\left[-\frac{1}{2\hbar} \int d\tau \, x \left(-\partial_\tau^2 + V''(x_0(\tau))\right) x\right] = \frac{N}{\sqrt{\det(-\partial^2 + V''(x_0(\tau)))}} \,, \tag{2.25}$$

where N is a normalization constant, and the determinant can be computed by analogy with matrices, i.e. by finding the eigenfunctions of the operator $(-\partial^2 + V''(x_0))$ and then taking the product of their eigenvalues.

This is a natural point to give a definition of an instanton.
Definition: *An instanton is a solution to the Euclidean equations of motion with finite, non-zero action.* This definition ensures that the instanton is a saddle point that will contribute to a path integral.

Let us now get back to our double-well problem. We set out to compute the tunneling amplitude $< -a \,|\, e^{-HT/\hbar} \,|\, a >$ with the path integral given in (2.11). To apply the semiclassical approximation, we need to find the configurations with minimal Euclidean action. The kink in figure 2.2 is the absolute minimum of the action, so we should compute the path integral by means of a Gaussian integral centered around the kink. However, the kink is not the only minimum, it is only the absolute one. The action (2.12) has several local minima which have to be summed over too. One can take a sequence of kinks and *anti*-kinks as shown in figure 2.3. Any alternating sequence will do as long as it satisfies the boundary conditions of the path integral.[3]

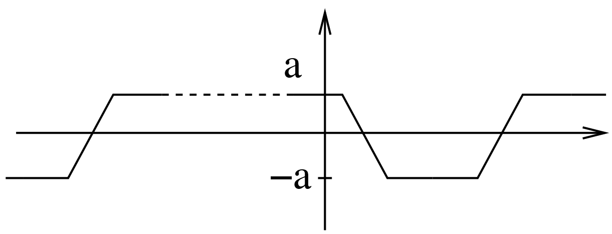


Figure 2.3: *An alternating sequence of kinks and anti-kinks. This interpolating trajectory is a local minimum of the Euclidean action.*

Another subtlety is that if the range T of τ is infinite, each kink or anti-kink can be displaced along the time axis by an arbitrary amount and yield a new trajectory whose action is equal to the

[2]In analogy with the one-dimensional case, this requires $\det(\mathbf{A}) \neq 0$

[3]Sequences of kinks and anti-kinks are only true stationary points in the limit where the range of Euclidean time $T \to \infty$, which is the limit we will always be interested in.

previous one. For instance, the one-kink trajectory can be centered around any value τ' and the value of its action will be independent of τ'. This means that we have to sum over the positions of the (anti-)kinks in each sector of the path integral. This is reflected in (2.25) by the fact that the operator $-\partial^2 + V''(x_0)$ will have some zero eigenvalues, or *zero modes*. This would *a priori* yield an infinite result for the amplitude calculation. Fortunately, there is a trick to "factor out" the infinity and cancel it against the N in (2.25). This is the Fadeev-Popov trick, which I will not derive here. For a pedagogical derivation of it, the reader is referred to [17].

The contribution to the amplitude from a single kink is the following:

$$\langle -a|e^{-HT/\hbar}|a\rangle_{(1)} = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2}\, K\, e^{-S_0/\hbar}\, T\,, \tag{2.26}$$

where $\omega \equiv V''(-a) = V''(a)$, and K is a constant which takes into account the calculation of the translational zero mode. Note that this is proportional to $e^{-S_0/\hbar}$, as expected. This is the biggest contribution to the tunneling process. Now we need to sum over all configurations with kink-anti-kink sequences. If we use the $T \to \infty$ approximation then, in most of the configurations, the kinks and antikinks will be far away from each other, in which case the action becomes additive, i.e. $S_{\text{kink+antikink}} = S_{\text{kink}} + S_{\text{antikink}} = 2\,S_{\text{kink}}$ [4]. Each (anti)kink also brings a power of K with it. In a tunneling trajectory from $-a$ to $+a$ there must always be one kink more than there are antikinks. Our task is then clear, the calculation and result are the following:

$$K(-a, a; T) = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \sum_{\text{odd } n} \frac{(K\, e^{-S_0/\hbar}\, T)^n}{n!} \tag{2.27}$$

$$= \frac{1}{2}\left(\frac{\omega}{\pi\hbar}\right)^{1/2} \left[\exp\left(-\tfrac{1}{2}\,\omega T + K\, e^{-S_0/\hbar}\, T\right) + \exp\left(-\tfrac{1}{2}\,\omega T - K\, e^{-S_0/\hbar}\, T\right)\right]. \tag{2.28}$$

2.1.3 True vacua

Consider again the particle in 1 + 1 dimensions subject to a double-well potential as depicted in fig 2.1(a). What is the vacuum structure of this problem?

If we neglected tunneling effects, our classical intuition would tell us that the ground state of the particle will be localized at one of the two wells. To find such a state, we would pick one of the wells (say, the one at $x = -a$), and approximate it with a parabolic or harmonic oscillator potential around its center,

$$V(x-a) = V(-a) + \frac{1}{2}\,\omega^2\, x^2 + O(x^3), \quad \text{where} \quad \omega^2 \equiv V''(-a)\,. \tag{2.29}$$

Then, we would solve the harmonic oscillator as usual, and do the same for the other well. This would lead us to conclude that the ground state is degenerate, namely, that there are two ground states, each localized at one well:

$$\psi_{-a}(x) = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{\omega}{\pi\hbar}\,(x+a)^2\right), \qquad E_{-a} = \tfrac{1}{2}\,\hbar\,\omega\,,$$

$$\psi_{a}(x) = \left(\frac{\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{\omega}{\pi\hbar}\,(x-a)^2\right), \qquad E_{a} = \tfrac{1}{2}\,\hbar\,\omega\,. \tag{2.30}$$

[4] an antikink has the same action as a kink

However, we know that a particle can tunnel from one well to the other, so these states we have constructed are not really stationary. This means that they are not energy eigenstates, and therefore, not vacuum states. A true vacuum state will have to be some linear combination of the two states we constructed in the naïve perturbative approach (2.30). As we will see next, instantons will give us all the information we need about this system.

Let us take a closer look at what the tunneling amplitudes we computed in the previous subsection tell us. Let $|E_n>$ be the set of true energy eigenstates of this system, then,

$$K(-a, a; T) \equiv \langle -a|e^{-HT/\hbar}|a\rangle \tag{2.31}$$

$$= \sum_n \langle -a|E_n\rangle \langle E_n|a\rangle \, e^{-E_n T/\hbar}, \tag{2.32}$$

which in the large T limit yields:

$$K(-a, a; T) = \sum_{\substack{\text{Lowest} \\ \text{energy states}}} \langle -a|E_n\rangle \langle E_n|a\rangle \, e^{-E_n T/\hbar}. \tag{2.33}$$

This provides us very valuable information. Comparing this to (2.28) we realize that the energies of the two lowest energy eigenstates are

$$E_\pm = \tfrac{1}{2}\hbar\omega \pm \hbar K e^{-S_0/\hbar}, \tag{2.34}$$

where E_- is the true ground state energy and E_+ is the energy of the second lowest level. Equation (2.33) also tells us what the wave functions of these states look like:

$$\langle -a|E_\pm\rangle \langle E_\pm|a\rangle = \langle -a|E_\pm\rangle \langle E_\pm|a\rangle = \mp\tfrac{1}{2}\left(\frac{\omega}{\pi\hbar}\right)^{1/2}. \tag{2.35}$$

The ground state wave function is spatially even and can be shown to coincide with an even linear combination of the two wave functions in (2.30) to leading order in the approximation of the potential. The next energy level is spatially odd.

The lesson instantons teach us is that when the vacuum of a system is *classically* degenerate, tunneling effects lift the degeneracy, and the quantum mechanical vacuum state will be a linear combination of the naive wave functions that respects the symmetry of the potential. In the case of the double-well problem, the vacuum state turned out to be even, just like the potential.

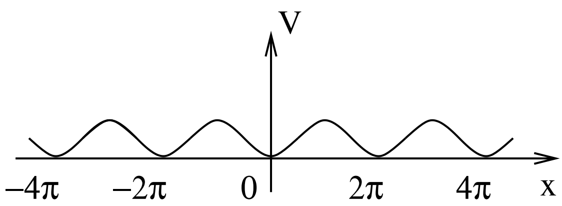


Figure 2.4: *The periodic potential.*

Let us now see what happens when the symmetry of the potential is larger than just $\mathbb{Z}_2$. Consider the periodic potential whose shape is depicted in figure 2.4. Again let us ask the question:

what is the vacuum structure of this system? Let us go through it as we did in the previous problem, starting from the naive approach. Naively, neglecting quantum tunneling effects, we would assume that the particle's wave function is centered around one of the infinitely many minima of the potential, say $x = 0$, thereby spontaneously breaking the $\mathbb{Z}$-symmetry of the system. At this point we would approximate the potential around $x = 0$ with a harmonic oscillator, and find the ground state wave function and energy. But in light of the above discussion, we are aware of tunneling effects. By computing the tunneling amplitude for the particle to go from one minimum $x = 2\pi N_1$ to another $x = 2\pi N_2$, and taking the limit $T \to \infty$, we will obtain information about the true vacuum states:

$$K(2\pi N_1, 2\pi N_2; T) = \sum_n \langle 2\pi N_1 | E_n\rangle \langle E_n | 2\pi N_2\rangle\, e^{-E_n T/\hbar}\,, \tag{2.36}$$

$$\longrightarrow \sum_{\substack{\text{Lowest}\\ \text{energy states}}} \langle 2\pi N_1 | E_n\rangle \langle E_n | 2\pi N_2\rangle\, e^{-E_n T/\hbar}\,, \tag{2.37}$$

namely, the lowest energy eigenvalues and their wave functions. To compute this amplitude, we again need to sum over the one-kink sector, and over all sequences with multiple kinks and antikinks. The one-kink contribution to the amplitude is the same as in that the in the double-well potential, namely equation (2.26), and the action is still additive, so the rules of the game are the same. The only difference is that, now, kinks do not have to be followed by antikinks and vice-versa, because the space where the particle moves has been enlarged to infinity. In other words, the instanton trajectories need not be confined to the interval $[2\pi N_1, 2\pi N_2]$, they just need to begin and end at $2\pi N_1$ and $2\pi N_2$ respectively. The sum is the following:

$$K(2\pi N_1, 2\pi N_2; T) = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{-\omega T/2} \sum_{n,\bar{n}} \frac{(K\, e^{-S_0/\hbar}\, T)^{n+\bar{n}}}{n!\,\bar{n}!}\, \delta_{N_2-N_1-n-\bar{n}}\,, \tag{2.38}$$

where the Kroenecker δ-function imposes the boundary conditions. This δ-function can be rewritten as follows:

$$\delta_{N_2-N_1-n+\bar{n}} = \int_0^{2\pi} \frac{d\theta}{2\pi}\, e^{i\theta(N_2-N_1-n+\bar{n})}\,. \tag{2.39}$$

By inserting this integral, the sums over n and $\bar{n}$ decouple. The result, which is also derived in Coleman's lectures [14] and in Rajaraman's book [15] is the following:

$$K(2\pi N_1, 2\pi N_2; T) = \int_0^{2\pi} \frac{d\theta}{2\pi}\, e^{i\theta(N_2-N_1)} \left(\frac{\omega}{\pi\hbar}\right)^{1/2} \exp\!\left(-\tfrac{1}{2}\,\omega T + 2K\, e^{-S_0/\hbar}\, T\, \cos(\theta)\right). \tag{2.40}$$

Notice that the discrete sum over low energy states has become an integral over a continuum of energy states labeled by θ. The energy of a *theta*-state is then given by the following:

$$E_\theta = \hbar\left(\tfrac{1}{2}\,\omega T - 2K\, e^{-S_0/\hbar}\, T\, \cos(\theta)\right), \quad \text{where} \quad 0 \le \theta \le 2\pi\,, \tag{2.41}$$

where the state of lowest energy is the one with $\theta = 0$. These energy levels are reminiscent of the band structures exhibited by systems with periodic potentials. This is the limit where the number of minima of the potential goes to infinity (in other words, this is the limit where the periodic potential goes on forever). In this limit, the band of energy levels becomes continuous, yielding the energy formula (2.41). The double-well problem could be regarded roughly as the opposite limit, where the number of potential minima is two. In that case θ could only have two discrete values, 0 and π. We also have the following information about the wave function of the θ-state:

$$\langle 2\pi N_1|\theta\rangle\ \langle\theta|2\pi N_2\rangle = \left(\frac{\omega}{\pi\hbar}\right)^{1/2} e^{i\theta(N_2-N_1)}\,. \tag{2.42}$$

The wave function of a θ-state is quasi-periodic: Under a translation by 2π it gains a phase $e^{i\theta}$. So these states restore the symmetry of the system. In fact it can be shown that, to leading order in the approximation of the potential, the wave function of a θ-state is the following:

$$|\theta\rangle = \sum_N e^{i\theta N}\ |\psi_{2\pi N}\rangle\ , \tag{2.43}$$

where the $|\psi_{2\pi N}\rangle$ are the naively constructed harmonic oscillator ground states of each potential minimum, when tunneling effects are neglected. This is analogous to what we noted in the double-well case except that now, instead of just having two possible linear combinations of the naive states, we have a whole continuum of them.

In this section we have learned that the classical vacua of a system do not always correspond to the quantum mechanical ones. In basic quantum mechanics we learn that for "small" $\hbar$ a particle will tend to be "smeared" around its classical vacuum equilibrium point. The more orders of $\hbar$ we keep in our approximation, the better we know the shape of the wave function and its energy. Instantons tell us, however, that tunneling effects drastically modify this picture. The particle will actually tend to be "smeared" around all of its classical vacua, thereby restoring the symmetry of the theory. We could have never seen this effect in an order-by-order approximation of the wave function in $\hbar$. This effect is non-perturbative.

In the next section we will see that gauge theories can also have tunneling effects that modify the vacuum structure.

2.2 Yang-Mills instantons

Now that we have seen the basics about instantons through simple examples, we are ready to take a look at a more sophisticated example. We will study instantons in a quantum field theory; specifically Yang-Mills theory. Although everything we have seen up to now in this chapter were instantons in quantum mechanics, we will be able to generalize the knowledge we have gathered to field theories very easily, thanks to the wonderful language of path integrals. This section will not be as technical as the previous one, as it is only meant to illustrate how the *ideas* we have seen so far apply to Yang-Mills theory. For an introduction to Yang-Mills theory and a full derivation of the Yang-Mills instanton and all of its properties, the reader is again referred to [14] and [15].

The goal is to find the vacuum structure of the Yang-Mills quantum field theory. We will work specifically with the structure group SU(2), because the results can be generalized for SU(N) with arbitrary N. The action is the following:

$$S_{YM} = -\frac{1}{2g^2}\int d^4x\, Tr\left[F_{\mu\nu}F^{\mu\nu}\right]\ , \tag{2.44}$$

where g is the coupling constant of the theory. $F_{\mu\nu}$ is the field-strength defined as follows:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]\,, \tag{2.45}$$

and the conection $A_m(x)$ is a Lie algebra valued vector field:

$$A_\mu = g\, A_\mu^a\, T^a\,, \tag{2.46}$$

The T^a are the generators of SU(2), which can be expressed in terms of the Pauli matrices as $T^a = -i\,\sigma^a/2$. SU(2) is a connected manifold, so any group element can be written in terms of the Lie algebra as follows:

$$g(x) = \exp(\alpha^a(x)\, T^a)\,, \tag{2.47}$$

where the $\alpha^a(x)$ are arbitrary smooth functions. The trace in (2.44) runs over SU(2) indices. The action (2.44) is invariant under the following gauge transformations of A_μ:

$$A_\mu \to g\, A_\mu\, g^{-1} + g\, \partial_\mu g^{-1}\,, \tag{2.48}$$

under which the field-strength transforms as follows:

$$F_{\mu\nu} \to g\, F_{\mu\nu}\, g^{-1}\,. \tag{2.49}$$

There are two kinds of gauge transformations, which we must distinguish: "small" and "large" gauge transformations. "Small" gauge transformations are those that satisfy $\alpha(|\vec{x}| = \infty) = 0$. Those that do not satisfy this restriction are denominated "large" gauge transformations. The reader should note that the physical interpretation of a gauge symmetry is different from that of a global symmetry. A global symmetry relates physically inequivalent solutions of a system. In a gauge theory, however, one considers configurations that are related via "small" gauge transformations as being physically equivalent. In fact, the physical states (in the classical sense) are defined by the gauge equivalence classes (equivalence under "small" g. t.'s) of the solutions for the gauge field.

First things first, we need to understand the classical vacua of this system. To simplify the task we take the so-called *static* gauge $A_0 = 0$, which is left invariant by time-independent gauge transformations. Now we can rewrite the Lagrangian density for (2.44) as follows:

$$\mathcal{L} = \frac{1}{g^2}\, Tr\left(\tfrac{1}{2}\,(\partial_0 A_i)^2 - \tfrac{1}{4}\, F_{ij}\, F_{ij}\right)\,. \tag{2.50}$$

This looks like the kinetic term minus a potential for the A_i fields. So we immediately notice that the classical vacua of this action are the so-called *static pure gauges*. Static, means $A_i(x) = A_i(\vec{x})$, and pure gauge means gauge equivalent to $A_i = 0$. These configurations can be written as follows:

$$A_i(\vec{x}) = e^{-\alpha(\vec{x})}\partial_i\, e^{\alpha(\vec{x})} \quad \text{where} \quad \alpha(\vec{x}) = \alpha^a(\vec{x})\, T^a\,. \tag{2.51}$$

It can be shown that it is enough to restrict our search to configurations that satisfy $e^\alpha = \mathbb{I}$ at spatial infinity $|\vec{x}| = \infty$. Since α tends toward the same value in any direction at spatial infinity, we can actually identify all of spatial infinity to a point. In other words, we can reformulate the problem of finding the static pure gauges with $\alpha \to 0$ for $|\vec{x}| \to \infty$ as the problem of finding maps from S^3 into SU(2). As a manifold, SU(2) is diffeomorphic to S^3; hence we are looking for maps $\alpha : S^3 \to S^3$. Maps that are homotopic (can be continuously deformed into each other) correspond to field configurations that are related by "small" gauge transformations. Hence, the vacua can be classified in homotopy classes. In this case, the homotopy group is $\Pi_3(S^3) \cong \mathbb{Z}$. To each homotopy class we can associate an integer, which counts the number of times S^3 is "wrapped" around S^3 by the map α. Given such a map, its homotopy class is determined by computing the following:

$$N = \frac{1}{24\pi^2} \int_{S^3} d^3x\, \epsilon_{ijk}\, Tr[(e^{-\alpha}\partial_i e^{\alpha})(e^{-\alpha}\partial_j e^{\alpha})(e^{-\alpha}\partial_k e^{\alpha})]\,. \tag{2.52}$$

This is called the Pontryagin index, it literally yields the integer representing the homotopy class of the vacuum configuration. Because a homotopy class is invariant under continuous deformations one usually calls these configurations *topological* vacua. The classical N-vacuum can be thought of as the analogue of the $x = 2\pi N$ vacuum in the periodic potential problem. They are physically inequivalent because no "small" gauge transformation can relate them. However, they can be related via "larger" gauge transformations, just like $x = 2\pi N$ is related to $x = 2\pi(N+1)$ via a 2π shift. From the classical N-vacuum, one can build a naive perturbative quantum state $|N\rangle$, just as we did with in the previous examples, and deduce that the vacuum is infinitely degenerate. However, Yang-Mills theory also has instantons, and tunneling between the different $|N\rangle$ states takes place. By computing tunneling amplitudes in analogy with the periodic potential problem, one sees that the true low energy eigenstates form a band parametrized by an angle θ; and in terms of the $|N\rangle$, a θ-state is given by the following:

$$|\theta\rangle = \sum_N e^{i\theta N}\, |N\rangle\;, \tag{2.53}$$

which restores the symmetry under "large" gauge transformation. This is analogous to the restoration of the $\mathbb{Z}$-symmetry by the θ-vacua of the periodic potential system. One other important property of these θ-states is that they can never talk to each other. In other words, there can never be a physical transition from one such state to another. For any gauge invariant operator B, it can be shown that

$$\langle\theta\,|\,B\,|\,\theta'\rangle = 0\,, \tag{2.54}$$

for any choice of θ and θ'. Therefore, we can make a paradigm shift and consider each $|\theta\rangle$ as the vacuum of a separate theory. For each value of θ we have a theory whose *unique* vacuum state is $|\theta\rangle$. In quantum field theory, one is interested in the vacuum-to-vacuum amplitude $\langle 0\,|\,e^{-HT/\hbar}\,|\,0\rangle$, also known as the partition function Z. In this case, to compute the partition function we have to choose a theory by choosing a value of θ and use its vacuum state. Then, we can write Z as follows:

$$Z = \langle\theta\,|\,e^{-HT/\hbar}\,|\,\theta\rangle = \sum_{N,Q} e^{-i\theta Q}\,\langle N+Q\,|\,e^{-HT/\hbar}\,|\,N\rangle\;, \tag{2.55}$$

using the fact that $\langle N+Q\,|\,e^{-HT/\hbar}\,|\,N\rangle$ is independent of N[5] , we write

$$Z = K\sum_Q e^{-i\theta Q}\int_Q d[A_\mu]\, e^{-S_E}\,, \tag{2.56}$$

[5]The amplitude is invariant under all gauge transformations, "small" and "large", because the Yang-Mills action is. Since N can be changed to any value via a "large" gauge transformation, the amplitude must be independent of N

where S_E is the Euclidean version of (2.44), and the subscript Q indicates that the path integral corresponds to a tunneling amplitude between two topological states whose Pontryagin indices differ by Q. K is just an normalization constant encoding the infinity coming from the summation over N. It is not physically relevant, as all quantum field theoretic amplitudes are normalized by dividing by Z.

Let us summarize what we have learned so far. Yang-Mills theory for SU(2) has classical vacua, which are classified by the third homotopy group of the 3-sphere $\Pi_3(S^3)$. Each class consists of static pure gauge field configurations, which are related by "small" gauge transformations, and it is labeled by the Pontryagin index N. For each N, we have a topological naive vacuum, which can tunnel into another topological naive vacuum, and, just as in the case of the periodic potential, the true energy eigenstates are combinations of the $|N\rangle$, labeled by an angle θ. Since different θ-vacua can never physically interact, we consider θ as a parameter labeling a theory, whose *unique* vacuum is $|\theta\rangle$. To compute the partition function of the theory, we have to sum over all possible tunneling amplitudes, weighing each by $e^{-i\theta Q}$. However, this whole language of topological $|N\rangle$ states is not gauge invariant. It only works in the static gauge. Therefore, the partition function as we wrote it in (2.56) is not gauge invariant. Fortunately, there is a way to remedy this.

Instead of classifying classical vacua, let us classify instantons; i.e. finite action Euclidean configurations. In order for a field configuration to have finite action, its Lagrangian density must be non-zero only in a localized area and vanish at the boundary of Euclidean spacetime. The Euclidean version of the Yang-Mills action (2.44) is positive-definite:

$$S_E = -\frac{1}{2g^2} \int d^4x\, Tr\left[F_{\mu\nu} F^{\mu\nu}\right] = -\frac{1}{2g^2} \int d^4x\, Tr\left[F_{\mu\nu} F_{\mu\nu}\right] . \tag{2.57}$$

The minus sign is due to the fact that the SU(2) trace is negative in the basis we have chosen. Hence, in order for a configuration to have $\mathcal{L}$ vanish at infinity it must be pure gauge at infinity. It must satisfy the following:

$$A_\mu(x) \to g^{-1}(x)\, \partial_\mu\, g(x)\,, \tag{2.58}$$

$$\text{as} \quad |x| \to \infty\,. \tag{2.59}$$

If we define the boundary of R^4 as a 3-sphere[6] whose radius is taken to infinity, then we can think of instanton configurations as maps $g : S^3 \to \mathrm{SU}(2) = S^3$. These are again classified by $\Pi_3(S^3) = \mathbb{Z}$. The Pontryagin index can be computed using formula (2.52), but this time integrating over the S^3 that represents the spacetime boundary. Since we are integrating over the boundary, we can use Stokes' theorem and rewrite the formula as a total derivative:

$$Q = -\frac{1}{24\pi^2} \int_{\partial\mathbb{R}^4} d^3x\, \epsilon^{\nu\rho\sigma}\, Tr[A_\nu A_\rho A_\sigma] = -\frac{1}{24\pi^2} \int_{\mathbb{R}^4} d^4x\, \epsilon^{\mu\nu\rho\sigma}\, Tr[\partial_\mu A_\nu A_\rho A_\sigma]\,, \tag{2.60}$$

which can be shown to be equivalent to

$$Q = -\frac{1}{16\pi^2} \int_{R^4} d^4x\, \epsilon^{\mu\nu\rho\sigma}\, Tr[F_{\mu\nu} F_{\rho\sigma}] = -\frac{1}{16\pi^2} \int_{R^4} d^4x\, Tr[F_{\mu\nu} \tilde{F}^{\mu\nu}]\,, \tag{2.61}$$

[6] A note of caution: the S^3 we previously considered was a one-point compactification of the *space* R^3, which we used in order to classify the state of the system at a certain point in time. The S^3 we are considering now is the boundary of Euclidean space-*time* $\mathbb{R}^4$, which we are using in order to classify instanton configurations.

where $\tilde{F}^{\mu\nu} \equiv \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}$ is the Hodge dual of the field-strength. This expression is manifestly gauge invariant. It is also known as the second Chern class, due to its interpretation as the characteristic class of an SU(2)-principal bundle over the base manifold S^4.

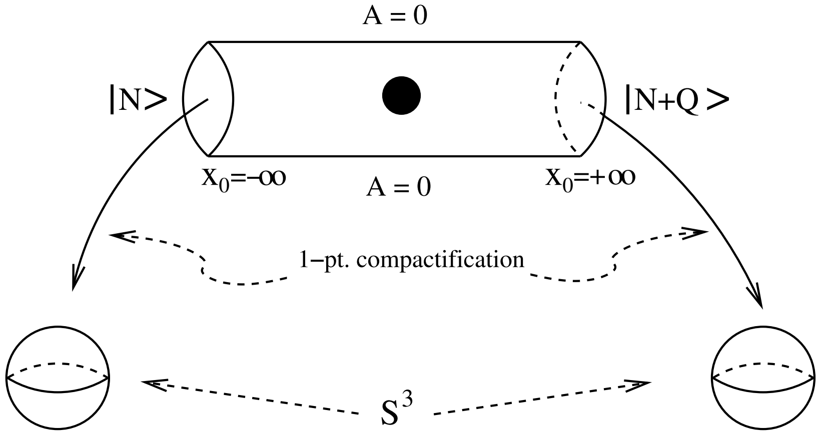


Figure 2.5: *The boundary of spacetime as a time-like cylinder $\mathbb{R} \times S^2$, with one suppressed dimension. The initial and final topological states reside at the caps of the cylinder. The latter, which are two D^3, are compactified to two S^3 to determine their topological indices N and $N + Q$, respectively. The black filled circle represents the localization of the instanton.*

This topological term classifies the boundary conditions of all instanton configurations in a gauge invariant way. However, any such configuration with second Chern class Q can be interpreted as a tunneling process from a topological state $|N\rangle$ to a state $|N + Q\rangle$ by performing a gauge transformation to go to the static gauge. In the static gauge, if we view the boundary of Euclidean spacetime as a generalized cylinder $\mathbb{R} \times S^2$ as in figure 2.5, where $\mathbb{R}$ is the Euclidean time range, then the only contribution to (2.61) will come from the two 3-discs at $x_0 = \pm\infty$ (i.e. the caps of the cylinder):

$$Q = -\frac{1}{16\pi^2} \int_{D^3} d^3x\, \epsilon^{\nu\rho\sigma} \; Tr[A_\nu A_\rho A_\sigma]\Big|_{x=-\infty}^{x=\infty} , \tag{2.62}$$

$$= (N + Q) - N . \tag{2.63}$$

Hence, the second Chern class computes the change in N of the tunneling process. We can now finally rewrite the partition function (2.56) in a gauge invariant way:

$$Z = \int_{\text{all } Q} d[A_\mu] \exp\left[-S_E - i\,\frac{\theta}{16\pi^2} \int d^4x Tr[F_{\mu\nu}\tilde{F}^{\mu\nu}]\right] . \tag{2.64}$$

The θ-term has a physical effect on the theory. It breaks parity. This actually makes θ a physically measurable quantity in gauge theories.

2.3 Solitons vs. instantons

Having studied the mathematics and physics of instantons, we should also look at a special class of solutions to classical equations of motion called *solitons*. These will be interesting to us for a number of reasons: first of all, they have a similar mathematical structure to instantons in that they are *topologically non-trivial*. They too, are in some sense interpolating configurations. Secondly, in some cases, there exists a precise correspondence between instantons and solitons. In the next chapter, we will actually see an explicit example of this. Because solitons are not the main focus of this text, I will only briefly introduce them and will refer the interested reader to Coleman 's book [14] and Rajaraman's book [15] for a careful introduction, and Zee's book [17] for a short but very clear exposition of the topic.

2.3.1 Solitons: Definition and examples

Definition: A soliton is a time-independent extremum of the *Mikowskian* action with finite non-zero energy.[7]

Note that we are now back to Minkowski spacetime. Time-independent means that the field configuration has no non-trivial time-dependence that could for instance be obtained by boosting a static solution.

Let us take a look at the simplest soliton, the *kink* solution. We define the following field theory in (1 + 1)-dimensions:

$$\mathcal{L} = -\tfrac{1}{2}\,(\partial\phi)^2 - V(\phi)\,, \tag{2.65}$$

with

$$V(\phi) = \frac{\lambda}{4}\left(\phi^2 - v^2\right)^2\,, \tag{2.66}$$

where ϕ is the field, and λ and v are parameters. This is a double-well potential. Note that we are working in the *mostly plus* convention, which is why the kinetic term has a minus sign. We instinctively know that this Lagrangian has two very simple solutions, namely the two vacua $\phi(t,x) = \pm v$. They both have energy zero. In standard perturbation theory we are instructed to pick one of the two vacua and study the fluctuations around it. In practice, this means rewriting the scalar field as $\phi \to v + \chi$, and treating the fluctuation χ as the fundamental field. Plugging this back into (2.65) we will find that χ is a scalar particle with mass $\mu = (\lambda\, v^2)^{1/2}$.

One can, however, also look for a solution with non-trivial conditions, namely a configuration that interpolates between those two vacua, i.e. $\phi \to \pm v$ for $x \to \pm\infty$. Such a solution will look qualitatively like the kink we saw in section 2.1, see figure 2.6. In fact, this solution is also known as the kink solution. Because it is time-independent, we can write its energy density as follows:

$$\mathcal{E} = \tfrac{1}{2}\,\phi'^2 + V\,, \tag{2.67}$$

where the prime denotes differentiation w.r.t. the spatial coordinate x. Because $\phi \to \pm v$ for $x \to \pm\infty$, the energy density is non-zero only within a localized region. This means that the

[7] This is not the only possible definition. A stricter one, stated in [15], also requires that a soliton's shape be left unaffected by scattering against another soliton, but we will not be exploring this property here.

total energy will be finite. Since this energy density is positive, we can rewrite it as a square plus a positive term:

$$\tfrac{1}{2}\left(\phi' \pm \sqrt{2\,V}\right)^2 \mp \phi' \sqrt{2\,V}\,. \tag{2.68}$$

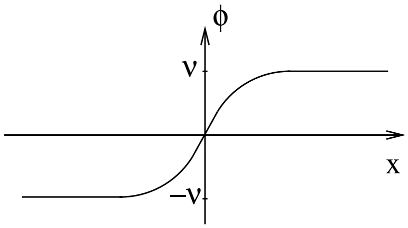


Figure 2.6: *The kink solution: a classical field configuration trajectory that interpolates between the two classical vacua of the double-well potential.*

This means that the energy of any solution to this system satisfies a bound:

$$E \geq \left|\int dx\,\phi' \sqrt{2\,V}\right| = \left|\int_{\phi(x=-\infty)}^{\phi(x=+\infty)} d\phi\, \sqrt{2\,V}\right|\,. \tag{2.69}$$

This is known as the *Bogomol'nyi bound*. Because we are choosing a time-independent Ansatz, we can easily see that the Lagrangian density of this system (2.65) is equal to minus the energy density (2.67), i.e. $\mathcal{L} = -\mathcal{E}$. This is more than a mere curiosity, this is at the heart of the instanton-soliton correspondence. Therefore, solving the equations of motion with this Ansatz means extremizing not only the action, but also the energy. This means that the soliton actually saturates the Bogomol'nyi bound (2.69). In other words, a soliton is the configuration of least energy within its class of boundary conditions or topological class. To saturate the bound, the field has to satisfy:

$$\phi' = \pm\sqrt{2\,V}\,. \tag{2.70}$$

This is often referred to as the BPS condition. Note that if a field satisfies this equation, it automatically satisfies the equations of motions. However, we have now simplified the task of solving a second order differential equation into solving a first order equation. In supergravity, p-branes are solutions, which satisfy an analogous form of the BPS condition. The latter implies that the solution preserves a certain amount of the supersymmetry of the theory it lives in. Using (2.68) and (2.70) we find that the energy is given by:

$$E = \left|\int_{\phi=-\nu}^{\phi=\nu} d\phi\, \sqrt{2\,V}\right|\,. \tag{2.71}$$

This depends only on the potential and the boundary conditions, and not on any parameters of the solution. In our case, $E \sim \mu^3/\lambda$. So the kink is very massive (energetic) for small coupling constant. This means that object is non-perturbative, i.e. it cannot be found by doing some sort of pertubation theory around the vacuum. The kink is at least pertubatively a stable configuration. Its non-trivial boundary conditions prevent it from simply decaying into an object with lower energy. It is not a simple ripple in the field. Mathematically this translates into the statement that the kink has a conserved *topological current*[8]

$$J^\mu = \frac{1}{2}\,\epsilon^{\mu\nu}\,\partial_\nu \phi\,, \tag{2.72}$$

yielding a conserved *topological charge*

$$Q = \int_{+\infty}^{-\infty} dx\, J^0 = \frac{1}{2\,v}\Big(\phi(+\infty) - \phi(-\infty)\Big). \tag{2.73}$$

Solitons are also present in more complicated field theories, such as gauge theories. Magnetic monopoles are an example of solitons. Depending on the dimensionality of the soliton it may be called, *monopole, string or vortex, membrane*, or *texture*, if it 'stretches' over $0, 1, 2$ and 3 spatial directions respectively. If it only has one transverse spatial direction, such as the kink in 1 + 1 dimensions, it is called a *domain wall*. All of these objects are characterized by some topological charge. In gauge theories this charge will be a Pontryagin index.

In gravitational theories, there are objects analogous to solitons. The simplest one is the Schwarzschild black hole. Its metric is the following:

$$ds^2 = -\left(1 - \frac{2\,G\,M}{r}\right) dt^2 + \left(1 - \frac{2\,G\,M}{r}\right)^{-1} dr^2 + r^2\, d\Omega^2_{S^2}\,, \tag{2.74}$$

where G is the Newton constant and M is a parameter of the solution. For an introduction to black holes, the reader is referred to the pedagogical lecture notes by S. Carroll [18] (or his book [19]), and to Townsend's extensive lecture notes [20]. The spacetime geometry of the Schwarzschild black hole is non-trivial in that it interpolates between flat Minkowski spacetime at spatial infinity, and $AdS_2 \times S^2$ near its horizon at $r = 2\,G\,M$. Although energy is a tricky subject in General Relativity, it can be defined via the ADM mass formula, which can be found in [20]. Once it is calculated, one finds that it is equal to the parameter M in the solution for the Schwarzschild metric (2.74). From the solution, we see that this object is also non-perturbative. No matter how 'small' we make the mass, its effect will be very dramatic near the horizon. In supergravity, p-branes play the role of the soliton. They are the higher-dimensional generalization of the charged Reissner-Nordström black hole. A p-brane has a $p+1$-dimensional world-volume and is charged under a $p+2$-form field-strength. For an introduction into p-brane solutions, the reader is referred to "String Solitons" [21], and to "Gravity and Strings" [22].

2.3.2 The correspondence

Now that we have seen the definition of solitons and have seen some examples of them, let us study their correspondence with instantons. We will first look at the simplest example of this correspondence, and then explain it in a more general context.

Taking the example of the scalar field in 1 + 1 dimensions from the previous subsection, the reader will recall that if we make take the time-independent Ansatz, which is what we do

[8] Note that this current is not a Noether current, as it does not follow from a continous symmetry.

when looking for solitons, and substitute it into the Lagrangian density (2.65), the latter takes the following form:

$$\mathcal{L} = -\tfrac{1}{2}\,\partial_x\phi^2 - V = -\mathcal{E}\,, \tag{2.75}$$

where $\phi = \phi(x)$, and $\mathcal{E}$ is the energy density of the system. A soliton is defined as being an extremum of the action defined by this Lagrangian density *and* as having finite energy. Note that this Lagrangian density is, up to a minus sign, equivalent to that of a scalar field in *one* Euclidean dimension if we define Euclidean time τ as $\tau \equiv x$. Hence, the equations of motion for a soliton in $1+1$ dimensions are the same as the equations for an instanton in one Euclidean dimension, and the requirement that the soliton have finite *energy*

$$E = \int dx\,\mathcal{E}\,, \tag{2.76}$$

is equivalent to the requirement that the instanton in one dimension have finite *action*. So the kink-soliton in $1+1$ dimensions corresponds to the instanton in one dimension[9]. The relation is simply $\phi_{sol}(x) = \phi_{inst}(\tau)$.

This is not specific to the kink model, one can show a more general correspondence. Let us define a system in $d+1$ spacetime dimensions with general degrees of freedom, which we denote by ϕ_I, where the I can stand for a collection Lorentz indices, or internal indices, and a Lagrangian density

$$\mathcal{L} = \mathcal{L}(\phi_I, \partial\phi_I)\,, \tag{2.77}$$

where both temporal and spatial derivatives are implied by the symbol '∂'. The conjugate momenta of the system are defined as follows:

$$\pi_I^\mu \equiv \frac{\delta\mathcal{L}}{\delta\left(\partial_\mu\phi_I\right)}\,. \tag{2.78}$$

Using the time-dependent Ansatz we can write the energy as follows:

$$E = \int d^dx\,\left[\pi_I^0\,\dot{\phi}_I(t,\vec{x}) - \mathcal{L}\left(\phi_I(t,\vec{x})\right)\right] = -\int d^dx\mathcal{L}\left(\phi(\vec{x})\right)\,, \tag{2.79}$$

where the dot, as usual, represents a time derivative, and the first term on the LHS vanishes due to the time-independence of the solution. A soliton solution will be an extremum of this energy (since S=-E), and will have finite energy. Since all degrees of freedom depend only on the spatial directions, we can view this Lagrangian density as that of a d-dimensional Euclidean system (up to a minus sign), and this energy can be viewed as its action. The soliton can then be called an instanton in d dimensions. In practice, all one has to do is a Kaluza-Klein reduction over time, but without the interpretation that time is compactified. One is simply truncating time.

To summarize all this, the statement is the following: *A soliton in $d+1$ dimensions is equivalent to an instanton in d dimensions*. In the next chapter, we will see that charged black holes can be viewed as a certain kind of supergravity instantons called *D-instantons*. An interesting

[9]The kink instanton solution in $(1+1)$-dimensional quantum mechanics can be viewed as an instanton in $(0+1)$-dimensional quantum field theory.

question that comes to mind based on the statement we have made, is whether its converse is true. In other words: *When is an instanton in d dimensions equivalent to a soliton in $d + 1$ dimensions*? The answer depends on the Lagrangian. If a Euclidean Lagrangian can be obtained as the time truncation of a $d + 1$-dimensional Lagrangian, in other words, if it can be *uplifted* to $d + 1$ dimensions, then the instanton will give rise to a soliton. In the next chapter we will establish the necessary condition for a D-instanton to give rise to a black hole.

In this chapter, we studied the basics about instantons in quantum mechanics and quantum field theory. We learned that instantons provide us with non-perturbative information, by telling us that a naïve perturbative vacuum is not really the vacuum state of a theory, because the system can tunnel out of it. This requires that one rewrite a path integral with a new topological term that properly takes this fact into account.

In the next chapter, we will be looking at instantons in gravitational field theories, such as supergravities. Although defining a path integral for a gravitational theory is tricky business and requires unnatural adjustments in order to be well-defined, it is possible to talk about instantons and non-perturbative tunneling effects in gravity.

Chapter 3

Non-extremal D-instantons

3.1 Introduction

In the previous chapter, we studied instantons in quantum mechanics and quantum field theory. In this chapter we will be looking at instantons in gravitational theories. Instantons, as we have seen, are inherently linked to path integrals. However, a path integral formulation of quantum gravity is not as straight forward as one might wish. In an ideal world, we would simply write down the following:

$$\langle\, h_F \,|\, e^{-HT} \,|\, h_I \rangle = \int d[g] \exp\left(- \int d^D x\, R \right) , \tag{3.1}$$

where $h_{I,F}$ are the induced metrics on the initial and final spacelike hypersurfaces of spacetime, respectively, R is the Ricci scalar, and the path integral sums over all metrics satisfying the boundary condition that they asymptote to $h_{I,F}$ in the early past and late future, respectively. However, this path integral is not well-defined because the action is not bounded from below. In fact, even flat Euclidean space is not a minimum of the Einstein-Hilbert action. Suppose we wanted to perform a semiclassical approximation around the Wick rotated Minkowski spacetime, i.e. flat Euclidean space. There are infinitely many possible fluctuations around the flat metric, but let us restrict to summing over metrics that are related to flat space via a Weyl transformation; i.e. *conformally flat* metrics:

$$\tilde{g} = e^{2\sigma}\, \eta \,, \tag{3.2}$$

where η is the flat metric. Then, the action for $\tilde{g}$ will roughly go as follows:

$$\int d^D x R \sim - \int d^D x\, (\partial\sigma)^2 \ , \tag{3.3}$$

which means that the action can be made arbitrarily negative by quantum fluctuations, making flat spacetime a local maximum (or at best a saddle point), and making the whole path integral divergent. Fixing this problem requires a new formalism, which is developed in [23], but is not yet widely agreed upon. The idea is to first sum over conformal classes of metrics, and then, within each class of conformally related metrics, one rotates the contour of integration to imaginary conformal factors. In (3.3) this manifests itself in that only imaginary σ are allowed, thus keeping the action positive. We will not really be using any of this formalism in this thesis. The purpose of this paragraph was to show how severely different path integration becomes when dealing with gravity.

Despite difficulties with path integrals, gravitational instantons do exist and have been applied to many different problems in quantum gravity such as the renormalization of the constants of nature, the adjustment of the cosmological constant, spacetime topology fluctuations, and the creation of baby universes (see [24–28]).

In the field theory limit of string theory, instantons can give rise to non-perturbative effects (for an overview see [29]). The standard *D-instanton* is an instanton solution of type IIB supergravity, which was discovered in [30], and was later shown to give higher derivative correction terms, specifically R^4 terms, to the effective action of type IIB string theory [31]. The coefficient of such terms was conjectured to be an $\mathrm{SL}(2,\mathbb{Z})$ invariant modular function. In [32], the high-energy limit of this conjecture was tested. Other instantons have been obtained through dimensional reductions of supergravity by wrapping Euclidean D-branes around compact cycles of the internal space. This yields non-perturbative effects, which give rise to interesting lower-dimensional effective actions that have applications in cosmology [33].

The standard D-instanton is a solution of a truncation of type IIB supergravity with the metric, the dilaton, and the RR scalar known as *axion* as its field content. The solution has a flat Euclidean metric, preserves $1/2$ of the supersymmetry of the theory, and is characterized by the axion 'charge'[1]. The fact that it is 'charged' under a 0-form potential makes the D-instanton mathematically similar to p-branes. In this case it, could be thought of as a (-1)-brane, meaning it is localized in space *and* time. In this chapter, we will be studying solutions that generalize the standard D-instanton in many ways: their metrics will be non-trivial, and they will not preserve any supersymmetry. The solutions that will be presented are not new, but will be studied in a novel way. For earlier work on generalized D-instanton solutions see [25, 28, 34–41]

In this chapter, we will generalize the Lagrangian of type IIB supergravity to arbitrary dimensions, and arbitrary dilaton coupling. However, one important property of type IIB supergravity will be preserved: the scalars (dilaton and axion) are coupled in such a way that they parametrize an $\mathrm{SL}(2,\mathbb{R})/\,\mathrm{SO}(1,1)$ coset space. By conveniently reorganizing the fields into 2×2 matrices, the $\mathrm{SL}(2,\mathbb{R})$ symmetry will become manifest, and we will see that solutions to the field equations will have a 'conserved' *charge matrix* Q, as implied by Noether's theorem. This charge matrix Q transforms under the adjoint representation of $\mathrm{SL}(2,\mathbb{R})$, which means that its determinant is invariant under the symmetry. This implies that there are three families of solutions that are not related via $\mathrm{SL}(2,\mathbb{R})$, i.e. those with $\det Q > 0, = 0$ and < 0. This is analogous to the fact that Minkowski spacetime admits three families of vectors: Timelike, lightlike, and spacelike. In this chapter, we will see that all D-instanton solutions can be classified into three classes, whereby the standard D-instanton falls under the $\det Q = 0$ class.

A similar discovery was made in [42], where three classes of $\mathrm{SL}(2,\mathbb{R})$-unrelated seven-branes were found. Seven-branes can be seen as the magnetic duals of D-instantons. They are carried by the same fields; however, instead of being *electrically* charged under the axion, they

[1]We will give this 'charge' a physical interpretation later on.

are *magnetically* charged under it. This means that, in contrast to the D-instantons, seven-branes are not localized in spacetime. Given that seven-branes were shown to occupy all three possible *conjugacy* classes of $SL(2,\mathbb{R})$, it is natural to ask whether D-instantons do the same.

At the end of chapter 2 we saw that instantons in D Euclidean dimensions can sometimes be viewed as the spacelike sections of solitons in $D+1$ spacetime dimensions. In this chapter, we will show that the three $SL(2,\mathbb{R})$ classes of D-instantons can sometimes be seen as spacelike sections of electrically charged black holes, i.e. Reissner-Nordström black holes. As we will see, the three families of solutions, $\det Q > 0, = 0, < 0$, correspond to underextremal, extremal, and overextremal black holes (i.e. black holes with electric charges lower than, equal to, and greater than their masses). The condition for such a correspondence to hold will be worked out, and the correspondence will be extended to *uplift* the D-instantons to p-branes in higher dimensions.

This chapter is based on a collaboration with E. Bergshoeff, U. Gran, D. Roest, and S. Vandoren, entitled *Non-extremal D-instantons* [43]. It is organized as follows: in section 3.2, we will present the metric-scalar system we are interested in and discuss the realization of the $SL(2,\mathbb{R})$-duality group for the Euclidean case. In section 3.3, we will give the generalized instanton solutions mentioned above. At this point we only construct the bulk solutions without taking care of boundary terms and/or boundary conditions. Next, in section 3.4, we will discuss the relation to wormholes corresponding to non-extremal Reissner-Nordström black holes one dimension higher. In section 3.5 we will consider generalizations that uplift to non-extremal p-branes in $D+p+1$ dimensions. The application as true instantons of type IIB string theory will be investigated in section 3.6. Finally, we will discuss our results in section 3.7.

3.2 The system and its symmetries

3.2.1 Lagrangian

The system we will be interested in is described by the following Minkowskian Lagrangian density:

$$\mathcal{L}_M = \tfrac{1}{2}\sqrt{|g|}\,[R - \tfrac{1}{2}(\partial\phi)^2 - \tfrac{1}{2}e^{b\phi}(\partial\chi)^2]\,, \tag{3.4}$$

where ϕ and χ are scalars. We will work in D arbitrary dimensions, and will keep the coupling b unspecified. This theory occurs, for example, as the scalar section of IIB supergravity in $D=10$ Minkowski spacetime with coupling parameter $b=2$. In this case, the scalar ϕ corresponds to the string theory *dilaton*, and the scalar χ is the Ramond-Ramond scalar known as the *axion*. Other values of b can arise when considering (truncations of) compactifications of IIB supergravity. For instance, in $D=3$ one has supersymmetry for $b=2, b=\sqrt{2}, b=\sqrt{4/3}$ and $b=1$. In order to study instanton solutions of this system we not only need to Wick rotate the theory, but we also need to change the sign of the axion kinetic term, yielding the following Euclidean Lagrangian:

$$\mathcal{L}_E = \tfrac{1}{2}\sqrt{|g|}\,[R - \tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}e^{b\phi}(\partial\chi)^2]\,, \tag{3.5}$$

The effect of the Wick rotation on the scalar is a very subtle issue, which I will further develop in section 3.6. I will now summarize the three basic arguments to justify the sign change in the kinetic term:

- In the context of type IIB supergravity the axion is considered a *pseudo*scalar. In that case one could claim that the Wick rotation is the 'square root of time reversal', and hence a pseudoscalar should get multiplied by an 'i' upon transforming. This argument, however, is neither rigorous, nor widely agreed upon. Since we want to study D-instanton solutions in theories with arbitrary D and b that are not necessarily imbeddable in supergravity, we will not endorse this claim.

- A theory with a scalar is *dual* to a theory with a $(D-1)$-form field strength. Dual means that there exists a procedure to show that the path integrals of the two theories are equivalent. This procedure allows one to move back and forth from the one path integral to the other. In our case, the Lagrangian of the dual theory is the following:

$$\mathcal{L} = \tfrac{1}{2}\sqrt{|g|}\,[R - \frac{1}{2}(\partial\phi)^2 - \frac{1}{2\cdot(D-1)!}e^{-b\phi}F_{D-1}^2]\,, \tag{3.6}$$

where F_{D-1} is a $(D-1)$-form field-strength. Contrary to common belief, the quantum mechanical dualization *starting from* the $(D-1)$-form theory *does not* yield a scalar theory with the wrong kinetic term sign, but a scalar theory with the normal sign. However, one quickly notices that the Euclidean scalar theory does not have any non-trivial *real* saddle points, so instead of performing the semiclassical approximation on the scalar theory, one does it on the dual $(D-1)$-form theory, which does have non-trivial real saddle points. After writing down the classical Euclidean equations of motion to do the semiclassical approximation one notices that, if one rewrites the $(D-1)$-form field-strength as the *Hodge*-dual of a 1-form field-strength as follows:

$$F_{D-1} = -e^{b\phi} \,*\, d\chi\,, \tag{3.7}$$

then the Euclidean equations of motion of the $(D-1)$-form look like the equations of motion of a would-be scalar theory with the wrong sign for the kinetic term. In other words, looking for the saddle points of the $(D-1)$-form theory is *effectively* the same as looking for the saddle points of (3.5). I would like to emphasize that quantum mechanical dualization and Hodge dualization are two different things.

- In a quantum field theory, imposing Dirichlet boundary conditions on the field yields a transition amplitude between eigenstates of the field operators. In our case, this means that the path integral is actually computing the following:

$$\langle\,\phi_F,\chi_F\,|\,e^{-HT}\,|\,\phi_I,\chi_I\rangle\ . \tag{3.8}$$

However, one can also compute a transition amplitude between axionic charge-eigenstates by means of Fourier transformation:

$$\langle\,\phi_F,\pi_F\,|\,e^{-HT}\,|\,\phi_I,\pi_I\rangle = \\ \int d[\chi_I]\,d[\chi_F]\,\exp\left(-i\int_{\Sigma_I}\pi_I\,\chi_I + i\int_{\Sigma_F}\pi_F\,\chi_F\right)\langle\,\phi_F,\chi_F\,|\,e^{-HT}\,|\,\phi_I,\chi_I\rangle \tag{3.9}$$

where the path integral over $\chi_{I,F}$ runs over functions defined on the initial and final time hypersurfaces Σ_I and Σ_F, respectively; and $\pi_{I,F}$ are the time components of the conjugate momenta of the axion. This theory has no boundary conditions. The path integral (3.9) has no real saddle points. However, it can be computed in the semiclassical approximation; and it can be shown that the result of this path integration can also be obtained by looking for the saddle points of a would-be system with the wrong kinetic term sign (3.5). Effectively, it is as if we were looking for imaginary saddle points of the original system. This argument was first discovered by Lee in [44]. In [45–47] the argument was refined; however, the clearest and simplest explanation, in my view, can be found in [48].

In section 3.6.1, we will further develop the second method in order to evaluate the actions of our solutions, and in appendix A, a toy model will be used to illustrate the phenomenon of the 'wrong' sign in a simpler setting.

3.2.2 $\mathrm{SL}(2,\mathbb{R})$-symmetry

The Lagrangian (3.5) has a manifest $\mathrm{SL}(2,\mathbb{R})$ symmetry. In fact, in chapter 7 we will see that the scalar sector parametrizes a two-dimesional hyperboloid with Lorentzian signature; i.e. a dS_2 spacetime. The latter can be viewed as the following coset:

$$\frac{\mathrm{SO}(2.1)}{\mathrm{SO}(1,1)}\,, \tag{3.10}$$

where $\mathrm{SO}(2,1) \cong \mathrm{SL}(2,\mathbb{R})$. In this chapter, we will making the symmetry manifest by writing the Lagangian in a different form. Define the following matrix:

$$\mathcal{M} = e^{b\phi/2}\begin{pmatrix} \frac{1}{4}b^2\chi^2 - e^{-b\phi} & \frac{1}{2}b\chi \\ \frac{1}{2}b\chi & 1 \end{pmatrix}. \tag{3.11}$$

Now we can write (3.5) as follows:

$$\mathcal{L}_E = \frac{1}{2}\sqrt{g}\,[R + b^{-2}\mathrm{Tr}(\partial\mathcal{M}\,\partial\mathcal{M}^{-1})]\,. \tag{3.12}$$

It is clear that this is invariant under the following transformation:

$$\mathcal{M} \to \Omega\,\mathcal{M}\,\Omega^T \quad \text{with} \quad \Omega = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})\,. \tag{3.13}$$

The attentive reader will probably have noticed that any invertible matrix $\Omega \in \mathrm{GL}(2,\mathbb{R})$ will do. However, only elements of $\mathrm{SL}(2,\mathbb{R})$ yield a transformed matrix $\mathcal{M}$ that is consistent with the scalar parametrization (3.11) of the coset space[2].

This symmetry, like any continuous symmetry, has Noether current:

$$J_\mu = (\partial_\mu\mathcal{M})\mathcal{M}^{-1} = \begin{pmatrix} j^{(3)}_\mu & j^{(+)}_\mu \\ -j^{(-)}_\mu & -j^{(3)}_\mu \end{pmatrix}, \tag{3.14}$$

[2]Throughout this chapter we assume that $b \neq 0$. Note that for $b = 0$ the Euclidean $\mathrm{SL}(2,\mathbb{R})$ symmetry degenerates to an $\mathrm{ISO}(1,1)$ symmetry, and the scalar coset becomes a two-dimensional Minkowski spacetime.

which is a current matrix, with the following components:

$$j_\mu^{(3)} = \tfrac{1}{2}\, e^{b\phi}\partial_\mu(e^{-b\phi} - \tfrac{1}{4}b^2\chi^2)\,, \qquad j_\mu^{(-)} = \tfrac{1}{2}be^{b\phi}\partial_\mu\chi\,,$$
$$j_\mu^{(+)} = -b\chi\, j_\mu^{(3)} + (e^{-b\phi} - \tfrac{1}{4}b^2\chi^2)\, j_\mu^{(-)}\,. \tag{3.15}$$

Althought this is a Euclidean theory, we can still regard this current as giving rise to 'charges' that are 'conserved' with respect to a Euclidean time direction. Throughout this section, we will choose it to be the radial direction. However, for a proper tunneling interpretation of the instantons, we will choose a Cartesian direction in subsection 3.6.2. For a spherical boundary defined by a radial normal unit vector n^μ, the conserved charge matrix is the following:

$$Q = \frac{(2\,(D-1)\,(D-2))^{-1/2}}{b\,\mathcal{V}ol(S^{D-1})} \int_{S^{D-1}} J_\mu n^\mu\,, \tag{3.16}$$

where the S^{D-1} is transverse to the unit vector. Under an SL(2, $\mathbb{R}$) transformation (3.13) the corresponding charge matrix transforms as

$$Q \to \Omega\, Q\, \Omega^{-1}\,. \tag{3.17}$$

Note that the determinant of Q is invariant under SL(2, $\mathbb{R}$). Thus, solutions with different values of det(Q) can never be related via SL(2, $\mathbb{R}$)-transformations. Hence, as discussed in the introduction the cases $\det(Q) = 0, \det(Q) > 0$ and $\det(Q) < 0$ define the three different 'conjugacy classes' of SL(2, $\mathbb{R}$).

3.3 The solutions and their geometries

In this section we will consider solutions to the bulk equations of motion of (3.5). Issues like boundary terms and the value of the action are postponed to section 6, where we will determine which solutions can be considered as instantons.

3.3.1 Solutions

We consider the Euclidean gravity-dilaton-axion system in $D \geq 3$ dimensions given by the Lagrangian (with arbitrary dilaton coupling parameter b)

$$\mathcal{L}_E = \tfrac{1}{2}\sqrt{g}\,[R - \tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}\,e^{b\phi}(\partial\chi)^2]\,, \tag{3.18}$$

and search for generalized D-instanton solutions with manifest SO(D) symmetry of the form[3]

$$ds^2 = e^{2\,B(r)}(dr^2 + r^2 d\Omega_{D-1}^2)\,, \qquad \phi = \phi(r)\,, \qquad \chi = \chi(r)\,. \tag{3.20}$$

[3]Note that by using reparameterizations of r one can obtain different, but equivalent, forms of the metric in which the SO(D) symmetry is non-manifest, in particular

$$ds^2 = e^{2\,B(r)}(e^{-2\,f(r)}dr^2 + r^2 d\Omega_{D-1}^2)\,, \tag{3.19}$$

in analogy to what we will encounter later, see (3.78). We choose to take as our starting point a conformally flat metric, i.e. $f(r) = 0$.

The standard D-instanton solution [30] is obtained for the special case where $B(r)$ is constant. In order to obtain an SO(D) symmetric generalized D-instanton solution, we allow for a non-constant $B(r)$ and solve the field equations following from the Euclidean action (3.18), which read

$$\begin{aligned} R_{\mu\nu} &= \frac{1}{2}\,\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\,e^{b\phi}\partial_\mu\chi\partial_\nu\chi\,, \\ 0 &= \partial_\mu\left(\sqrt{g}g^{\mu\nu}e^{b\phi}\partial_\nu\chi\right)\,, \\ 0 &= \frac{b}{2}\,e^{b\phi}(\partial\chi)^2 + \frac{1}{\sqrt{g}}\,\partial_\mu\left(\sqrt{g}g^{\mu\nu}\partial_\nu\phi\right)\,. \end{aligned} \tag{3.21}$$

The expression for the Ricci tensor for the Ansatz (3.20) is given by

$$\begin{aligned} R_{rr} &= -(D-1)\left(B''(r) + \frac{B'(r)}{r}\right)\,, \\ R_{\theta\theta} &= -e^{-2\,B(r)}\,g_{\theta\theta}\,[B''(r) + (D-2)\,B'(r)^2 + (2\,D-3)\,\frac{B'(r)}{r}]\,, \end{aligned} \tag{3.22}$$

where the prime denotes differentiation with respect to r, and θ stands for all angular coordinates. In addition to the SL(2, $\mathbb{R}$) symmetry these field equations are invariant under a constant Weyl rescaling of the metric[4]

$$g_{\mu\nu} \rightarrow e^{2\omega}g_{\mu\nu}\,. \tag{3.23}$$

However, this is only a symmetry of the field equations and not of the action. In our Ansatz (3.20), this has the effect of shifting B by a constant, i.e. $B \rightarrow B + \omega$.

In order to solve for $B(r)$, one can consider the angular component of the Einstein equation of (3.21). Having solved for $B(r)$ the expressions for the dilaton and axion scalars can be obtained from the remaining two equations of (3.21). We thus obtain the following solution[5] for $B(r)$, $\phi(r)$ and $\chi(r)$, which extends the solution given in [37] to arbitrary b:

$$\begin{aligned} e^{(D-2)\,B(r)} &= f_+(r)\,f_-(r)\,, \\ e^{b\phi(r)} &= \left(\frac{q_-}{2\,q}\,[e^{C_1}\,(f_+(r)/f_-(r))^{b\,c/2} - e^{-C_1}\,(f_+(r)/f_-(r))^{-b\,c/2}]\right)^2\,, \\ \chi(r) &= \frac{2}{b\,q_-}\,[q\left(\frac{e^{C_1}\,(f_+(r)/f_-(r))^{b\,c/2} + e^{-C_1}\,(f_+(r)/f_-(r))^{-b\,c/2}}{e^{C_1}\,(f_+(r)/f_-(r))^{b\,c/2} - e^{-C_1}\,(f_+(r)/f_-(r))^{-b\,c/2}}\right) - q_3]\,. \end{aligned} \tag{3.24}$$

The solution is given in terms of the two flat-space harmonic functions

$$f_\pm(r) = 1 \pm \frac{q}{r^{D-2}} \tag{3.25}$$

[4]The constant Weyl rescaling symmetry is broken by $O(\alpha')$ corrections.

[5]For practical purposes we omit an overall ± sign corresponding to the $\mathbb{Z}_2$ symmetry of the axion, which defines the difference between between the instanton and anti-instanton. This sign affects some signs in the SL(2, $\mathbb{R}$) charges of the solution, but does not change its conjugacy class.

and the four integration constants q, q_3, q_- and C_1. The integration constant q is defined as the square root of q^2, which is an integration constant that can be positive, zero or negative[6]. Finally, the constant c is given by

$$c = \sqrt{\frac{2(D-1)}{(D-2)}}\,. \tag{3.26}$$

Note that the metric, specified by $B(r)$ given in (3.24), only depends on the product of f_+ and f_-, whereas the scalars only depend on the quotient of f_+ and f_-. This reflects the presence of the scale symmetry (3.23), whose effect is to scale both $f_\pm$ with the same factor. The constants q^2 and q_- occur with inverse powers and have been taken non-zero in the above solution. Below, we will see that sending them to zero yields interesting limits.

The solution (3.24) carries electric $\mathrm{SL}(2,\mathbb{R})$ charges given by

$$Q_E = \begin{pmatrix} q_3 & q_+ \\ -q_- & -q_3 \end{pmatrix}, \tag{3.27}$$

where we have defined the dependent integration constant q_+ via

$$q^2 = -q_+ q_- + {q_3}^2 = -\det(Q_E)\,. \tag{3.28}$$

Thus, the solution (3.24) has general $\mathrm{SL}(2,\mathbb{R})$ charges (q_+, q_-, q_3).

The appearance of the four independent integration constants, q^2, q_-, q_3 and C_1, can be understood as follows. As can be inferred from the solution (3.24), the constant q_3 corresponds to the freedom to apply $\mathbb{R}$ transformations, which shift the axion. Similarly, the constant q_- corresponds to $\mathrm{SO}(1,1)$ transformations, which scale the axion and shift the dilaton. By applying such transformations one can shift q_3 with arbitrary numbers while q_- can be rescaled with a positive number. The constant C_1 is shifted as follows

$$C_1 \to C_1 - 2\,\lambda\, q \tag{3.29}$$

under the $\mathrm{SL}(2,\mathbb{R})$ transformation, with parameter λ, whose generator is given by the electric charge matrix:

$$\Omega_E = \exp(\lambda\, Q_E)\,. \tag{3.30}$$

Since Q_E is invariant under such transformations (see (3.17)), while C_1 is shifted, this explains why C_1 does not appear in (3.27). The remaining constant, q^2, is invariant under $\mathrm{SL}(2,\mathbb{R})$ and hence does not correspond to these symmetry transformations. Rather, this constant corresponds to the freedom to perform rescalings of the metric (3.23). To retain a metric that asymptotically goes to 1, this must be combined with an appropriate rescaling of r. The resulting effect of this transformation is a rescaling of q^2 with a positive number. One therefore always stays in the same conjugacy class under such transformations.

The solution (3.24) can be written in a more compact form by using, instead of the two functions f_+ and f_- which are harmonic over D-dimensional flat space, a function $H(r)$ which is

[6]Note that this implies that the solution (3.24) is not manifestly real, since q can be imaginary. Below, we discuss this issue separately for the three cases q^2 positive, negative or zero.

harmonic over a conformally flat space with the conformal factor specified by the function $B(r)$ given in (3.24), i.e.

$$\Box H(r) = \frac{1}{\sqrt{g}}\frac{\partial}{\partial r}\left(r^{D-1}e^{(D-2)\,B(r)}\frac{\partial H(r)}{\partial r}\right) = 0\,. \tag{3.31}$$

The general solution to this equation is of the following form:

$$H(r) \propto \log(f_+(r)/f_-(r))\,. \tag{3.32}$$

We can, therefore, rewrite the solutions (3.24) as follows:

$$\boxed{\begin{aligned} ds^2 &= \left(1-\frac{q^2}{r^{2\,(D-2)}}\right)^{2/(D-2)}(dr^2+r^2d\Omega^2_{D-1})\,,\\ e^{b\,\phi(r)} &= \left(\frac{q_-}{q}\,\sinh(H(r)+C_1)\right)^2\,,\\ \chi(r) &= \frac{2}{b\,q_-}\,(q\,\coth(H(r)+C_1)-q_3)\,,\end{aligned}} \tag{3.33}$$

where

$$H(r) = \frac{b\,c}{2}\,\log(f_+(r)/f_-(r))\,. \tag{3.34}$$

The solutions (3.33) are valid both for q^2 positive, negative and zero. Below, we will discuss the reality and validity of the solutions for each of these three cases. Note that we are using the Einstein frame.

- $\mathbf{q^2 > 0}$:

  In this case q is real and the solution is given by (3.33) with all constants real. However, the metric poses a problem: it becomes imaginary for

  $$r^{D-2} < r_c^{D-2} = q\,. \tag{3.35}$$

  One can check that there is a curvature singularity at $r = r_c$. However, this curvature singularity happens at strong string coupling:

  $$e^{\phi(r)} \to \infty\,, \qquad\qquad r \to r_c\,. \tag{3.36}$$

  Between $r = r_c$ and $r = \infty$, H varies between ∞ and 0, and with an appropriate choice[7] of C_1, i.e. a positive value of C_1, the scalars have no further singularities in this domain. One might hope to have a modification of this solution by higher-order contributions to the effective action of IIB string theory [38]. Alternatively, one can consider the possible

[7] According to (3.29), the constant C_1 can be changed by an $SL(2,\mathbb{R})$ transformation, leading to singular scalars (but non-singular currents, which are independent of C_1). However, since these are related to regular scalars by a global $SL(2,\mathbb{R})$ transformation, this does not pose a problem.

resolution of this singularity upon uplifting. In the next section, we will see that this indeed happens for the special case of

$$b = \sqrt{\frac{2(D-2)}{D-1}}\,, \tag{3.37}$$

equivalent to $bc = 2$.

In the case with $q^2 > 0$, there is an interesting limit in which $q_- \to 0$. For generical values of the other three constants, this yields a non-sensible solution with infinite scalars. To avoid this, one must simultaneously impose

$$C_1 \to -\log(\frac{q_-}{2q})\,, \qquad q_3 \to q - \frac{q_+ q_-}{2q}\,, \qquad q_- \to 0\,. \tag{3.38}$$

This yields a well-defined limit, in which the scalars read

$$e^{\phi/c} = \frac{f_+}{f_-}\,, \qquad \chi = \frac{-q_+}{bq}\,, \tag{3.39}$$

while the metric is unaffected and given by (3.24). This solution can also be deduced by simply solving the equations of motion from scratch, with the constant axion Ansatz. Note that in this limit the dilaton becomes independent of b: when the axion is constant, the dilaton coupling drops out of the field equations. In this limit, one is left with two independent integration constants, q_+ and q^2. The range of validity of this solution is equal to that of the above solution with $q_- \neq 0$: it is well-defined for $r > r_c$, while at $r = r_c$ the metric has a singularity and the dilaton blows up. We will find that this singularity is resolved upon uplifting for all values of $bc \geq 2$.

- $\mathbf{q^2 = 0}$

  We now consider the limit $q^2 \to 0$ of the general solution (3.33). Taking this limit for generic values of C_1, one sees that $e^{\phi(r)} \to \infty$ for all r. The only way to avoid this bad behaviour is to have $C_1 \to 0$, as $q^2 \to 0$. Thus, to obtain a well-defined limit, we simultaneously take

$$C_1 \to g_s^{b/2}\,\frac{q}{q_-}\,, \qquad q^2 \to 0\,. \tag{3.40}$$

The constant g_s is assumed positive and will correspond to the value of $e^{\phi(r)}$ at $r = \infty$. Taking the limit (3.40) of the general solution (3.33) yields the extremal solution:

$$\boxed{ds^2 = dr^2 + r^2\,d\Omega^2_{D-1}\,, \qquad e^{b\,\phi(r)/2} = h \qquad \chi(r) = \frac{2}{b}\,(h^{-1} - \frac{q_3}{q_-})\,,} \tag{3.41}$$

where $h(r)$ is the harmonic function:

$$h(r) = g_s^{b/2} + \frac{b\,c\,q_-}{r^{D-2}}\,. \tag{3.42}$$

This is the extremal D-instanton solution of [30]. It can also be obtained by solving the equations from scratch with a flat metric in the Ansatz. This solution is regular over the range $0 < r < \infty$ provided one takes both g_s and $b\,c\,q_-$ positive; at $r = 0$ however, the harmonic function blows up and the scalars are singular. Again, string theory corrections may resolve these scalar singularities.

- $\mathbf{q^2 < 0}$:

In this case q is imaginary. To obtain a real solution we must take C_1 to be imaginary. We therefore redefine

$$q \to i\,\tilde{q} \qquad\qquad C_1 \to i\,\tilde{C}_1\,, \tag{3.43}$$

such that $\tilde{q}$ and $\tilde{C}_1$ are real. One can now rewrite the solution (3.33) by using the relation[8]

$$\log(f_+/f_-) = 2\,\mathrm{arctanh}(q/r^{D-2})\,, \tag{3.44}$$

and, next, replacing the hyperbolic trigonometric functions by trigonometric ones in such a way that no imaginary quantities appear. We find that, for $q^2 < 0$, the general solution (3.33) takes the following form:

$$\boxed{\begin{aligned} ds^2 &= (1 + \frac{\tilde{q}^2}{r^{2(D-2)}})^{2/(D-2)}\,(dr^2 + r^2\,d\Omega^2_{D-1})\,,\\ e^{b\phi(r)} &= \left(\frac{q_-}{\tilde{q}}\ \sin(b\,c\ \arctan(\frac{\tilde{q}}{r^{D-2}}) + \tilde{C}_1)\right)^2\,,\\ \chi(r) &= \frac{2}{b\,q_-}\,(\tilde{q}\ \cot(b\,c\ \arctan(\frac{\tilde{q}}{r^{D-2}}) + \tilde{C}_1) - q_3)\,. \end{aligned}} \tag{3.45}$$

The metric and curvature are well behaved over the range $0 < r < \infty$. However, the scalars can only be non-singular over the same range by an appropriate choice of $\tilde{C}_1$ provided that $bc < 2$. This can be seen as follows: the arctan varies over a range of $\pi/2$ when r goes from 0 to ∞. Since it is multiplied by bc, the argument of the sine varies over a range of more than π if $bc > 2$. Therefore, for $bc > 2$ there is always a point r_c such that $\chi \to \infty$ as $r \to r_c$. Note that the breakdown of the solution occurs at weak string coupling: $e^\phi \to 0$ as $r \to r_c$. In the next section we will find that this singularity is not resolved upon uplifting and will correspond to a black hole with a naked singularity. The same holds for the limiting case of $bc = 2$. Therefore the case $q^2 < 0$ only yields regular instanton solutions for $bc < 2$, together with the condition that C_1 and $C_1 + bc\pi/2$ are on the same branch of the cotangent.

3.3.2 Wormhole geometries

It is known [30] that the standard D-instanton, i.e. $D = 10, b = 2$, in string frame has the geometry of a wormhole, i.e. it has two asymptotically flat regions connected by a neck, see figure 3.1. It will therefore be interesting to investigate whether there exist frames, in which the non-extremal instantons also have the geometries of wormholes.

[8] Here we have used the general relation $\log((1 + x)/(1 - x)) = 2\,\mathrm{arctanh}(x)$.

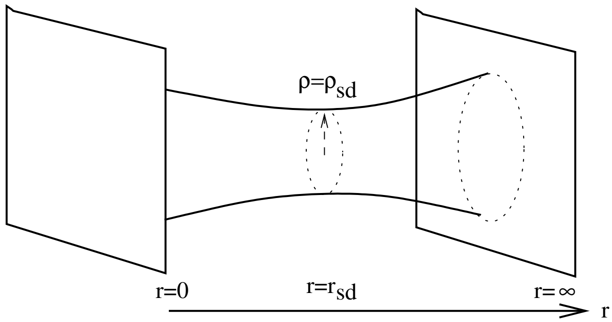


Figure 3.1: *The geometry of a wormhole. The two asymptotically flat regions at $r = 0$ and $r = \infty$ are connected via a neck with a minimal physical radius ρ_{sd} at the self-dual radius r_{sd}.*

We consider a general wormhole metric of the form

$$ds^2 = f(r)^{2/(D-2)}\,(dr^2 + r^2 d\Omega^2)\,, \qquad f(r) = \alpha + \beta r^{2-D} + \gamma r^{4-2D}\,, \tag{3.46}$$

where α, β and γ are constants. The metric has a $\mathbb{Z}_2$ isometry corresponding to the transformation $r^{D-2} \to \gamma\, r^{2-D}/\alpha$ which interchanges the two asymptotically flat regions. The physical radius ρ is the square root of the coefficient of the angular part of the metric, given by $\rho^{D-2} = f(r) r^{D-2}$. The minimum of this physical radius of the neck occurs at the fixed point of the transformation above, i.e. at the so-called self-dual radius $r_{\rm sd}^{D-2} = \sqrt{\gamma/\alpha}$, and is given by $\rho_{\rm sd}^{D-2} = 2\sqrt{\alpha\gamma} + \beta$. We will now study the three conjugacy classes in order to see for each case if there exists a frame[9] in which the metric takes the form (3.46).

- $\mathbf{q^2 > 0}$: As we will see in section 3.4, the appropriate frame in this case is the frame dual to the instanton, i.e. the $(D-3)$-brane frame, given by

  $$g^{\rm dual}_{\mu\nu} = e^{b\phi/(D-2)}\, g^{\rm E}_{\mu\nu}\,. \tag{3.47}$$

  In the special case of $b\,c = 2$, the metric takes the form (3.46) in the dual frame with

  $$f(r) = \frac{q_-}{q}\sinh(C_1) + 2q_-\cosh(C_1) r^{2-D} + q_- q \sinh(C_1) r^{4-2D}\,. \tag{3.48}$$

  This gives the self-dual radius $r_{\rm sd}$ and the minimal physical radius $\rho_{\rm sd}$

  $$r_{\rm sd}^{D-2} = q\,, \qquad \rho_{\rm sd}^{D-2} = 2q_- e^{C_1}\,. \tag{3.49}$$

  Note that the self-dual radius $r_{\rm sd}$ coincides with the critical radius r_c of the previous section: the curvature singularity in Einstein frame becomes the center of the wormhole in

[9] In arbitrary dimension one can define three different frames as follows: in the Einstein frame, the Einstein-Hilbert term has no dilaton factor; in the string frame, the kinetic term for the axionic field strength comes without a dilaton factor (like all Ramond-Ramond field strengths); and in the dual frame, the Einstein-Hilbert term, the dilaton kinetic term and the kinetic term for the dual field strength (i.e. F^2_{D-p-2} for the frame dual to a p-brane) come with the same dilaton factor (see e.g. [49, 50] for a more detailed discussion).

the dual frame. The limit $q_- \to 0$, with appropriate scaling of C_1 as given in (3.38), yields $\rho_{\text{sd}}^{D-2} = 4q$. For generic values of bc, the instanton metrics cannot be written in the form (3.46) in any frame.

- $\mathbf{q^2 = 0}$: It turns out that for any value of b the wormhole geometry is made manifest by going to the string frame

$$g_{\mu\nu}^{\text{str}} = e^{2b\phi/(D-2)}\, g_{\mu\nu}^{\text{E}} \,. \tag{3.50}$$

In this frame, the metric is given by (3.46) with

$$f(r) = g_s^b + 2bcq_- g_s^{b/2} r^{2-D} + (bcq_-)^2 r^{4-2D} \,. \tag{3.51}$$

This gives the self-dual and minimal physical radii

$$r_{\text{sd}}^{D-2} = bcq_-/g_s^{b/2} \,, \qquad \rho_{\text{sd}}^{D-2} = 4bcq_- g_s^{b/2} \,. \tag{3.52}$$

- $\mathbf{q^2 < 0}$: Here, the metric has the appropriate form already in Einstein frame, hence, from (3.45) we get, for any value of b,

$$r_{\text{sd}}^{D-2} = \tilde{q} \,, \qquad \rho_{\text{sd}}^{D-2} = 2\tilde{q} \,. \tag{3.53}$$

We thus see that for all three conjugacy classes there exists frames, in which the solutions have the geometries of wormholes.

3.3.3 Instanton solutions with multiple dilatons

We will now consider extensions of the instanton solution described in the previous sections, which is carried by the $\mathrm{SL}(2,\mathbb{R})$ scalars ϕ and χ. We will extend this system with n dilatons φ_α ($\alpha = 1, \ldots, n$), which are $\mathrm{SL}(2,\mathbb{R})$ singlets and do not couple to the axion (this can always be achieved by field redefinitions provided one allows for an arbitrary dilaton coupling b to the original dilaton ϕ). We will call the corresponding solution a multi-dilaton instanton. The multi-dilaton action is given by

$$\mathcal{L}_E = \tfrac{1}{2}\sqrt{g}\,[R - \tfrac{1}{2}\sum_{\alpha=1}^{n}(\partial\varphi_\alpha)^2 - \tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}e^{b\phi}(\partial\chi)^2] \,, \tag{3.54}$$

with field equations (3.21) plus n equations, requiring φ_α to be harmonic in the curved space. The case of one extra dilaton was considered in [51].

The solution to this system has the same metric as given in (3.24), see also [51]. Then the extra dilatons φ_α satisfy a d'Alembertian equation in a conformally flat background specified by $B(r)$ as given in (3.24):

$$\frac{\partial}{\partial r}\left(r^{D-1} e^{(D-2)\,B(r)} \frac{\partial\varphi(r)}{\partial r}\right) = 0 \,. \tag{3.55}$$

This equation is solved by the harmonic function as given in (3.32), yielding dilatons given by

$$\varphi_\alpha = \nu_\alpha + \mu_\alpha \log\left(\frac{f_+(r)}{f_-(r)}\right), \tag{3.56}$$

with $2n$ integrations constants ν_α and μ_α.

Of course, due to the presence of the extra dilatons φ_α, the Einstein equation in (3.21) is modified. It turns out that the contribution of φ_α to the energy-momentum tensor is cancelled by similar μ_α-dependent contributions of the dilaton ϕ and the axion χ to the energy-momentum tensor. Since all μ_α-dependent contributions of the dilatons and the axion to the energy-momentum tensor cancel each other, this extension allows for a μ_α-independent metric.

3.4 Uplift to black holes

In this section, we will find an explicit example of the soliton-instanton correspondence mentioned in chapter 2. We will show that a D-instanton can sometimes be viewed as a spacelike section of a charged black hole, and more generally a p-brane.

3.4.1 Kaluza-Klein reduction

In this section we consider the possible higher-dimensional origin of the Euclidean system (3.18) as a consistent truncation of the $(D+1)$-dimensional Lagrangian, defined over Minkowski space,

$$\mathcal{L}_{D+1} = \sqrt{-\hat{g}}\,[\hat{R} - \tfrac{1}{2}(\partial\hat{\phi})^2 - \tfrac{1}{4}\,e^{a\hat{\phi}}\,\hat{F}^2]\,, \tag{3.57}$$

with the two-form field strength $\hat{F} = d\hat{A}$. It consists of an Einstein-Hilbert term (for a metric of Lorentzian signature), a dilaton kinetic term and a kinetic term for a vector potential with arbitrary dilaton coupling, parametrized by a. The corresponding Δ value [52] is given by

$$\Delta = a^2 + \frac{2\,(D-2)}{D-1}\,, \tag{3.58}$$

which characterizes the dilaton coupling in $D+1$ dimensions.

The reduction Ansatz over the time coordinate is

$$\hat{ds}^2 = e^{2\alpha\varphi}\,ds^2 - e^{2\beta\varphi}\,dt^2\,, \qquad \hat{A} = \chi\,dt\,, \qquad \hat{\phi} = \phi\,, \tag{3.59}$$

with the constants

$$\alpha^2 = \frac{1}{2\,(D-1)\,(D-2)}\,, \qquad \beta = -(D-2)\,\alpha\,, \tag{3.60}$$

which are chosen such as to obtain the Einstein frame in the lower dimension with appropriate normalization of the dilaton φ. Note that the dilaton factor in front of the spatial part of the metric $\hat{g}_{\mu\nu}$ coincides, for $bc = 2$, with the dual frame defined in section 3.3.2.

With the above Ansatz, the Einstein-Maxwell-dilaton system reduces to the D-dimensional Euclidean system

$$\mathcal{L}_D = \sqrt{-g}\,[R - \tfrac{1}{2}(\partial\phi)^2 - \tfrac{1}{2}(\partial\varphi)^2 + \tfrac{1}{2}\,e^{a\,\phi-2\beta\,\varphi}\,(\partial\chi)^2]\,. \tag{3.61}$$

Next, we perform a field redefinition corresponding to a rotation in the (ϕ, φ)-plane such that we obtain

$$\mathcal{L}_D = \sqrt{-g}\,[R - \tfrac{1}{2}(\partial\tilde{\phi})^2 - \tfrac{1}{2}(\partial\tilde{\varphi})^2 + \tfrac{1}{2}e^{b\tilde{\phi}}(\partial\chi)^2]\,, \tag{3.62}$$

with dilaton coupling b given by

$$b = \sqrt{a^2 + \frac{2(D-2)}{D-1}}\,. \tag{3.63}$$

The corresponding value of Δ is equal to the original value (3.58). This system can be truncated to the one we are considering by setting $\tilde{\varphi} = 0$.

Therefore, the system that we consider in section 3.3 has a higher-dimensional origin if the dilaton coupling satisfies $bc \geq 2$ or

$$b \geq \sqrt{\frac{2(D-2)}{D-1}}\,. \tag{3.64}$$

The case which saturates the inequality, i.e. $a = 0$, can be uplifted to an Einstein-Maxwell system without the dilaton $\hat{\phi}$. For $bc > 2$ one needs to include an explicit dilaton $\hat{\phi}$ in the higher-dimensional system; i.e. one must consider the Einstein-Maxwell-dilaton system (3.57) with $a \neq 0$. Note that in string theory toroidal reductions, under which the combination Δ is preserved, only lead to values of b with $bc \geq 2$.

Since the Euclidean gravity-axion-dilaton system we are considering can be obtained as a consistent truncation of the higher-dimensional Minkowskian Einstein-Maxwell-dilaton system (3.57), it is natural to look for a higher-dimensional origin of the non-extremal instanton solutions within this system. In the following two sections we consider the cases $bc = 2$ and $bc > 2$ separately. The instantons with $bc < 2$ have no physical higher-dimensional origin from toroidal reduction.

3.4.2 Reissner-Nordström black holes: $bc = 2$

It is not difficult to see that for $bc = 2$ the generalized instanton solutions uplift to the $(D + 1)$-dimensional Reissner-Nordström (RN) black hole solution

$$ds^2 = -g_+(\rho)\,g_-(\rho)\,dt^2 + \frac{d\rho^2}{g_+(\rho)\,g_-(\rho)} + \rho^2 d\Omega_{D-1}^2\,, \quad F_{t\rho} = -\partial_\rho A_t = (D-2)\,c\,\frac{Q}{\rho^{D-1}}\,, \tag{3.65}$$

where

$$g_\pm(\rho) = 1 - \frac{\rho_\pm^{D-2}}{\rho^{D-2}}\,, \qquad \rho_\pm^{D-2} = M \pm \sqrt{M^2 - Q^2}\,, \tag{3.66}$$

and Q and M are the charge and mass of the black hole, respectively. The RN black hole has naked singularities for $M^2 < Q^2$, while these are cloaked for $M^2 \geq Q^2$, yielding a physically acceptable spacetime. Note that the coordinate ρ coincides with the physical radius of the previous section, for which the angular part of the metric $d\Omega_{D-1}^2$ is multiplied by ρ^2.

In order to establish the precise relation between the charge Q and the mass M of the RN black hole and the SL(2, $\mathbb{R}$) charges of the $bc = 2$ instanton solutions given in (3.33) we must first cast the RN metric in isotropic form as follows:

$$ds^2 = -\frac{g(r)}{\rho(r)^{2(D-2)}}\, dt^2 + \frac{\rho(r)^2}{r^2}\,(dr^2 + r^2\, d\Omega^2_{D-1})\,, \tag{3.67}$$

where

$$\rho(r) = \left(r^{D-2} + M + \frac{M^2 - Q^2}{4\, r^{D-2}}\right)^{1/(D-2)}, \qquad g(r) = \left(r^{D-2} - \frac{M^2 - Q^2}{4\, r^{D-2}}\right)^2 . \tag{3.68}$$

To relate the instanton and black hole solutions, we need to choose proper boundary conditions for the instanton solutions (3.33), which are implied by the boundary conditions of the RN black hole:

$$\begin{array}{lcl} \lim_{r\to\infty} g_{tt} = -1\,, & \Longleftrightarrow & \lim_{r\to\infty} e^{\phi} = 1\,, \\ \lim_{r\to\infty} A_t = 0\,, & & \lim_{r\to\infty} \chi = 0\,. \end{array} \tag{3.69}$$

This fixes the constants C_1 and one of the three SL(2, $\mathbb{R}$) charges q_3 in (3.33) as follows:

$$C_1 = \text{arcsinh}(\frac{q}{q_-})\,, \qquad q_3 = q\, \coth(C_1) = \sqrt{q^2 + q_-^2}\,. \tag{3.70}$$

The relation between the charge Q and the mass M of the RN black hole and the two unfixed SL(2, $\mathbb{R}$) charges q_- and q^2 is:

$$Q = -2\, q_-\,, \qquad M = 2\, \sqrt{q^2 + q_-^2}\,, \tag{3.71}$$

such that

$$q^2 = \frac{M^2 - Q^2}{4}\,. \tag{3.72}$$

From (3.72) we see that the physically acceptable non-extremal RN black holes with $M^2 \geq Q^2$ coincide with the uplifted instanton solutions in the $q^2 = 0$ and $q^2 > 0$ conjugacy classes:

$$\begin{array}{lcl} M^2 > Q^2 & \Leftrightarrow & q^2 > 0\,, \\ M^2 = Q^2 & \Leftrightarrow & q^2 = 0\,. \end{array} \tag{3.73}$$

More specifically, we find that the non-extremal (extremal) RN metric in isotropic coordinates (3.67) reduces to the $q^2 > 0$ ($q^2 = 0$) instanton solution in the dual frame metric (3.47). Note that the $q^2 > 0$ instanton has a wormhole geometry in the dual frame metric. It turns out that the minimal physical radius ρ_{sd} for this case is given by $\rho_{\text{sd}} = \rho_+$, where ρ_+ is the position of the outer event horizon given in (3.66).

3.4.3 Interpretation of instantons as BH wormholes

In the previous section we have seen that the non-extremal D-instanton solutions (3.33) in the dual frame metric (3.47) with $b\,c = 2$ and $M^2 \geq Q^2$ can be viewed as $t = constant$ space-like sections of the RN black hole metric (3.67). In the Kruskal-Szekeres-like extension of the RN black hole, the spatial part of the metric (3.67) has the geometry of an Einstein-Rosen bridge or wormhole, which connects two asymptotically flat regions of space (see [20] for a general introduction to black holes). Indeed, the spatial part of (3.67) has, for $M^2 > Q^2$, the $\mathbb{Z}_2$ isometry

$$r^{D-2} \rightarrow \frac{M^2 - Q^2}{4\,r^{D-2}}\,, \tag{3.74}$$

which relates each point on one side of the Einstein-Rosen bridge to a point on the other side.

It is instructive to consider the special case of the Schwarzschild black hole, (i.e. $Q = 0$). Due to (3.71), this corresponds to the uplift of instantons with $q_- = 0$, i.e. the solutions given in (3.38). As shown in figure 3.2, in the Kruskal-Szekeres extension of the Schwarzschild black hole, every $t = constant$ section of space time corresponds to a straight space-like line going through the origin of this coordinate system, with slope determined by the constant value of t.

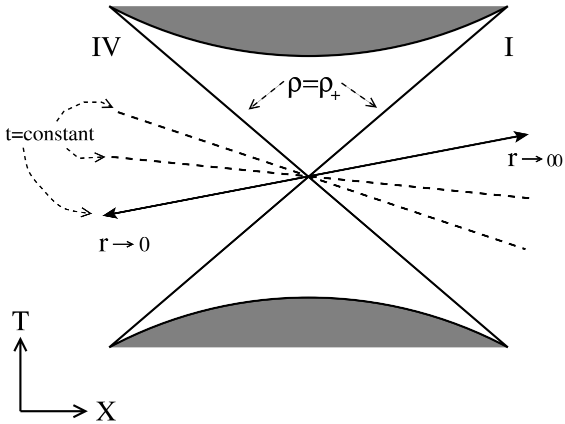


Figure 3.2: *Schwarzschild black hole in Kruskal-Szekeres coordinates. Spatial sections with $t = constant$ are space-like lines through the origin, going from region IV to region I. T and X are the Kruskal-Szekeres time-like and space-like directions respectively. The horizons are at $\rho = \rho_+$, which coincides with the minimal physical radius at the center $\rho = \rho_{\rm sd}$.*

Notice that on each line, the coordinate r from (3.67) runs from $r = 0$ at the spatial infinity on the left-hand-side, to $r = \infty$ on the right-hand-side. The fixed point of the $\mathbb{Z}_2$-isometry (3.74) (now with $Q = 0$) is positioned at the center of figure 3.2. The value of r at this fixed point and the corresponding minimal physical radius are given by

$$r_{\rm sd}^{D-2} = \tfrac{1}{2}M\,, \qquad\qquad \rho_{sd}^{D-2} = 2M\,. \tag{3.75}$$

Note that this value of the physical radius corresponds to the horizon of the black hole, as can also be seen from figure 3.2. One can make the wormhole geometry visible by associating to

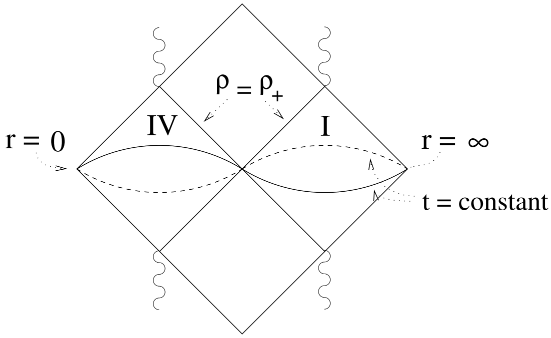


Figure 3.3: *Carter-Penrose diagram of RN black hole. The lines with $\rho = \rho_+$ are the horizons, which coincide with the minimal physical radius $\rho = \rho_{sd}$ in the center.*

every value of r a $(D-1)-$sphere. Representing every $(D-1)-$sphere by a circle one obtains the wormhole picture of figure 3.1.

In the more general case (i.e. $Q \neq 0$), the $t = constant$ sections are still paths connecting two regions of the RN black hole. To see what these regions correspond to, it is helpful to draw a Carter-Penrose diagram, see figure 3.3. The wormhole geometry is qualitatively the same as in the Schwarzschild case. The position of the wormhole neck and the value of the minimal physical radius are given by

$$r_{\rm sd}^{D-2} = \tfrac{1}{4}(M^2 - Q^2)\,, \qquad\qquad \rho_{sd}^{D-2} = M + \sqrt{M^2 - Q^2}\,, \tag{3.76}$$

which again coincide with the horizon at $\rho = \rho_+$. The curvature singularity of the D-instanton solutions with $q^2 > 0$ (3.33) at $r_{\rm c} = (q)^{1/D-2}$ are resolved in this uplifting and can now be understood as the usual coordinate singularity of the RN black hole outer event horizons (i.e. $\rho = \rho_+$, or $r^{2\,(D-2)} = (M^2 - Q^2)/4$).

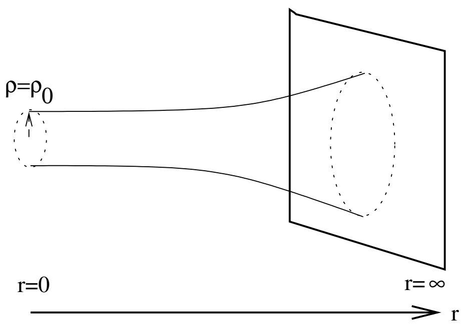


Figure 3.4: *The geometry of the extremal black hole as a "one-sided" wormhole with minimal physical radius ρ_0.*

The extremal RN black hole (i.e. $M^2 = Q^2$) is qualitatively different from the other cases. As one can see from (3.74), the $\mathbb{Z}_2$-isometry is gone. By taking the limit $M^2 \to Q^2$ of a non-extremal black hole we see that the wormhole stretches to an infinitely long neck. The fixed point of the isometry goes to spatial infinity at $r = 0$. This means that the extremal black hole has a "one-sided" wormhole with a minimal physical radius $\rho_0^{D-2} = M$, and the full Kruskal-like extension is geodesically complete without need for a region IV. This situation is illustrated in figure 3.4.

3.4.4 Dilatonic black holes: $bc > 2$

The instantons with $bc > 2$ uplift to non-extremal *dilatonic* black holes, i.e. black hole solutions carried by a metric, a vector and a dilaton. In fact, the uplift is identical to a version of the black hole solution presented in [53]. To be more precise, the non-extremal dilatonic black hole solutions of [53] contain an extra parameter μ. For generic values of this parameter the black hole solution is singular[10]. One only obtains a regular solution if[11] $\mu \sim q$.

The uplift of the $bc > 2$ instantons equals the $\mu \to 0$ limit of the non-extremal black hole solutions of [53]. Therefore, in contrast to the $bc = 2$ case, we obtain a singular black hole solution. This singularity can only be avoided in two limiting cases. The singularity disappears both in the extremal limit (3.40) when $q^2 \to 0$ and in the Schwarzschild limit (3.39) when $q_- \to 0$, where the dilaton decouples.

3.5 Uplift to p-branes

In section 4 we have discussed the uplift of the instantons of section 3 to higher-dimensional black hole solutions. It is therefore natural to consider the uplift to higher-dimensional p-branes. To this end, it will be useful to first introduce the following nomenclature.

Non-extremal deformations of general p-branes have been considered in [53, 55]. These are solutions of the $(D + p + 1)$-dimensional Lagrangian, defined over Minkowski space,

$$\mathcal{L}_{D+p+1} = \sqrt{-\hat{g}}\,[\hat{R} - \frac{1}{2}(\partial\hat{\phi})^2 - \frac{1}{2\,(p+1)!}\, e^{a\hat{\phi}}\, \hat{G}^2_{(p+2)}]\,, \qquad (3.77)$$

with the rank-$(p + 2)$ field strength $\hat{G}_{(p+2)} = d\hat{C}_{(p+1)}$. For a p-brane in $D + p + 1$ dimensions the metric (in Einstein frame) is of the form

$$ds^2 = e^{2A}(-e^{2f}dt^2 + dx_p{}^2) + e^{2B}(e^{-2f}dr^2 + r^2 d\Omega_{D-1}{}^2)\,, \qquad (3.78)$$

where A, B and f are functions that depend on the radial coordinate r only. It is convenient to introduce the quantity

$$X = (p + 1)A + (D - 3)B\,. \qquad (3.79)$$

The extremal p-brane solutions with equal mass and charge, preserving half of the supersymmetry, are obtained by taking $X = f = 0$.

Assuming that $D \geq 3$ there exist two types of non-extremal p-brane solutions in the literature. Following [53], we will call them type 1 and type 2 non-extremal p-branes:

[10] These (singular) solutions are a generalization of the (regular) black holes of [54].
[11] The parameter q^2 can be identified with the parameter k of [53].

- **Type 1 non-extremal p-branes:** $X = 0$ and $f \neq 0$.

  These are the non-extremal black branes of [55, 56]. The deformation function f is given by

$$e^{2f} = 1 - \frac{k}{r^{D-2}} \,, \tag{3.80}$$

  where k is the deformation parameter. In a different coordinate frame, with radial coordinates ρ, these branes can be expressed in terms of the two harmonic functions

$$f_{\pm}(\rho) = 1 - \left(\frac{\rho_{\pm}}{\rho}\right)^{D-2} \,. \tag{3.81}$$

  Physical branes without a naked singularity have more mass than charge, which corresponds to $\rho_+ > \rho_-$ or $k > 0$. For this type of non-extremal deformation, the dilaton $\hat{\phi}$ is proportional to A and B, which are linearly related since $X = 0$.

- **Type 2 non-extremal p-branes:** $X \neq 0$ and $f = 0$.

  These are the non-extremal black branes of [53]. The deformation function X reads

$$e^{X} = 1 - \frac{k}{r^{2(D-2)}} \,, \tag{3.82}$$

  where k is the deformation parameter. The absence of naked singularities requires k to be positive. In this case, the dilaton $\hat{\phi}$ is not proportional to A or B, which are not linearly related.

  The non-extremal D-instanton solutions (3.33) fit exactly in this chain of non-extremal p-branes for $p = -1$. Although the type 2 non-extremal p-branes are defined in Minkowski space, we find that one can extend the formulae of [53] to $p = -1$ branes in Euclidean space, i.e. generalized D-instantons, by taking $f = 0$ and $B \neq 0$.

Both types of non-extremal p-branes break supersymmetry. A special case is $p = 0$, for which the regular type 1 and type 2 non-extremal 0-branes are equivalent up to a coordinate transformation in r. From the form of the metric (3.78), which has different world-volume isometries for $f = 0$ and $f \neq 0$, it is clear that this is not the case for $p > 0$.

To relate the (multi-dilaton) instanton solutions of section 3 to the non-extremal p-branes, it is instructive to reduce the p-branes over their $(p+1)$-dimensional world-volume, including time. In complete analogy with the reduction over time of section 4.1, this will give rise to $p+1$ dilatons from the world-volume of the p-brane. However, these are not all unrelated: for one thing, the dilatons corresponding to the spatial world-volume will be proportional to each other, and can therefore be truncated to a single dilaton. We will denote the dilaton from the spatial metric components by φ, while the time-like component of the metric gives rise to $\tilde{\varphi}$. In general, the reduction of non-extremal p-branes will therefore give rise to a multi-instanton solution with three different dilatons, including the explicit dilaton ϕ:

$$\hat{g}_{tt} \to \tilde{\varphi} \,, \qquad \hat{g}_{xx} \to \varphi \,, \qquad \hat{\phi} \to \phi \,. \tag{3.83}$$

For the two types of non-extremal deformations considered here, however, there is always a relation between the three dilatons, allowing a truncation to two dilatons[12]. For the type 1 deformations the dilatons ϕ and φ are related, as can be seen from the metric with $X = 0$. Similarly, the type 2 deformations yield a relation between φ and $\tilde{\varphi}$ since $f = 0$. Therefore, these non-extremal p-branes reduce to multi-dilaton instanton solutions with two inequivalent dilatons. Conversely, two-dilaton instanton solutions can uplift to either types of non-extremal p-branes, by embedding these dilatons in different ways in the higher-dimensional metric and dilaton.

It is interesting to investigate when these two dilatons can be related or reduced to one, therefore corresponding to our explicit SL$(2,\mathbb{R})$ instanton solution (3.24) with only one dilaton. For the type 1 deformations, this is only possible for the special case with $p = 0$ and $a = 0$. For these values, the dilatons ϕ and φ vanish, leaving one with only $\tilde{\varphi}$. The constraint on a implies $bc = 2$ which, as discussed in section 3, gives rise to the Reissner-Nordström black hole.

For the type 2 deformations there are more possibilities to eliminate the dilaton ϕ. It can be achieved by requiring $a = 0$, as we did for the uplift to black holes. For general p, this leads to the following constraint on b:

$$b = \sqrt{\frac{2(p+1)(D-2)}{D+p-1}}\,. \tag{3.84}$$

Note that this yields $bc = 2$ for black holes with $p = 0$. For these values of b, the instanton solution (3.24) can be uplifted to regular non-extremal non-dilatonic p-branes. For higher values of b, the instanton solution uplifts to singular non-extremal dilatonic p-branes. For these solutions to become regular, one must take either $q^2 \to 0$ or $q_- \to 0$, exactly like we found in the $bc > 2$ discussion of section 4.3.

The uplift of the SL$(2,\mathbb{R})$ instanton solution (3.24) to p-branes is therefore very similar to the uplift to black holes. There is one value of b (3.84) for which the instanton solution can be uplifted to a regular non-extremal non-dilatonic p-brane of type 2. For higher values of b one can obtain singular non-extremal dilatonic p-branes of type 2, which only become regular on either of the limits $q^2 \to 0$ and $q_- \to 0$. By adding an extra dilaton to the instanton solution one can also make a connection to the regular type 1 and type 2 non-extremal dilatonic p-branes.

3.6 Instantons

In the previous section we focused on the bulk behavior of the three conjugacy classes of instanton-like solutions. In this section we will investigate which of these solutions can be interpreted as instantons. Instantons, as we have seen in chapter 2 are defined to be solutions of the Euclidean equations of motion with finite, non-zero value of the action. They have a tunneling interpretation, and generically contribute to certain correlation functions in the path integral with terms that are exponentially suppressed by the instanton action. These correlation functions then induce new interactions in the effective action, and for the extremal, 1/2 BPS, D-instantons in type IIB in $D = 10$, these effects are captured by certain SL$(2,\mathbb{Z})$ modular functions that

[12]This seems to indicate a generalization of the non-extremal deformations with both $X \neq 0$ and $f \neq 0$, reducing to a three-dilaton instanton.

multiply higher derivative terms such as R^4 and their superpartners [31]. Before we study correlation functions and effective interactions induced by non-extremal D-instantons, we must first discuss the properties and show the finiteness of the non-extremal instanton action. We will do this using a method that will allow us to recover the special case of extremal D-instantons easily.

3.6.1 Instanton action

The first thing we notice is that the action (3.18), evaluated on *any* solution of (3.21) vanishes. What is also bothersome about the Euclidean action (3.18) is that it is not bounded from below, not even in the scalar sector. Such actions cannot be used in a path integral, since fluctuations around the instanton will diverge. However, this should not be a surprise at all. After all, the Lagrangian (3.5), whose equations of motion we have been solving, is not the true Lagrangian of the full quantum field theoretic system, but an *effective* Lagrangian that is only meant to be used for finding 'saddle points'[13]. It was never meant to appear in a path integral. In order to evaluate the true value of the action of the non-extremal D-instanton we will use the dualization procedure and replace this dilaton-axion system with a system containing the dilaton and a $(D-1)$-form field-strength, which *does* have true saddle points. This procedure was briefly mentioned at the beginning of this chapter. We will now fully develop it here. For a toy model illustration of this procedure, see appendix A.

The goal is to prove that two different systems can be regarded as the effective path integrals of one and only one common parent path integral. Let us first write down the Euclidean path integral for a dilaton coupled to a $(D-1)$-form field-strength[14], subject to the constraint of being a closed form, i.e. $dF_{D-1} = 0$:

$$\int d[F_{D-1}]\, d[\lambda]\, \exp\left(\int_{\mathcal{M}} -\tfrac{1}{2}\,(d\phi \wedge *d\phi + e^{-b\phi}\, F_{D-1} \wedge *F_{D-1}) + i\,\lambda\, dF_{D-1}\right), \qquad (3.85)$$

where λ is *real* and acts as a quantum Lagrange multiplier that imposes the constraint $dF_{D-1} = 0$, by means of the following identity:

$$\int d[\lambda]\exp[i\,\lambda\, G] = \delta[G]\,, \qquad \text{for any function} \quad G\,, \qquad (3.86)$$

where the $\delta[]$ stands for δ-function*al*. Notice that we are treating the field-strength as fundamental, not the gauge potential. This path integral is defined with 'Dirichlet' boundary conditions on F_{D-1}, i.e. some of the components of F_{D-1} are fixed on the boundary. The constraint that the former be closed implies that it is locally exact, i.e. locally, $F_{D-1} = dC_{D-2}$, for some C_{D-2}. The path integral (3.85) is well-defined because the action is positive-definite, and it is straightforward to find its saddle points, by treating C_{D-2} as fundamental, and deriving the usual higher-dimensional Maxwell equations.

Let us now change the order of integration and perform the path integral over F_{D-1} first. In order to do this, we need to rewrite the action in such a way that the field-strength appears

[13] They are not the true saddle points of the scalar system, but they still provide a semiclassical approximation of the path integral.

[14] We will not worry about the gravitational sector in the following derivation, since it is not relevant. The integration over the dilaton is also omitted.

without derivatives acting on it:

$$\begin{aligned}\mathcal{S}_E &= \int_{\mathcal{M}} \tfrac{1}{2}\left(d\phi \wedge *d\phi + e^{-b\phi}\, F_{D-1} \wedge *F_{D-1} - 2\, i\, \lambda\, dF_{D-1}\right) \\ &= \int_{\mathcal{M}} \tfrac{1}{2}\left[d\phi \wedge *d\phi + e^{-b\phi}\left(F_{D-1} + i\, e^{b\phi} * d\lambda\right) \wedge *\left(F_{D-1} + i\, e^{b\phi} * d\lambda\right)\right. \\ &\qquad \left. + e^{b\phi}\, d\lambda \wedge *d\lambda - 2\, i\, d\,(\lambda\, F_{D-1})\right],\end{aligned} \tag{3.87}$$

where we have used partial integration and the fact that, in a Euclidean space, $**A_p = (-)^{(D-1)p} A_p$, where A_p is a p-form. The last term in (3.87) is a surface term, and since boundary conditions have been imposed on the field-strength, it will not participate in the path integral over the latter. The term can be interpreted as an external current J^μ. Defining $*J = F_{D-1}$, we have

$$\int_{\mathcal{M}} d\,(\lambda\, F_{D-1}) = \int_{\partial\mathcal{M}} \lambda\, *\, J = \int_{\partial\mathcal{M}} \lambda\, J_\mu\, n^\mu\,, \tag{3.88}$$

where n^μ is an outward normal vector.

To integrate F_{D-1} in (3.87), we first perform the following shift of integration variables:

$$F_{D-1} \to \bar{F}_{D-1} + i\, e^{b\phi} * d\lambda\,. \tag{3.89}$$

We are allowed to do this even though λ is real. This is *not* a rotation of the contour of integration, it is just a shift in the imaginary direction. The resulting integration over $\bar{F}_{D-1}$ is nothing other than the plain old Gaussian integral, yielding a determinant $e^{b\phi/2}$ in the path integral. We can absorb the latter in the measure of the dilatonic path integral by changing variables as follows:

$$e^{b\phi/2}\, d[\phi] = 2/b\; d[e^{b\phi/2}]\,. \tag{3.90}$$

This means we are treating the exponential of the dilaton as fundamental. As long as we only sum over positive values of the exponential, this does not affect anything. The change of variables is valid because the exponential is a strictly monotonic function of the dilaton, and hence injective. When the smoke clears, we are left with the following system:

$$\int d[\lambda]\, \exp\left(-\int_{\mathcal{M}} \tfrac{1}{2}\left[d\phi \wedge *d\phi + e^{b\phi}\, d\lambda \wedge *d\lambda\right] + i \int_{\partial M} *J\, \lambda\right), \tag{3.91}$$

where no boundary conditions are imposed on λ. The constraint $dF_{D-1} = 0$ translates to $d * J = 0$, i.e. the external current must be divergenceless. The important thing to notice is that the kinetic term of λ has the 'normal' sign. Contrary to common belief, a quantum mechanical dualization does *not* yield a negative action scalar. The boundary term in this path integral corresponds to the two surface[15] terms in (3.9). This boundary term, combined with the fact that boundary values of λ are being integrated over, plays the role of a Fourier transformation of the boundary states. The path integral does not compute a transition amplitude between field eigenstates $|\,\lambda\,\rangle$, but between momentum eigenstates $|\,\pi\,\rangle \equiv \int d[\lambda]\, \exp(i\,\pi\,\lambda)\,|\,\lambda\,\rangle$.

[15] There is ambiguity in defining the boundary at infinity of a manifold. Although the surface terms in (3.9) are only defined on disconnected 'initial' and 'final' hypersurfaces, I believe that defining a single, connected, radial boundary at $r = \infty$ leads to equivalent results.

Note that the shift $\mathbb{R}$-symmetry of the axion is now broken to a $\mathbb{Z}$-symmetry by the surface term:

$$\lambda \to \lambda + \frac{2\pi n}{c}\,, \qquad \text{where} \quad c \equiv \int_{\partial M} *J\,, \qquad \text{and} \quad c \in \mathbb{Z}\,. \tag{3.92}$$

In theories where λ is periodically identified, the single-valuedness of the path integral imposes a quantization condition on c. String theory effects are expected to induce such a quantization, [57, 58]

Let us naïvely try to approximate (3.91) by means of the saddle point approximation. Because there are no boundary conditions on λ, variations need not vanish on the boundary. The Euler-Lagrange variation of the action then yields

$$\delta S = \int_{\mathcal{M}} d\left(e^{b\phi} * d\lambda\right)\delta\lambda - \int_{\partial\mathcal{M}} \left(e^{b\phi} * d\lambda - i * J\right)\delta\lambda\,. \tag{3.93}$$

For arbitrary $\delta\lambda$, this imposes a rather normal equation of motion for the axion in the bulk

$$d\left(e^{b\phi} * d\lambda\right) = 0\,. \tag{3.94}$$

However, it also imposes the following boundary condition on the current of the axion shift symmetry:

$$e^{b\phi}\, d\lambda\big|_{\partial\mathcal{M}} = i\,J\,. \tag{3.95}$$

This constraint is rather strange, as it would imply that the saddle point approximation requires λ to be imaginary. Hence, the path integral has no real saddle points. However, it is possible to perform a semiclassical approximation of it in two ways: the first method consists in using the fact that this path integral is at most quadratic in λ to compute it. The idea is that one can split up the integral into an integration over bulk fields with Dirichlet boundary conditions followed by one over the boundary fields. The former can be evaluated in the usual way by using the variational principle, since it is just a Gaussian. Then, by performing the integral over the boundary fields, one is basically Fourier transforming this result. However, this method is very cumbersome, as it requires an explicit choice of the boundary. The second method relies on the dualization procedure we described. This is a far simpler and more covariant approach, and we will be using it to evaluate the actions of our solutions. The idea is that, since the axion path integral (3.91) and the field-strength path integral (3.85) are equal to each other, instead of trying to evaluate the former, which has no real saddle points, one can just evaluate the latter, which does have saddle points. This indirectly yields a semiclassical approximation of the axion theory.

If we use the constraint dF_{D-1}, we can treat the $(D-1)$-form as locally exact; i.e $F_{D-1} = dC_{D-2}$. Then, we can derive the following equation of motion:

$$d(e^{-b\phi} * F_{D-1}) = 0\,, \tag{3.96}$$

which means that, locally, one can rewrite the field-strength as follows:

$$F_{D-1} = e^{b\phi} * d\chi\,, \tag{3.97}$$

where χ is a scalar. The equation of motion of the dilaton is the following:

$$d * d\phi + \frac{b}{2} e^{-b\phi} F \wedge *F = 0\,. \tag{3.98}$$

Substituting the definition of χ into this yields the following:

$$d * d\phi + \frac{b}{2}\, d\chi \wedge *d\chi = 0\,. \tag{3.99}$$

This equation of motion has the wrong sign in front of the χ term. One can similarly show that the Einstein equation also 'sees' a dilaton with the wrong sign. Hence, the remaining equations of motion of the resulting system are the ones we have been solving in this chapter; i.e. those of a system with a wrong sign kinetic term for the axion. At the end of the day, the result of solving the F_{D-1} equations and substituting the solution into (3.85) is effectively the same as performing a saddle point approximation of a 'would-be' imaginary scalar field χ with the following action:

$$S = \int_{\mathcal{M}} \tfrac{1}{2} \left[d\phi \wedge *d\phi - e^{b\phi}\, d\chi \wedge *d\chi + 2\, d\left(\chi\, e^{b\phi}\, * d\chi\right)\right], \tag{3.100}$$

and with the following Neumann boundary conditions for the axion current:

$$e^{b\phi}\, d\chi\big|_{\partial M} = J\,. \tag{3.101}$$

where J is the external current in (3.91) and the Hodge dual of the boundary value of F_{D-1} in (3.85). The equations of motion of the would-be scalar field χ seem to imply that J is divergenceless, which is equivalent to the constraint $dF_{D-1} = d * J = 0$. Therefore, the path integral yields a selection rule that enforces momentum conservation.

From now on, we will use the F_{D-1} action in (3.85) to evaluate the action of the non-extremal D-instanton, and the on-shell duality relation (3.97) to translate our 'electric' axionic solutions into dual 'magnetic' solutions.

It is now easy to show that this action satisfies a Bogomol'nyi bound [31]. We can rewrite the action as follows:

$$\mathcal{S}_E = \int_{\mathcal{M}} \tfrac{1}{2}\,(d\phi \wedge *d\phi + e^{-b\phi}\, F_{D-1} \wedge *F_{D-1})\,, \tag{3.102}$$

$$= \int_{\mathcal{M}} \tfrac{1}{2}[(d\phi \pm e^{-b\phi/2} * F_{D-1}) \wedge *(d\phi \pm e^{-b\phi/2} * F_{D-1}) \mp (-)^D \tfrac{4}{b} d(e^{-b\phi/2} F_{D-1})]\,, \tag{3.103}$$

where we have used the fact that $dF_{D-1} = 0$. Since the first term is positive semi-definite $\mathcal{S}_E$ is bounded from below by a topological surface term given by the last term in (3.103). The bound is saturated when the Bogomol'nyi equation

$$*F_{D-1} = \mp e^{b\phi/2} d\phi\,, \tag{3.104}$$

is satisfied. The $\mp$ distinguishes instantons from anti-instantons, and for simplicity, we will use the upper sign from now on. Using (3.97), one can write the Bogomol'nyi equation as

$$d\chi = -e^{-b\phi/2} d\phi\,, \tag{3.105}$$

and one can check explicitly that the instanton solutions with $q^2 = 0$, given in (3.41), satisfy this bound. They are therefore rightfully called extremal. The instanton action can then easily be evaluated, and has only a contribution from the boundary at infinity,

$$\mathcal{S}^{\infty}_{inst} = \frac{4}{b^2}(D-2)\mathcal{V}ol(S^{D-1})\frac{|bcq_-|}{g_s^{b/2}}\,, \tag{3.106}$$

while the contribution from $r = 0$ vanishes.

For $D = 10$ and $b = 2$, this value of the instanton action precisely coincides with [30]. For other values of b, we notice the dependence of g_s on b. In ten dimensions, the only possible value for b compatible with maximal supersymmetry is $b = 2$. One then finds that the instanton action depends linearly on the inverse string coupling constant. In lower dimensions this is not necessarily so, and more values for b are possible, depending on whether χ comes from the RR sector or from the NS sector. This would imply different kinds of instanton effects, with instanton actions that depend on different powers of the string coupling constant. This indeed happens for instance in four dimensions, after compactifying type IIA strings on a Calabi-Yau threefold. There are D-instantons coming from wrapping (Euclidean) D2 branes around a supersymmetric three-cycle, and there are NS5-brane instantons coming from wrapping the NS5-brane around the entire Calabi-Yau. As explained in [59], such instanton effects are weighted with different powers of g_s in the instanton action. This was also explicitly demonstrated in [60–62]. In our notation, they correspond[16] to $b = 1$ and $b = 2$. Our results in (3.106) are consistent with these observations.

Notice also that the instanton action is proportional to q_-. For extremal instantons, this is precisely the mass of the corresponding black hole one dimension higher, see (3.71). This is the generic characteristic of the instanton-soliton correspondence that we explained in subsection 2.3.2. There, the Euclidean action of the instanton in D dimensions equals the mass or Hamiltonian of the black hole soliton in $D+1$ dimensions. It is interesting to note that this also happens for theories with gravity.

We now turn to the case of non-extremal instantons, and focus first on the case of $q^2 > 0$. The solutions (3.33) for the dilaton and axion fields can be written as

$$d\phi = \frac{2}{b}\coth(H + C_1)dH\,, \qquad e^{-b\phi/2}F_{D-1} = \frac{2}{b}\frac{*dH}{\sinh(H + C_1)}\,, \tag{3.107}$$

and do not satisfy the Bogomol'nyi equation (3.104). To evaluate the action on this non-extremal instanton solution, we substitute these expressions into the bulk action (3.102), and find

$$\mathcal{S}_{\text{scalars}} = \frac{2}{b^2}\int d\Big(\{H - 2\coth(H + C_1)\} * dH\Big)\,, \tag{3.108}$$

which is a total derivative term. Evaluating the Ricci scalar on the solution in (3.33) we find the following:

$$S_R = -\int_{\mathcal{M}} R = -\frac{2}{b^2}\int_{\mathcal{M}} d(H * dH)\,, \tag{3.109}$$

which precisely cancels the first term of the scalar action (3.108). Hence, the bulk action is given by the following:

$$S_R + S_{\text{scalars}} = -\frac{4}{b^2}\int d\left(\coth(H + C_1)\, * dH\right)\,, \tag{3.110}$$

[16] This corrects a minor mistake in the previous version and in the version published in *JHEP*. In our conventions, the $D = 4$ dilaton is related to the $D = 10$ string dilaton by a factor of 2, see [63] for further details and implications of this correction.

which is again a total derivative. In fact, had we used the pseudo action in (3.100), we would have also ended up with a total derivative of the form $d\big(\chi\, e^{b\,\phi}\, *\, d\chi\big)$, which would have yielded the same result.

Using Stokes theorem, we only pick up contributions from the boundaries. Since the $q^2 > 0$ instantons have a curvature singularity at $r = r_c$ (see section 3.1), one can take these boundaries at $r = \infty$ and at $r = r_c$. In terms of the variable H, this corresponds to $H = 0$ and $H = \infty$ respectively[17]. We stress again that we have taken C_1 to be positive, in order to avoid further singularities in the scalar sector when $H + C_1 = 0$.

Besides the bulk action, one also needs to include the Gibbons-Hawking term [64], to make the action consistent with the Einstein equations:

$$\mathcal{S}_{GH} = -2\int_{\partial\mathcal{M}} (K - K_0)\,, \tag{3.111}$$

where $\mathcal{M}$ is the D-dimensional Euclidean space and $\partial\mathcal{M}$ is the boundary. In the second term, K is the trace of the extrinsic curvature of the boundary and K_0 the extrinsic curvature one would find for flat space, which is subtracted to normalize the value of the action. The extrinsic curvature is defined in terms of a unit vector n^μ that is normal to the boundary as follows:

$$K \equiv h_\mu{}^\nu\, \nabla_\nu\, n^\mu\,, \tag{3.112}$$

where $h_\mu{}^\nu$ is the tensor that projects components onto the boundary.

Let us now evaluate the total action at both $r = \infty$ and $r = r_c$: we first discuss the boundary at $r = \infty$. The contribution from (3.111) vanishes, while (3.110) yields a contribution

$$\begin{aligned} \mathcal{S}^{\infty}_{inst} &= \frac{4}{b^2}\,(D-2)\mathcal{V}ol(S^{D-1})\, b\, c\Big(q\ \coth C_1\Big)\,, \\ &= \frac{4}{b^2}\,(D-2)\,\mathcal{V}ol(S^{D-1})\, b\, c\Big(\sqrt{q^2 + \frac{q_-^2}{g_s^b}}\Big)\,. \end{aligned} \tag{3.113}$$

In the second line, we have used the relation between C_1 and the asymptotic value of the dilaton, $g_s^b = (q_-/q)^2\ \sinh^2 C_1$.

For $q^2 = 0$, (3.113) precisely yields back the result for the extremal instanton, see (3.106). There we made the relation between the instanton action and the black hole mass one dimension higher. Also for the non-extremal instanton, such a relation seems to hold. Indeed, from the mass formula for the non-extremal black hole in terms of the instanton parameters, one has that $q\ \coth C_1 = \sqrt{q^2 + q_-^2}$, and the string coupling constant is set to unity. One therefore sees that the contribution to the instanton action from the boundary at infinity is proportional to the black hole mass one dimension higher.

The boundary at $r = r_c$ receives contributions from both integrals (3.110) and (3.111), which add up to

$$\mathcal{S}^{r_c}_{inst} = \frac{4}{b^2}\,(D-2)\mathcal{V}ol(S^{D-1})\, b\, c\Big(q\,(\frac{bc}{2} - 1)\Big)\,. \tag{3.114}$$

[17] Without loss of generality, we can choose $q > 0$.

Because the dilaton *and* the curvature blow up at r_c, the supergravity approximation and string perturbation theory both break down. Hence, it is not clear whether this contribution is meaningful. One might take the point of view that string theory corrections, which are expected to take over at r_c, would actually smooth the singularities out. In that case, there would be no need to consider this point as a boundary, and no need to take this contribution into account. It is also plausible, however, that string theory corrections completely modify the geometry, 'opening up' a wormhole that leads into a whole new space. In that case, a second boundary would exist, but the values of the fields might be different there.

Note that this contribution vanishes for the case $bc = 2$, while it is positive for $bc > 2$. However, as discussed above, it is not at all clear whether this contribution to the integrals (3.110) and (3.111) should be included in the instanton action, since it is calculated in a region of space where the supergravity approximation is no longer valid.

We now turn to the case of $q^2 < 0$, or with $q = i\tilde{q}$, a positive $\tilde{q}^2 > 0$. A similar calculation as for $q^2 > 0$ shows that, for the solution (3.45), we have

$$d\phi = \frac{2}{b}\cot(\tilde{H} + \tilde{C}_1)d\tilde{H}\,, \qquad e^{-b\phi/2}F_{D-1} = \frac{2}{b}\frac{*d\tilde{H}}{\sin(\tilde{H} + \tilde{C}_1)}\,, \tag{3.115}$$

where

$$\tilde{H} = bc\arctan(\frac{\tilde{q}}{r^{D-2}})\,, \tag{3.116}$$

is a harmonic function over the geometry given by the metric in (3.45). Plugging in these expressions into the bulk action (3.102), we find

$$\mathcal{S}_{inst} = -\frac{2}{b^2}\int d\Big(\{\tilde{H} + 2\cot(\tilde{H} + \tilde{C}_1)\} * d\tilde{H}\Big)\,. \tag{3.117}$$

Since this is a total derivative, we can use Stokes theorem again to reduce it to an integral over the boundaries. These boundaries are at $r = \infty$ and $r = 0$, where we required that $bc < 2$, as discussed in section 3.1. In contrast to the discussion of the $r = r_c$ boundary for $q^2 > 0$, the instanton solution is perfectly regular everywhere, in particular at both boundaries. Therefore the contribution from the boundary at $r = 0$ can also be trusted.

In addition to the above action, one also needs to include the gravitational contribution (3.111). Similar to the case of $q^2 > 0$, the first term of (3.117) is cancelled by the contribution from the Ricci scalar. We anticipate the Gibbons-Hawking term not to contribute, since the two asymptotic geometries at $r = 0$ and $r = \infty$ are equivalent due to the $\mathbb{Z}_2$-symmetry (3.74). Hence, their contributions should cancel.

Therefore the $q^2 < 0$ instanton action has contributions only from the second term of (3.117) from both boundaries at $r = 0$ and $r = \infty$:

$$\begin{aligned}\mathcal{S}^{\infty}_{inst} &= \frac{4}{b^2}(D-2)\mathcal{V}ol(S^{D-1})\,b\,c\,\tilde{q}\Big(\cot\tilde{C}_1\Big)\,,\\ \mathcal{S}^{0}_{inst} &= \frac{4}{b^2}(D-2)\mathcal{V}ol(S^{D-1})\,b\,c\,\tilde{q}\Big(-\cot(\tilde{C}_1 + bc\frac{\pi}{2})\Big)\,.\end{aligned} \tag{3.118}$$

Due to the fact that $\tilde{C}_1$ and $\tilde{C}_1 + bc\pi/2$ are on the same branch of the cotangent (due to the restriction of regular scalars for $0 < r < \infty$, which can only be achieved for $bc < 2$, see section 3.1), the total instanton action is manifestly positive definite. In the neighborhood of $bc \approx 2$, the instanton action becomes very large, and the limit to the extremal point where $bc = 2$, is discontinuous. This shows that this instanton is completely disconnected from the extremal D-instanton.

Using the asymptotic value of the dilaton in (3.45), we have $g_s^b = (q_-/\tilde{q})^2 \sin^2 \tilde{C}_1$, and therefore $\tilde{q}^2 < q_-^2/g_s^b$. Assuming that $\cot \tilde{C}_1 > 0$, the contribution from infinity is positive and can be rewritten as

$$\mathcal{S}_{inst}^{\infty} = \frac{4}{b^2}\,(D-2)\,\mathcal{V}ol(S^{D-1})\,b\,c\,\sqrt{\frac{q_-^2}{g_s^b} - \tilde{q}^2}\,, \tag{3.119}$$

which is the analytic continuation of the result with $q^2 > 0$.

3.6.2 Tunneling interpretation

The reader may wonder what the tunneling interpretation of a D-instanton is. In a standard non-gravitational QFT, the metric is fixed and one always knows what the Euclidean time direction is, because one knows how the theory was Wick rotated in the first place. In a theory where the metric is dynamical, however, this is not straightforward at all. Since the Euclidean spacetime is not part of the input, but rather the outcome of the equations of motion, which direction is viewed as time-like is not determined *a priori*. For our solutions, one might be tempted to think of r as the Euclidean time parameter, since all fields depend on it. However, this wouldn't lead to the tunneling interpretation we are after. Take for instance the case $q^2 = 0$, which has a flat space. Let us Wick rotate this back to Lorentzian signature taking the r direction to be time:

$$dr^2 + r^2\, d\Omega^2_{S^{D-1}} \to -dt^2 + t^2\, d\mathbb{H}^2_{D-1}\,. \tag{3.120}$$

See chapter 7 for a derivation of this Wick rotation. The initial slice $t = 0$ is singular, and the later slices are hyperbolic spaces. These are not the initial and final states one would like to have for a tunneling interpretation. The more natural Wick rotation takes place in Cartesian coordinates. Letting $r = (x_0^2 + ... + x_{D-1}^2)^{1/2}$, and rotating $x_0 \to i\,t$.

Another reason not to pick r as a time direction is the fact that, for our solutions, the axion current $e^{b\,\phi}\,\partial\chi$ would be conserved in the r direction, since our Ansatz is such that the axion equation of motion is the following:

$$\partial_\mu \left(e^{b\,\phi}\, \nabla^\mu \chi\right) = \partial_r \left(e^{b\,\phi(r)}\, \nabla^r \chi(r)\right) \sim \delta(r)\,. \tag{3.121}$$

This means that, in the r direction, there would be no charge conservation violation due to tunneling, and hence no interesting tunneling effect in any way. If we pick x_0, however, then the point $r = 0$ will act as a source-like singularity[18] (for the cases with $q^2 \geq 0$) and will generate a charge difference between the initial and final states. See figures 3.5(a) and 3.5(b). One can calculate that this difference will be $\sim q_-$ for our solutions. Classically one could say that the δ-function in the equations of motion for χ can be reproduced by adding a source term in the action of the form $\chi\,\delta(r)$. From the point of view of the path integral in (3.91), one should

[18] This is basically because $\Box\, H(r) \sim \delta(r)$ for our harmonic functions.

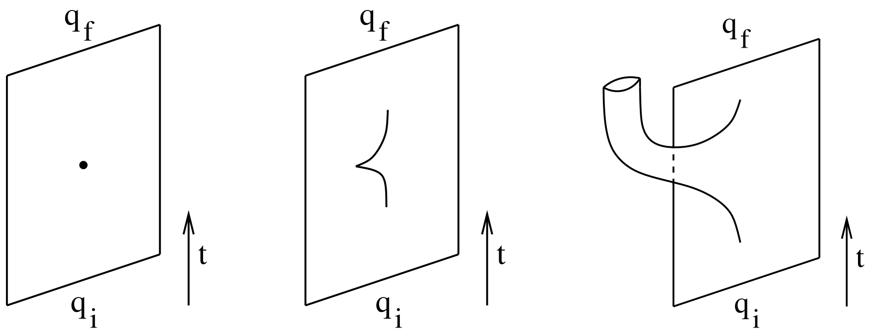


Figure 3.5: *The tunneling interpretation. Figures (a), (b), and (c) depict the* $q^2 = 0, > 0$, *and* < 0 *respectively. The first two solutions have a charge conservation violation because they have electric source-like singularities at the origin* $r = 0$*. The worhole (c) conserves its total charge, but splits up into two disconnected spaces* $\mathbb{R}^{D-1} \oplus S^{D-1}$*, so that an observer on* $\mathbb{R}^{D-1}$ *will see a charge loss.*

add a term of the form $i\,\lambda\,\delta(r)$ to the action; and in the path integral in (3.85) this corresponds to adding $i\,dF_{D-1}\,\delta(r)$. This will supplement the charge conservation or closedness constraint, respectively:

$$dF_{D-1} + \delta(r) = d * J + \delta(r) = 0\,. \tag{3.122}$$

Such a term is a local operator insertion or vertex operator in the path integral, i.e. $\langle \exp(i\,\lambda(r_0))\rangle$.

In the case of the wormhole, the point $r = 0$ is not included in the manifold, and there is no δ-function-like singularity. In order to find the tunneling interpretation, one must first cut the wormhole in half at its neck, and suggestively redraw the shape of the remaining geometry as in figure 3.5(c), as was done in [28]. The axion charge is globally conserved, but the manifold splits up into two disconnected spaces as follows:

$$\mathbb{R}^{D-1} \to \mathbb{R}^{D-1} \oplus S^{D-1}\,. \tag{3.123}$$

Although the total charge is conserved, an observer who stays on the $\mathbb{R}^{D-1}$ will see a charge loss, because the S^{D-1} *baby universe* will carry off some charge with it. From the string theory point of view, it is possible that, for the case $q^2 > 0$, string theory corrections will change the singular geometry into a smooth one, perhaps by ‘opening up’ a wormhole-like geometry where the singularity was. This would restore global axion charge conservation.

3.6.3 Correlation functions

Once the instanton solutions are established, one would like to study their effect in the path integral. As for D-instantons in ten-dimensional IIB, they contribute to certain correlation functions via the insertion of fermionic zero modes. For the D-instanton, which is 1/2 BPS, there are sixteen fermionic zero modes. These are solutions for the fluctuations that satisfy the linearized Dirac equation in the presence of the instanton. All of these zero modes can be generated by acting with the broken supersymmetries on the purely bosonic instanton solution. For the non-extremal instantons, no supersymmetries are preserved, so there are more fermionic zero modes. Let us focus for simplicity on ten-dimensional type IIB. Since all supercharges are broken, one can generate 32 fermionic zero modes. The path integral measure contains an integration over these fermionic collective coordinates, and to have a non-vanishing result, one must therefore insert 32 dilatinos in the path integral. Based on this counting argument of fermionic zero modes, a 32-point correlator of dilatinos would be non-zero, and induce new terms in the effective action, containing 32 dilatinos. In the full effective action, such terms are related to higher curvature terms like e.g. certain contractions of R^8. An explicit instanton calculation should be done to determine the non-perturbative contribution to the function that multiplies R^8. As for the D-instanton, we expect that the contributions of the instantons with different q^2-values build up a modular form with respect to $SL(2,\mathbb{Z})$, possibly after integrating over q^2.

These issues, though important, lie beyond the scope of this chapter, and are left open for investigation.

3.7 Discussion

In this chapter we investigated non-extremal instantons in string theory that are solutions of a gravity-dilaton-axion system with dilaton coupling parameter b. In particular, we constructed an $SL(2,\mathbb{R})$ family of spherically symmetric instanton-like solutions in all conjugacy classes labelled by q^2. Among these is the (anti-)D-instanton solution with $q^2 = 0$. For special values of the dilaton coupling parameter this solution is half-supersymmetric. The instanton solutions in the other two conjugacy classes, with $q^2 > 0$ and $q^2 < 0$, are non-supersymmetric and can be viewed as the non-extremal versions of the (anti-)D-instanton. This view is confirmed by the property that instantons in these two conjugacy classes, for $bc \geq 2$ with c defined in (3.26), can be uplifted to non-extremal black holes.

We stressed the wormhole nature of the instanton solutions. We found that each conjugacy class leads to a wormhole geometry provided the corresponding instanton is given in a particular metric frame:

$$\begin{aligned} q^2 > 0 \quad &\leftrightarrow \quad \text{dual frame metric (only for } bc = 2 \text{ or } q_- = 0) \\ q^2 = 0 \quad &\leftrightarrow \quad \text{string frame metric} \\ q^2 < 0 \quad &\leftrightarrow \quad \text{Einstein frame metric} \end{aligned} \tag{3.124}$$

For all these cases the metric takes the form (3.46), with the specific values given in section 3.2.

Not all instanton solutions we constructed are regular and not all can be uplifted to black holes. The non-extremal instantons in the $q^2 > 0$ conjugacy class all have a curvature singularity at $r = r_c$, see (3.35). Only the $bc = 2$ instanton can be uplifted to a regular non-extremal RN black hole with the singularity being resolved as a coordinate singularity at the outer event horizon of the RN black hole. The singularity remains for $bc > 2$ and in that case can be resolved by adding an extra dilaton to the original system [51]. Two exceptions are the limits $q^2 \to 0$ or $q_- \to 0$, which correspond to the extremal and Schwarzschild black hole solutions, respectively. Finally, the instantons in the $q^2 < 0$ conjugacy class are only regular for $bc < 2$. These instantons can never be uplifted to black holes.

We have also considered the uplift of our instanton solutions to p-branes. It turns out that an instanton can only be uplifted over a $(p+1)$-torus to a p-brane provided the dilaton coupling satisfies (following from (3.84))

$$b\,c \geq \sqrt{\frac{4(p+1)(D-1)}{D+p-1}}\,. \tag{3.125}$$

For the case that saturates this bound, the instanton with $q^2 \geq 0$ uplifts to a regular non-dilatonic p-brane. For larger values of b, the instanton solution (3.24) with $q^2 > 0$ uplifts to a singular limit of the dilatonic p-branes of [53]. These solutions only become regular in the limit $q^2 \to 0$ or $q_- \to 0$. A summary of the possible regular solutions is given in table 3.1. Alternatively, we have discussed the possibility of adding an extra dilaton to the instanton solution [51], which allows for the uplift to the regular dilatonic p-branes of both type 1 and type 2.

bc	Dimension	Regular solutions
< 2	D	Instantons with $q^2 \leq 0$, see (3.45)
$= 2$	$D+1$	RN black holes with $q^2 \geq 0$, see (3.67), or Schwarzschild black holes with $q^2 > 0$, $q_- = 0$
> 2	$D+1$	Dilatonic black holes with $q^2 = 0$ or Schwarzschild black holes with $q^2 > 0$, $q_- = 0$
$=$ in (3.125)	$D+p+1$	Non-dilatonic p-branes with $q^2 \geq 0$
$>$ in (3.125)	$D+p+1$	Dilatonic p-branes with $q^2 = 0$ or $q^2 > 0$, $q_- = 0$

Table 3.1: *The regular instanton, black hole and p-brane solutions that are obtained, depending on the dilaton coupling parameter b, the conjugacy class q^2 and the charge q_-.*

For the particular value $b = 2$, corresponding to $\Delta = 4$, there is another higher-dimensional origin. In this special case, the D-dimensional extremal instanton can be uplifted to a gravitational wave in $D+2$ dimensions [35]. Similarly, the other two conjugacy classes uplift to purely gravitational solutions in $D+2$ dimensions which we denominate “non-extremal waves”. The terminology is slightly misleading since the uplift only leads to a time-independent solution. Whether this solution can be extended to a time-dependent wave-like solution remains to be seen. It is also interesting to note the following curiosity: the source term for a pp-wave is a massless particle, i.e. a particle with a null-momentum vector: $p^2 = 0$. This suggests that we associate the source terms for the other two conjugacy classes with massive particles ($p^2 > 0$) and tachyonic particles ($p^2 < 0$). We leave this for future investigation.

In the second part of this chapter, we investigated whether the non-extremal instantons might contribute to certain correlation functions in string theory. For this application, it is a prerequisite that there be a well-defined and finite instanton action. Mimicking the calculation of the standard D-instanton action, we found that for $q^2 > 0$ the contribution from infinity to the instanton action, for all values of b, is given by the elegant formula (3.113). This action reduces to the standard D-instanton action for $q^2 = 0$. Having a finite action, the non-extremal instantons might contribute to certain correlation functions. In the case of type IIB string theory, we conjectured that non-extremal instantons contribute to the R^8 terms in the string effective action in the same way that the extremal D-instantons contribute to the R^4 terms in the same action. Whether the fact that all supersymmetries are broken by the non-extremal instantons poses problems remains to be seen. An explicit instanton calculation should decide whether our conjecture is correct. We leave this for future investigation.

Finally a few comments on some work in progress [65]. A natural and very interesting generalization to the solutions in this chapter can be achieved by adding a negative cosmological constant in the action. Just as the solutions we have studied here are asymptotically flat, solutions in a system with a cosmological constant are asymptotically *anti-de Sitter* or AdS. Asymptotically AdS spaces are particularly interesting in light of Maldacena's breakthrough in [1], where he conjectured that type IIB string theory in an $AdS_5 \times S^5$ background is completely equivalent to $\mathcal{N} = 4, d = 4$ super-Yang-Mills theory. The stronger version of his conjecture states that string theory on an asymptotically $AdS_5 \times S^5$ background is dual to some deformation of super-Yang-Mills. This duality has been used to show that the extremal D-instanton of type IIB supergravity corresponds to the super-Yang-Mills self-dual instanton [66–70]. It would be interesting to see what the field theory dual of a non-extremal D-instanton is. Perhaps it contains information about non-self-dual Yang-Mills instantons.

This concludes the first part of this thesis, which covered the topic of instantons. In the next two chapters, we will look at a different kind of scalar-gravity solutions that also have interpolating behavior: cosmological solutions. These are solutions of the Einstein equations that also depend on only one parameter, however, that parameter is Lorentzian time.

Chapter 4

Introduction to Cosmology

4.1 FLRW cosmology

To begin our studies of cosmology, we must first introduce a bit of formalism and terminology that is now part of what is called *the standard cosmology*. The language and formulae in which we will state facts about cosmology are deceitfully simple. They hide the massive amounts of observational data and research required to arrive at them. Doing justice to the topic of modern cosmology would obviously require a lot more than one chapter. For a proper introduction to standard cosmology and cosmology in the context of string theory, the reader is referred to the lecture notes [71–73], on which this chapter is mainly based. Often in physics one tries to reproduce or model complicated phenomena by defining a fundamental[1] theory that is simple to begin with, but requires all kinds of approximations and truncations in order to describe realistic physics. In cosmology, one does the exact opposite. One tries to model complicated phenomena with simple models, which are not really *derived* from a fundamental theory. They can ultimately be seen as large scale gross approximations of some unknown fundamental theory. When discussing inflation, F. Quevedo describes it as "a scenario in search of an underlying theory" [72]. A fundamental theory that could account for cosmology would also have to explain the Big Bang. General Relativity breaks down for highly curved spacetimes, where quantum effects become important. String theory is a current candidate as an underlying theory of cosmology because it is a theory of quantum gravity.

4.1.1 The FLRW Anstatz: Motivation and definition

We begin by defining the FLRW, or *Friedmann-Lemaître-Robertson-Walker* spacetime metric. It is actually a class of metrics defined by two properties as follows: a metric is FLRW if there exists a frame (i.e. a family of geodesic observers), in which it is *spatially* homogeneous and isotropic (see appendix C for definitions and examples). These two properties that are imposed are based on the observations that the universe "looks the same" at every point in space, and it

[1] Of course, the concept of a *fundamental* theory is only relative. So far there is no such thing as a fundamental theory that is valid in all regimes.

"looks the same" in every direction about a point. Of course this is only true on a very very large scale, a cosmological scale. Our lives would be pretty difficult if we were not capable of telling the difference between our boss' office and our bathroom, and driving would be impossible if we couldn't make a distinction between the right and the wrong way of a one-way street. But we as humans are looking too closely at things and what we see are only tiny fluctuations from homogeneity and isotropy.

So the Ansatz for an FLRW metric is the following:

$$ds^2 = -f^2(t)\,dt^2 + g^2(t)\,d\Sigma_3^2\,, \tag{4.1}$$

where $f(t)$ an $g(t)$ are two undetermined functions of time, and $d\Sigma_3$ is the line element of some homogeneous and isotropic spatial manifold. It can be shown that in three dimensions there are only three possible metrics that satisfy the requirement of homogeneity and isotropy :

$$d\Sigma_3^2 = \frac{dr^2}{1-k\,r^2} + r^2\left(d\theta^2 + \sin(\theta)^2\,d\phi^2\right) \qquad \text{with} \quad k = +1, 0, 1\,. \tag{4.2}$$

This can also be written as follows:

$$d\Sigma_3^2 = d\rho^2 + f^2(\rho)\left(d\theta^2 + \sin(\theta)^2\,d\phi^2\right)\,, \tag{4.3}$$

where

$$f(\rho) = \begin{cases} \sin(\rho) & \text{if } k = +1 \\ \rho & \text{if } k = 0 \\ \sinh(\rho) & \text{if } k = -1 \end{cases}\,. \tag{4.4}$$

The parameter k lables the curvature of the spatial section of the metric. A spatial section of (4.1) with line element

$$ds^2_{spatial} = g^2(t)\,d\Sigma_3^2 \tag{4.5}$$

has the following Ricci scalar:

$$R_\Sigma = \frac{6\,k}{g^2(t)}\,. \tag{4.6}$$

We easily recognize the three spatial metrics as those of the 3-sphere, 3-plane and 3-hyperboloid respectively. But we must be careful not to confuse local with global statements about a manifold. The three spatial metrics in (4.2) contain only local information and do not imply anything about the topologies of their respective manifolds. For instance, the $k = 0$ metric may be defined on the 3-plane R^3 as well as on the 3-torus T^3. Similarly, the 3-hyperboloid H^3 can be compactified by means of discrete group identifications that do not affect curvature. So what does the metric $g^2(t)\,d\Sigma_3^2$ tell us about a spatial manifold? Any physically meaningful statement in General Relativity must be expressible in terms of "clock and rods", and in this case specifically, in terms of "rods".

Let us start with the spatially flat ($k = 0$) case. We place an observer at time $t = t_0$ at the origin of our coordinate system ($\rho = 0$) and at rest w.r.t. it ($\dot{\rho} = 0$). Let the observer pick a plane passing through him (without loss of generality the $\theta = \pi/2$ plane), and draw a circle on it around himself of radius

$$R = g(t_0)\rho' \quad \text{for some} \quad \rho'\,, \tag{4.7}$$

If the observer measures the circumference L of this circle instantaneously, or fast enough so that $g(t)$ does not change significantly, the metric (4.5) tells us that he will find it to be

$$L = 2\pi\, g(t_0)\, \rho' = 2\pi R\,, \tag{4.8}$$

as expected. For general k this will change. If we conduct the same experiment, (4.5) tells us that the circumference of a circle of radius $R = g(t_0)\, \rho'$ will be

$$L = 2\pi\, g(t_0)\, f(\rho') = \begin{cases} 2\pi\, g(t_0)\, \sin[R/g(t_0)] & \text{if } k = +1 \\ 2\pi R & \text{if } k = 0 \\ 2\pi\, g(t_0)\, \sinh[R/g(t_0)] & \text{if } k = -1 \end{cases} . \tag{4.9}$$

The first thing to notice about this result is that $g(t_0)$ completely drops out for the $k = 0$ case, making its value at any given time physically meaningless. The other thing to notice is that if we take R to be very small and expand $f(\rho)$, we see that, to leading order, the circumferences become $2\pi R$ for the $k \neq 0$ cases. If this were not the case, we would have what is called a *conical singularity* on our spatial manifold. Hence, the $k = +1$ case tells us that circles have smaller circumferences than we are used to, and the $k = -1$ tells us that they are larger than normal.

Now that we understand the spatial geometry of the FLRW metric, let us study the spacetime geometry. Once k is fixed, the only undetermined parts of the metric (4.1) are the time-dependent functions $f(t)$ and $g(t)$. However, these two functions are not independent of each other. If we perform the following simple coordinate transformation:

$$\tau(t') \equiv \int_0^{t'} f(t)\, dt\,, \tag{4.10}$$

we end up with the following metric:

$$ds^2 = -d\tau^2 + a^2(\tau)\, d\Sigma_3^2\,, \tag{4.11}$$

where we are now left with only one undetermined function $a(\tau)$, usually called the *scale factor*. The time coordinate τ as defined in (4.11) is called *cosmic time*. In the standard cosmology jargon, if the scale factor is an increasing or decreasing function of time we say that the universe is "expanding" or "contracting" respectively. Similarly, if its second time derivative is positive, we say that the universe is "accelerating". But these words can be misleading. If the spatial topology of the universe is compact, one can define a volume of the universe, and then it makes sense to talk about expansion or contraction. But if the universe has a non-compact spatial topology, such as R^3 or H^3, then this does not make sense. So what does the scale factor really tell us about the universe? Again, the only meaningful thing to do is to revert to our "clocks and rods". The only information we can and should infer from a metric is what geodesic observers see. So let us define two geodesic trajectories $x_1(t)$ and $x_2(t)$ as follows:

$$x^0(t) = \tau(t) = t, \quad x^i(t) = x^i(\tau) = a^i \tag{4.12}$$

$$y^0(t) = \tau(t) = t, \quad y^i(t) = y^i(\tau) = b^i\,, \tag{4.13}$$

where a^i and b^i are constants. Such geodesics are called *comoving*. Notice that for comoving observers the time coordinate τ in (4.11) measures their proper time, so all comoving observers can keep their clocks synchronized. The spatial separation of x_1 and x_2 in the comoving frame is given by:

$$d^2 = d^i\, d^j\, g_{ij}\,, \qquad \text{where} \quad d^i \equiv a^i - b^i\,. \tag{4.14}$$

Differentiating this w.r.t. time we find that

$$\dot{d} = H\, d\,, \tag{4.15}$$

where $H \equiv \dot{a}/a$ is called the *Hubble parameter*. Therefore, the scale factor tells us that two comoving observers will notice a relative velocity between them that is proportional to their separation, and the Hubble parameter. In a universe with accelerated expansion (i.e $H > 0$ and $\ddot{a} > 0$), this means that this relative velocity will eventually exceed the speed of light! Although this may seem like a violation of causality, it is not. No information is travelling from one point to another acausally. What this does mean, however, is that the two observers will eventually cease to be in causal contact, as no signal sent from one can ever catch up with the other.

4.1.2 The right-hand side of the Einstein equation

Having studied the general form of an FLRW cosmological metric, we should now study the kind of matter or energy that can coexist with or drive such a metric. The assumption of spatial isotropy leads us to consider perfect fluids as unique candidates. They have the property (which can be taken as a defining property [18]) of looking isotropic in their rest frames. The stress-energy tensor of a perfect fluid has the following form:

$$T_{\mu\nu} = (\rho + p)\, U_\mu\, U_\nu + p\, g_{\mu\nu}\,, \tag{4.16}$$

where $U^\mu(x)$ is the velocity field of the fluid, ρ is the energy density of the fluid in its rest frame, and p its pressure in its rest frame. This is the stress-energy tensor that will be on the right-hand side of the Einstein equation. In order for the fluid to coexist in equilibrium, or be consistent with the FLRW metric, its elements must be comoving. In other words, in comoving coordinates the velocity field of the fluid must be

$$U^\mu = (1, 0, 0, 0)\,. \tag{4.17}$$

Note that if the fluid is made of photons then U^μ cannot be interpreted as the velocity of the individual photons, but must be interpreted as an average displacement of energy. Using these assumptions we can write the Einstein equations and cleverly rearrange them into the following two equations:

$$H^2 = \frac{8\,\pi\, G}{3}\rho - \frac{k}{a^2}\,, \tag{4.18}$$

$$\frac{\ddot{a}}{a} = -\frac{4\,\pi\, G}{3}\,(\rho + 3\, p)\,, \tag{4.19}$$

where H is the Hubble parameter. The first equation is called the *Friedmann equation* and the second is called the *acceleration equation*. Note that if we want to include several species of fluid we can simply add up the ρ's and p's. The equations of motion for the fluid follow from the conservation laws of the stress-energy tensor:

$$\nabla_\mu T^{\mu\nu} = 0\,. \tag{4.20}$$

They imply the continuity equation for the fluid:

$$\dot{\rho} + 3\,H\,(\rho + p) = 0\,. \tag{4.21}$$

This equation can actually also be obtained by differentiating the Friedmann equation (4.18) w.r.t. time and combining it with the acceleration equation (4.19).

To be able to solve for $a(t)$, $\rho(t)$ and $p(t)$, we need to make one more assumption about the fluid, namely, that it obeys an equation of state. In other words, that the pressure is a function of density, $p = p(\rho)$. For ordinary matter, we can approximate the equation of state by the following instantaneous relation:

$$p = \omega\,\rho\,, \tag{4.22}$$

where ω is a constant that depends on the kind of matter that makes the fluid. For pressureless dust (i.e. non-interacting particles) $\omega = 0$. For radiation, meaning either photons or highly relativistic particles, $\omega = 1/3$. In the case of radiation, one can see this by writing the stress-energy tensor of the Maxwell field:

$$T_{\mu\nu} = -\frac{1}{4\,\pi}\,(F_{\mu\alpha}\,F_\nu{}^\alpha - \frac{1}{4}\,g_{\mu\nu}\,F^2)\,, \tag{4.23}$$

which is manifestly traceless in four dimensions. Our assumptions about comoving perfect fluids tell us that the trace of this tensor is $T_\mu{}^\mu = 3\,p - \rho$. Combining these two facts gives us $\omega = 1/3$.

Dust and radiation are part of a larger class of possible forms of "matter" called *ordinary matter*. Another form of matter is *dark matter*, which is essentially non-baryonic matter. There is another important kind of energy that can drive an FLRW metric, a cosmological constant Λ. It cannot be viewed as matter, it is regarded as a vacuum energy. The cosmological constant also satisfies an equation of state (4.22), with $\omega = -1$, and its energy density is equal to itself, $\rho = \Lambda$. It is part of a class of possible forms of energy called *dark energy*, which characteristically have equations of state with $\omega < -1/3$.

Observations show that our universe is not made of just one kind of fluid, but it is a combination of different kinds of fluids. Also, throughout the history of the universe, the different kinds of matter and energy have swapped the roles of dominance and subdominance. Therefore, a convenient notation for comparing the energy densities of the fluids has been developped. From the Friedmann equation (4.18) we see that the energy density required to have a spatially flat universe is

$$\rho_c = \frac{3\,H}{8\,\pi\,G}\,. \tag{4.24}$$

This is called the *critical* density. By computing the ratio of the actual energy density of a fluid to the critical density

$$\Omega \equiv \frac{\rho}{\rho_c}\,, \tag{4.25}$$

we can easily relate the matter content and observed Hubble parameter of the universe to its spatial geometry as follows:

$$\begin{aligned} \Omega > 1 \quad &\Longleftrightarrow \quad k = 1 \\ \Omega = 1 \quad &\Longleftrightarrow \quad k = 0 \\ \Omega < 1 \quad &\Longleftrightarrow \quad k = -1\,. \end{aligned} \tag{4.26}$$

In a universe with coexisting fluids Ω is simply decomposed into the fractional contributions of each species to the total ratio:

$$\Omega_{\text{total}} = \sum_i \Omega_i\,. \tag{4.27}$$

Observations indicate that our current universe is spatially flat, and it is composed of ordinary (baryonic) matter, dark matter, and dark energy in the following respective ratios:

$$\begin{aligned} \Omega_B &= 0.04 \\ \Omega_{DM} &= 0.26 \\ \Omega_\Lambda &= 0.7\,. \end{aligned} \tag{4.28}$$

A statement of modern cosmology is that the early universe (shortly after the Big Bang) would have been radiation dominated. It is puzzling that, presently, the energy densities of all three forms of matter and energy are of the same order (i.e. $\propto 1$). This puzzle is known as the *cosmic coincidence problem*.

4.1.3 Solutions

Given the matter or energy content of the universe one is trying to model, it is easy to solve for the scale factor by combining the Friedmann and acceleration equations (4.18) (4.19) with the proper equations of state. Since observations show that our current universe is spatially flat to a high degree of precision, we will focus on the $k = 0$ case. The solutions are the following:

$$\begin{aligned} a(t) &= a_0 \left(\frac{t}{t_0}\right)^{2/3\,(1+\omega)} \qquad \text{for} \quad \omega \neq -1 \\ a(t) &\propto e^{H\,t} \qquad\qquad\qquad \text{for} \quad \omega = -1 \end{aligned} \tag{4.29}$$

where H is now constant. The first solution is called *power law* solution. It is mainly used to model pre- and post-inflationary cosmology. Note that for $t = 0$ such a metric has a singularity, namely all spatial distances are zero. This is called the *Big Bang* singularity. The second metric is a solution to the Einstein equation with a *positive* cosmological constant. It is called *de Sitter* space, after Willem de Sitter, the great mathematician, physicist, and astronomer who studied at the University of Groningen. Solutions for $k \neq 0$ can also easily be found.

At this point a word of caution would be in order. Specifying FLRW metrics in terms of k and $a(t)$ is, as we said before, only a local statement about the spacetime manifold. For instance, we noted earlier that division by a discrete group can related two different manifolds with the same local geometry. More generally, we have to remember that a manifold is defined as a

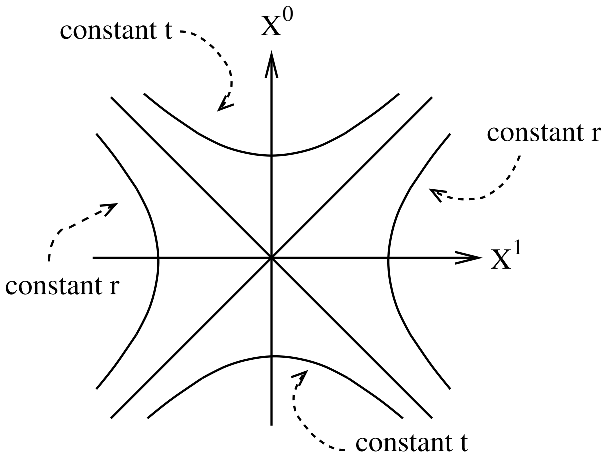


Figure 4.1: *Minkowski spacetime with two suppressed dimensions. A three-dimensional picture can be obtained by rotating everything about the X^0-axis. The constant t surfaces are the two-sheeted Euclidean hyperboloids covering the Milne patch. The constant r surfaces are the one-sheeted Lorentzian hyperboloids (dS) covering the Rindler patch.*

collection of *patches* (i.e. open sets of the underlying space) with *charts* (i.e. coordinates) and *transition functions* relating the charts of intersecting patches. In many cases a single patch may cover the whole space minus a finite set of points. For instance, polar coordinates cover the whole sphere except for the two poles. In such cases, that one patch is all we need. However, some coordinate systems cover only half of a space. So any metric that we write down may just represent one patch of a manifold.

Let us illustrate this with a familiar manifold, Minkowski spacetime. Minkowski spacetime is defined as the manifold $\mathbb{R}^4$ with a flat Lorentzian metric (i.e. Riemann tensor is zero). In cartesian coordinates we write this as follows:

$$ds^2 = -d(X^0)^2 + d(X^1)^2 + d(X^2)^2 + d(X^3)^2 \,. \tag{4.30}$$

So far so good. Now let us introduce the so-called *Milne* coordinates.

$$X^0 = t \cosh(\psi) \,,$$
$$X^1 = t \sinh(\psi) \sin(\theta) \sin(\phi) \,,$$
$$X^2 = t \sinh(\psi) \sin(\theta) \cos(\phi) \,, \tag{4.31}$$
$$X^3 = t \sinh(\psi) \cos(\theta) \,. \tag{4.32}$$

These coordinates don't cover all of Minkowski spacetime. They only cover the regions within the future and past light-cones of the origin of Minkowski spacetime:

$$(X^0)^2 - ||\vec{X}||^2 = t^2 > 0 \,. \tag{4.33}$$

Milne coordinates slice up the space with a one-parameter family of two-sheeted Euclidean hyperboloids, parametrized by t, see figure 4.1. In these coordinates, the flat metric (4.30) becomes

$$ds^2 = -dt^2 + t^2 \left(d\psi^2 + \sinh^2(\psi)\, d\Omega^2_{S^2}\right) . \tag{4.34}$$

In other words, the FLRW metric with $a(t) = t$ and $k = -1$ is nothing other than a patch of Minkowski spacetime in disguise!

For completness, and because it will come in handy in chapter 7, let us study the *Rindler* coordinates, which cover the complement of the region covered by the Milne coordinates, i.e. $(X^0)^2 - ||\vec{X}||^2 < 0$. Define the following parametrization of Minkowski spacetime:

$$\begin{aligned} X^0 &= r\ \sinh(t)\,, \\ X^1 &= r\ \cosh(t)\ \sin(\theta)\ \sin(\phi)\,, \\ X^2 &= r\ \cosh(t)\ \sin(\theta)\ \cos(\phi)\,, && (4.35) \\ X^3 &= r\ \cosh(t)\ \cos(\theta)\,. && (4.36) \end{aligned}$$

These coordinates slice up the spacetime with a one-paramteter family of one-sheeted Lorentzian hyperboloids, where the parameter is r, see figure 4.1. The metric (4.30) takes the following form:

$$ds^2 = dr^2 + r^2 \left(-dt^2 + \cosh^2(t)\, d\Omega_{S^2}{}^2\right) . \tag{4.37}$$

Although they are hyperboloids, the constant-r subspaces have Lorentzian signature and are positively curved. In fact, they are three-dimensional de Sitter spacetimes, as we will see next.

Having seen this familiar example, let us study de Sitter spacetime. It can be defined as a four-dimensional hyperboloid embedded in five-dimensional Minkowski spacetime:

$$-(X^0)^2 + (X^1)^2 + (X^2)^2 + (X^3)^2 + (X^4)^2 = \ell^2 \tag{4.38}$$

$$ds^2 = -d(X^0)^2 + d(X^1)^2 + d(X^2)^2 + d(X^3)^2 + d(X^4)^2\,, \tag{4.39}$$

where the first equation defines the hyperboloid, and the second defines the metric in the embedding space. The radius ℓ is related to the cosmological constant Λ in the Einstein equation as $\ell^2 = 3/\Lambda$. There are several coordinate systems that can be used to parametrize de Sitter spacetime, or at least a patch of it. In fact, it can be viewed as three different FLRW cosmologies with $k = 1, 0$, and -1 respectively. Let us start with the $k = 1$ form. Define the following coordinates:

$$\begin{aligned} X^0 &= \ell\ \sinh(t/\ell)\,, \\ X^1 &= \ell\ \cosh(t/\ell)\ \sin(\psi)\ \sin(\theta)\ \sin(\phi)\,, \\ X^2 &= \ell\ \cosh(t/\ell)\ \sin(\psi)\ \sin(\theta)\ \cos(\phi)\,, \\ X^3 &= \ell\ \cosh(t/\ell)\ \sin(\psi)\ \cos(\theta)\,, \\ X^4 &= \ell\ \cosh(t/\ell)\ \cos(\psi)\,. \end{aligned} \tag{4.40}$$

These coordinates solve the constraint (4.38) on the whole hyperboloid. The resulting four-dimensional metric is

$$ds^2 = -dt^2 + \ell^2\ \cosh^2(t/\ell)\, d\Omega^2_{S^3}\,. \tag{4.41}$$

This is called the de Sitter metric in *global* coordinates. It represents spacetime as a spacelike sphere that contracts from an infinite to a minimal radius ℓ (at $t = 0$), and then enters an eternal phase of accelerated expansion. The acceleration rate is constant at $\ddot{a}/a = 1$. This will cause causally connected spatial regions to become causally disconnected in the future. In other words, any two spatially separated observers will eventually become causally disconnected. To see this, we only need to look at null geodesics in de Sitter space. For simplicity, let us study a 'radial' geodesic emitted from the origin at time t_0:

$$-dt + \ell\ \cosh(t/\ell)\, d\psi = 0\,. \tag{4.42}$$

The solution is

$$\psi(t) = 2\Big(\arctan[\tanh(t/2\,\ell)] - \arctan[\tanh(t_0/2\,\ell)]\Big)\,. \tag{4.43}$$

If the light ray is emitted at time $t = 0$, it will asymptotically reach $\psi = \pi/2$ for $t \to \infty$. However, the later it is emitted the less it will travel as can be seen from the solution. This means that if we place a comoving observer at position $\psi = \epsilon$, it will at first be capable of receiving light rays emitted from the origin; however after a certain time (for $t > 2\,\text{arctanh}[\tan(\pi/4 - \epsilon)]$) it will be causally disconnected from the origin. This feature of de Sitter spacetime poses a serious problem in modern physics. One cannot define asymptotic states for a quantum field theory, or conservation laws for general relativity in the usual way.

Now, let us write down the $k = 0$ form of de Sitter spacetime. Once again, we implicitly define four-dimensional coordinates by solving the five-dimensional constraint (4.38):

$$\begin{aligned} X^0 + X^1 &= \ell\ \exp(t/\ell)\,, \\ X^i &= \ell\ \exp(t/\ell)\, x^i\,, \quad \text{for} \quad i = 2,3,4\,, \\ X^0 - X^1 &= \ell\ \exp(t/\ell)\left(\sum_{i=2}^{4}(x^i)^2 - \exp(-2\,t/\ell)\right), \end{aligned} \tag{4.44}$$

where the first equation defines a light-cone coordinate $\exp(t)$, the second equation defines cartesian coordinates x^i, and the third equation follows from the hyperboloid constraint (4.38). Note that the light cone coordinate is defined to be positive, which means that we are only covering half of the de Sitter manifold. Plugging this into (4.39) yields the following metric:

$$ds^2 = -dt^2 + \ell^2\ \exp(2\,t/\ell)\ \sum_i (dx^i)^2\,. \tag{4.45}$$

These are the de Sitter equivalent of Poincaré coordinates for anti-de Sitter spacetime. This form of de Sitter is the one used to model inflation because it has $k = 0$ and it is expanding for all t, unlike the global form (4.42). Finally, let us write down the $k = -1$ form. The trick is to put X^4 on the right-hand side of the constraint equation (4.38) and view the space as a one-parameter family of hyperboloids of radius $(X^4)^2 - \ell^2$, with the assumption that $|X^4| > \ell$:

$$\begin{aligned} X^4 &= \ell \cosh(t/\ell)\,, \\ X^0 &= \ell \sinh(t/\ell) \cosh(\psi)\,, \\ X^1 &= \ell \sinh(t/\ell) \sinh(\psi) \sin(\theta) \sin(\phi)\,, \\ X^2 &= \ell \sinh(t/\ell) \sinh(\psi) \sin(\theta) \cos(\phi)\,, \\ X^3 &= \ell \sinh(t/\ell) \sinh(\psi) \cos(\theta)\,, \end{aligned} \tag{4.46}$$

(4.47)

which yields the following metric:

$$ds^2 = -dt^2 + \ell^2 \sinh^2(t/\ell) \left(d\psi^2 + \sinh^2(\theta)\, d\Omega_{S^2}\right) . \tag{4.48}$$

The Anstatz for X^4 implies that this parametrization only cover half of the manifold. Note that this metric has a Big Bang singularity at $t = 0$.

Finally, we should briefly discuss anti-de Sitter spacetime or AdS. This is a solution to the Einstein equation with a negative cosmological constant. It can also be defined as a hyperboloid embedded in a higher dimensional spacetime, and many coordinate systems are available to cover it or at least partly cover it. However, AdS admits only one coordinate system such that its metric is in the FLRW form. The metric looks as follows:

$$ds^2 = -dt^2 + \ell^2 \sin^2(t/\ell) \left(d\psi^2 + \sinh^2(\theta)\, d\Omega_{S^2}\right) , \tag{4.49}$$

where ℓ is defined analogously to the de Sitter case. This is a $k = -1$ cosmology with a Big Bang singularity at $t = 0$ and a *big crunch* singularity at $t = \pi\,\ell$.

4.2 Physics of FLRW cosmologies

Having laid the foundations of cosmology we are ready to study the phenomena that drive the field of modern cosmology. The standard cosmology is a model of our universe that has been developed over decades by fitting observations from innumerably many experiments to theoretical models that rely upon the foundations of different fields such as general relativity, quantum field theory, thermodynamics, astrophysics, spectroscopy, etc.. Again, I would like to post my disclaimer here, and reiterate how extremely rich and complicated standard cosmology is, and that I in no way pretend to do justice to it. I will, however, try to give a condensed account of the history of our universe. Then, I will present three issues that arise in the standard cosmology, namely the *horizon problem*, the *flatness problem*, and the *relics problem*; and I will briefly explain the concept of inflation and show how it solves all three problems. I will then mention the presently observed acceleration of the universe, and finally, I will motivate the need for scalar cosmology models.

4.2.1 An ephemerally brief history of time

Let us start with an extremely brief history of the universe. In the beginning was the Big Bang. There are singularity theorems by Hawking and Penrose [74] that predict that any universe occupied by matter with $\rho > 0$ and $p > 0$ must have a Big Bang singularity. Since observations show that our early universe was mainly radiation dominated, the theorems would imply that our universe started with such a singularity. So what is a Big Bang singularity? A power law FLRW metric (4.29) provides us with a good metaphor for the Big Bang. At $t = 0$ the scale factor vanishes and the spatial section has 'zero size'. This is the 'beginning of time'. All the matter in the universe is condensed to a 'point', and thus ρ is really high. One must, however, realize that at the time of the Big Bang $t = 0$ the solution has a curvature singularity and the laws of General Relativity break down. No one knows, whether the singularity is a physical event, or a mere mathematical extrapolation from GR into uncharted territory. At this point a new theory is needed, namely one that can combine gravity and quantum mechanics. String theory is a strong candidate for this. For the time being, we must use GR within its regime of validity. This means that we cannot take $t = 0$ and $a = 0$ too literally. The standard cosmology is only meant to describe what happened after the first millisecond (or less) of the classically describable universe. So, although we cannot say that the universe 'started out' with 'zero size', or 'small' (unless $k = 1$, in which case a size can be defined), we can certainly say that it was occupied by very dense matter or radiation.

Since shortly after the Big Bang the universe was hot, dense and in thermal equilibrium, it started emitting light in every direction like a perfect blackbody. This radiation is observable today, especially its microwave component. This is the famous *Cosmic Microwave Background Radiation* or CMBR (or just CMB), which was almost accidentally discovered in 1965 by two radio astronomers, Arno Penzias and Robert Wilson. Its spectrum is so close to that of a perfect blackbody, that the CMBR is considered to be the strongest existing evidence of the Big Bang scenario. While the light was constantly scattering off of the rest of the matter constituents of the universe, the latter kept expanding. Expansion not only means that matter is driven apart at a rate proportional to the Hubble parameter, as we saw before, but it also means that the wavelengths of photons stretch. They get *redshifted*. Around 300,000 years after the Big Bang, the photons were so redshifted, that they no longer scattered off of particles. They decoupled, and simply went through everything. This is why the CMBR we observe today gives us such a perfectly undistorted picture of the universe as it was 300,000 years after the Big Bang. Before that, matter was constantly being ionized into plasma due to the constant scattering of photons. After that decoupling, the average temperature of the universe was low enough that atoms were able to form. This is called *recombination*. That is when galaxies and other structures started to form, leading to our present universe, at $t \sim 10^{10}$ years.

End of the schematic history of the universe.

4.2.2 Three problems

Like any great discovery in Physics, the CMBR has not only brought us answers, but also questions. It turns out that this radiation background has a remarkable property, it is almost perfectly isotropic. In any direction we look in the sky, this radiation has the same temperature to within 0.01%, about $2.7K$. Most of this variation by 0.01% is nowadays interpreted as proof that the Earth has a non-zero speed relative to the cosmological frame. We are not quite comoving. Taking this into account, the CMBR is ridiculously isotropic. This is puzzling from a causality point of view for the following reason: if one assumes that the universe has gone through a power-law expansion from the Big Bang until recently due to radiation and matter domination, then a calculation shows that the CMBR light that we see in the sky must have been emitted at recombination time ($t_{\rm CMBR} \sim 3 \times 10^5$ years) by points that could not have been in causal contact with each other. In other words, if we observe the light coming from two completely opposite directions in the sky, and we take the power-law expansion into account, we conclude that the two sources of light we are looking at were so distant from each other when they emitted it, that they had not had enough time to communicate since the Big Bang 300,000 years earlier. But why is the CMBR so isotropic, then? Why would causally disconnected regions of space emit such perfectly coordinated radiation? This is called the *horizon problem*.

The reader may find this paradox itself, paradoxical. One could ask the following question: "If the universe started with the Big Bang, and all spatial distances were (close to) zero in the beginning, then why couldn't all points in the universe simply have communicated back then, when they were so close to each other? How could 300,000 years not be enough for points that were at an initial distance of zero to communicate? As was pointed out before, no one knows, whether the universe really had 'zero size' in the past. The only trustworthy predictions of the standard cosmology are those regarding the history of the universe, beginning moments after the Big Bang. So, in this text, I will abandon the notion of a universe of 'zero size'. At most, one might say that a $k = 1$ model has an initially 'small' spatial section, in which case the above-mentioned paradox within the paradox becomes a valid one. Fortunately, it can be solved. Wald's book [75] discusses this very clearly. I will try, however, to explain this here. Let us start by defining the word *horizon*, or in this case *particle horizon*.

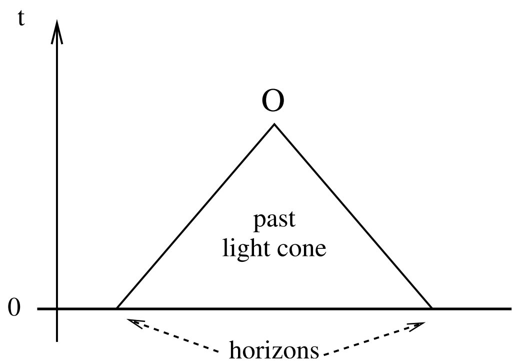


Figure 4.2: *The observer at event O can only see information emitted within the horizons. $t = 0$ marks the beginning of time.*

As observers, we can only see information coming from events that are within our past light-cones. We cannot, for instance, see something that happened one second ago in a galaxy that's three light-years away. When we look into the sky, the light that we see comes to us from the past. The farther the source is spatially, the older the information. But what if there was a 'beginning of time' such as in the Big Bang scenario? Then we would only be able to see information coming from a restricted area around our location. If spacetime were flat, but with a beginning of time, then only events that were within a distance d = (speed of light) $\times$ (age of the universe) of us at the time of emission could influence us. The spatial area that we can see is delimited by what is called a *particle horizon*. See figure 4.2. Now let us take a $k = 0$

FLRW metric with a power-law scale factor, and impose a cutoff minimal time t_i, which we will effectively treat as the beginning of time. Do horizons form? To see what happens, we must look at what light rays do. In comoving coordinates, a null geodesic has the following velocity:

$$\frac{dx}{dt} = \frac{1}{a(t)}\,, \tag{4.50}$$

where we use just one spatial axis for simplicity. This velocity is infinite at first, but decays more or less rapidly depending on the scale factor. We need to calculate how much comoving distance the geodesic can cover if it is emitted right after the Big Bang, at our cutoff time t_i, and observed at t_o:

$$\Delta x = x(t_o) - x(t_i) = \int_{\tau_i}^{t_o} \frac{dt}{a(t)}\,. \tag{4.51}$$

We can easily see that, for $a \propto t^{\alpha}$, this integral diverges as $t_i \to 0$, if $\alpha \geq 1$. In that case, there is a particle horizon, but the smaller t_i is, the bigger it gets. In other words, light coming from any point in the universe can reach the observer if it was emitted early enough. In the case where $\alpha < 1$, however, there is a particle horizon, and it is present even as $t_i \to 0$. Translating this into statements about matter

$$\alpha = \frac{2}{3\,(1+\omega)}\,, \tag{4.52}$$

we see that a radiation or matter dominated universe will generate horizons. Dark energy (i.e. $\omega < -1/3$), however, will generate horizons that are large at early time.

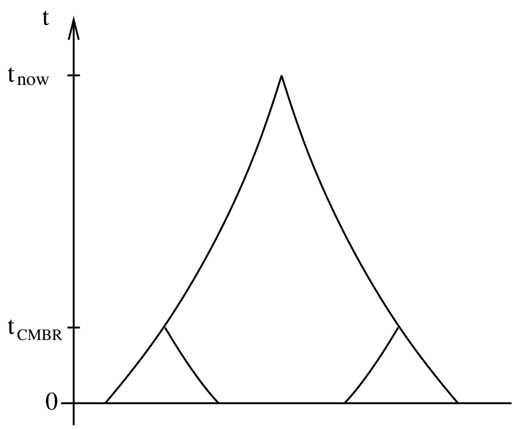


Figure 4.3: *The two sources of CMBR that we see today could not have been in causal contact.*

We are now ready to restate the horizon problem in the following oversimplified way:
At the present time, we can observe highly uniform CMBR rays. Choosing two widely separated CMBR sources in the sky will be separated by a comoving distance $\Delta_s \approx 4\,H_0$, where H_0 is the current Hubble parameter. The beams were emitted at t_{CMBR}. Assuming radiation domination ($a \propto t^{2/3}$), a null geodesic emitted at the Big Bang and observed at t_{CMBR} will travel a distance

$\Delta_l \approx 6 \times 10^{-2} H_0$. Hence we see that $\Delta_l \ll \Delta_s$, so the outermost sources of the CMBR could have never communicated, see figure 4.3. This is the horizon problem.

Another problem in the standard cosmology before inflation was known is the so-called *flatness problem*. Observations indicate that currently $\Omega \sim 1$ to a high degree of precision. However, in order for the universe to be so spatially flat in the present, it needs to have been extremely spatially flat from the get-go. This requires a high degree of fine-tuning that would have no apparent explanation. To understand how this comes about, let us start by rewriting the Friedmann equation as follows:

$$\Omega - 1 = \frac{k}{H^2 a^2} . \tag{4.53}$$

Differentiating this w.r.t. time this yields:

$$\dot{\Omega} = H (1 + 3\,\omega)\, \Omega\, (\Omega - 1) . \tag{4.54}$$

Note that, since $a(t)$ is a strictly monotonic function of t, we can treat the scale factor as a time parameter. This does not represent the proper time of any particular observer, but it allows us to look at the equations from the point of view of dynamical systems. We will do extensively in the next two chapters. Using $dt = da/H$ we rewrite the evolution equation for Ω as follows:

$$\frac{d\Omega}{da} = (1 + 3\,\omega)\, \frac{\Omega\, (\Omega - 1)}{a} . \tag{4.55}$$

We immediately see that $\Omega = 1$ is a critical point of this system (4.55), i.e. a point where $d\Omega/da = 0$. However, *assuming* the universe is dominated by ordinary matter or radiation (i.e. $\omega > -1/3$), this critical point is not an attractor, but a repellor or unstable critical point:

$$\frac{d}{d\Omega} \left(\frac{d\Omega}{da} \right) \bigg|_{\Omega=1} > 0 . \tag{4.56}$$

This means that, in order for Ω to be one today, it must have been incredibly close to one in the early universe. In fact, by looking at (4.55) we see that any slight deviation from the value one is magnified by the small scale factor (early universe) in the denominator. The fine tuning required to keep the rate of change of Ω small enough so that Ω is close to one today cannot be explained without inflation.

Finally, there is one more problem that arises in the standard cosmology, which is also solved by inflation. It is called the *unwanted relics problem*. I will not treat this problem in any detail whatsoever, but will merely state it. In spontaneously broken gauge theories, topologically non-trivial objects such as monopoles, strings, or textures naturally arise. The gauge theory that describes the matter in the universe is a GUT (Grand Unified Theory), and it has a gauge group, which is spontaneously broken to the standard model gauge group $SU(3) \times SU(2) \times U(1)$. It is possible to predict the density of monopoles that should be present in our universe today, by standard calculations using the assumptions about cosmology that we have been using so far. The result turns out to be far too big. The abundant number of monopoles as predicted by the standard cosmology is very generous, however, not one monopole has ever been observed.

4.2.3 Inflation saves the day

Inflation is a scenario for the evolution of the universe, which was created in the 80's [76–78] to solve a number of problems, among which are the three that were mentioned in the previous section. The idea is to have the universe go through a period of accelerated expansion (i.e. $\ddot{a} > 0$) starting $10^{-12}s$ after the Big Bang, and lasting long enough for the scale factor to increase by a factor of 10^{60}. Let us start by looking at how this could solve the horizon problem.

As was mentioned in the previous subsection, solving the horizon problem consists in explaining how regions that seem causally disconnected at $t = t_{\mathrm{CMBR}}$ under the assumption of power-law expansion could have actually been in causal contact at earlier times. As shown earlier, if the scale factor is a power law function with exponent $\alpha < 1$, then there is a finite horizon, no matter how early we take time to begin. However, if $1/a(t)$ blows up faster than $1/t$ for $t \to 0$, then the horizon can be made large (in comoving coordinates). By choosing a function that blows up fast enough, we can enlarge the horizons of the CMBR sources such that they will include each other, thereby solving the horizon problem. Note that this applied not only to power-law solutions with $\alpha \geq 1$, but also to the de Sitter solution, $a \propto \exp(H\,t)$. As mentioned before, in terms of matter or energy content, this requires $\omega < -1/3$. This can be a cosmological constant or some other form of dark energy.

Another way to see how this solves the problem is the following: take two comoving points separated initially by a distance $s = a(t_i)\,\Delta x$. Their proper relative speed is $\dot{s} = \dot{a}\,\Delta x$. If $\ddot{a} > 0$, this relative speed will increase with time, eventually exceeding the proper speed of light, which is

$$a\,\frac{dx}{dt} = a\,\frac{1}{a} = 1\,. \tag{4.57}$$

So regions that are initially causally connected can become causally disconnected by moving away from each other faster than the speed of light.

The flatness problem is also solved by inflation. Intuitively speaking, the period of accelerated expansion blows up small regions of space into huge ones in a short time, thereby flattening out any initial spatial curvature. This explains why the present universe is spatially flat without resorting to fine-tuning at early times. There are two ways to see how this works mathematically:

From the Friedmann equation, which I rewrite for the reader's convenience,

$$\Omega - 1 = \frac{k}{H^2\,a^2}\,, \tag{4.58}$$

we see that the right-hand side decreases with time if $\ddot{a} > 1$, leading to a spatially flat universe, even if the spatial curvature k/a was initially huge. We can also understand this in the language of critical points. From the acceleration equation (4.19) we read off that an accelerating universe requires $\omega < -1/3$. Analyzing (4.55) as we previously did, with this assumption about ω, we see that $\Omega = 1$ is now a stable critical point.

Finally, inflation also solves the problem of unwanted relics. The precise argument is beyond the scope of this chapter, so I will just state the intuitive one. Basically, inflation blows up small regions in space into huge ones, however the amount of monopoles and other topological relics does not increase. The consequences is that the latter are diluted in our universe, which provides us with a plausible explanation for why we have not detected them yet.

4.2.4 Present day acceleration

Another important piece of information about the physics of cosmology concerns the present. By measuring the redshift of light coming from supernovae, two independent teams [79, 80] have concluded that our universe is currently undergoing a period of accelerated expansion. From the acceleration equation (4.19), we see that this implies the presence of dark energy. In fact, these measurements imply that dark energy is the dominant form of energy in the universe today, providing us with the estimate $\Omega_\Lambda \sim 0.7$, mentioned in (4.29).

In this section, we described the history of the universe from moments after the Big Bang until the present day. We have seen that in order to solve the horizon, flatness, and relics problems, the early universe must have gone through a period of inflation lasting long enough to generate 60 e-foldings (i.e. $\log(a_{\text{now}}/a_i) = 60$). Inflation actually also solves a number of problems that I have not even mentioned here. Therefore, inflation is definitely a necessary scenario for modern cosmology. However, it is a 'passing the buck' solution to those problems. It merely merges several problems into one big problem: What drives inflation? Even though we know that dark energy is required for it, there is no known mechanism in physics to *derive* inflation from a fundamental theory. Similarly, there is no *derivation* of the current period of acceleration we are going through. To repeat the quote by Quevedo, inflation is "a scenario in search of an underlying theory." So is present acceleration. In recent years, new hope has arisen that string theory may be used to derive realistic cosmological scenarios. Especially, the latest very precise measurements of CMBR anisotropies have given theorists the hope of finding observational signatures of stringy or transplanckian physics. On one hand cosmology poses a challenge for string theory to come up with a mechanism to drive inflation and present day acceleration, on the other hand, it may provide string theorists with their first lab in which to test string theory ideas.

4.3 New challenges lead to new ideas

If string theory truly is the *theory of everything*, and especially if it is a theory of quantum gravity, then it must ultimately explain the Big Bang, inflation, and current acceleration. In this section we will be looking at some candidate mechanisms by means of which string theory might induce those two cosmological events. I will begin by introducing a new form of dark energy as a possible source for acceleration: the scalar field. Then, I will briefly introduce how gravity-scalar models with accelerating cosmological solutions can arise from string or M-theory. Consider this as an introduction for the next two chapters, which will be based on two articles about scalar cosmologies and their possible string/M-theory origins.

4.3.1 Scalar models for cosmology

As we pointed out before, in order to have accelerated expansion, be it for inflation or present day acceleration, we must have a perfect fluid with $\omega < -1/3$ in the universe. Having $\omega = -1$, a positive cosmological constant will do. It will source a de Sitter spacetime. However, it does have some drawbacks: being a constant by definition, it is non-dynamical. This means that the universe would be in a state of eternal inflation at a constant rate of acceleration, which is not quite consistent with observations. A more flexible and more interesting approach would be to have a form of dark energy that mimics a cosmological constant and is yet dynamical. This has two advantages: firstly, it could induce a de Sitter-like universe with a slowly varying acceleration rate, which would be more consistent with observations of current acceleration. Secondly, it would in principle allow for a dynamical start and end of inflation and current acceleration, and also for a dynamical resolution of the cosmic coincidence problem, which is more appealing from a theoretician's point of view.

Let us write down a simple gravity-scalar model, namely gravity with one scalar field and some potential for it:

$$\mathcal{L} = \sqrt{-g}\,\left(R - \tfrac{1}{2}\,(\partial\phi)^2 - V(\phi)\right) . \tag{4.59}$$

The equations of motion for a $k = 0$ FLRW Ansatz are the following:

$$H^2 = \tfrac{1}{12}\,\dot{\phi}^2 + \tfrac{1}{6}\,V\,, \tag{4.60}$$

$$\frac{\ddot{a}}{a} = \tfrac{1}{6}\,\left(-\dot{\phi}^2 + V\right)\,, \tag{4.61}$$

$$\ddot{\phi} + 3\,H\,\dot{\phi} + \frac{\partial V}{\partial\phi} = 0\,, \tag{4.62}$$

where we recognize the first two equations as the Friedmann and acceleration equations, respectively, and the third one is the equation of motion of the scalar field. To be consistent with homogeneity, we have assumed that ϕ depends only on time. Comparing this to (4.18) and (4.19), we see that

$$\rho = \frac{1}{16\,\pi\,G}\,\left(\tfrac{1}{2}\,\dot{\phi}^2 + V\right)\,, \tag{4.63}$$

$$p = \frac{1}{16\,\pi\,G}\,\left(\tfrac{1}{2}\,\dot{\phi}^2 - V\right)\,. \tag{4.64}$$

So, if ϕ varies slowly in time, its equation of motion approaches that of a cosmological constant, i.e. $\omega \sim -1$. We also see from the acceleration equation (4.61) that V acts in favor of acceleration like a cosmological constant, and the kinetic energy acts against it. This is why in scalar models for inflation such as *chaotic inflation*, one requires that the field be slowly varying, i.e. $\dot{\phi} \ll 1$, by restricting the form of the potential. However, a realistic model for inflation must have an inflationary period of at least 60 e-foldings. A scalar field will naturally roll down its potential until it reaches a minimum, and its kinetic energy will only increase in the meantime, leading to a non-accelerating or even decelerating cosmology. Therefore, in order to prevent a premature end of inflation one must also require that $\ddot{\phi} \ll 1$. These two conditions, $\dot{\phi} \ll 1$ and $\ddot{\phi} \ll 1$ are called the *slow roll* conditions. Of course, in a specific model, one usually parametrizes these constraints to obtain controlled results.

Introducing the scalar field allows for cosmologies that are more complicated than just power-law or de Sitter solutions. Because it is dynamical, it can source solutions that interpolate in time between those two basic solutions. Cosmological solutions that interpolate in time between two non-accecelerating regimes, but are separated by one or several periods of transient acceleration are of special interest. We will see a specific example of this in the next section, and in the next two chapters we will be looking at more general examples where we introduce several scalar fields with intricate potentials that couple them to each other.

4.3.2 Acceleration from string/M-theory

In principle one can obtain all kinds of interesting geometries to model inflation and current acceleration from scalar field models by having several scalar fields and the right potential, as we will see in the next two chapters. However, even if one can write down such a model the question remains: where do these fields and their potential come from? Often one refers to such scalar fields as *inflatons* and to their potentials as *quintessence*, meaning they are a fifth force in nature that drives acceleration. However, as string theorists, we do not like to invoke new forces unless we can derive them from a unified theory. In the past few years, string theorists have made numerous attempts to derive scalar cosmology models by dimensionally reducing 10-dimensional supergravities and making appropriate truncations leaving only scalar fields and scalar potentials in the four-dimensional spacetime. In chapter 6, we will look at what happens when one reduces supergravities on three-dimensional *group manifolds*. However, before jumping into that, I will attempt to give a brief review of what happens when one considers simpler schemes, such as reducing over Einstein spaces[2].

The standard torroidal Kaluza-Klein reduction scheme provides us with an easy way of going from ten dimensions to four *and* generating scalar fields (i.e. Kaluza-Klein modes) with potentials. However, the potentials it yields will not generate an accelerating four-dimensional universal. To make things worse, there is a *no-go* theorem [81, 82] that essentially states that compactifications of ten or eleven dimensional supergravities of string/M-theory over compact, non-singular, spaces without boundaries and with time-independent volume[3] never lead to accelerating universes. To circumvent the theorem, one must allow for time-dependent volume of the internal space. P.K. Townsend and M. Wohlfarth [83] showed that reducing gravity over a six or seven-dimensional hyperboloid with time-dependent volume yields a universe with a limited period of acceleration. The solution interpolates in time between two decelerating power-law periods at $t \to 0$ and $t \to \infty$, which are joined by an accelerating epoch. The Ansatz in $4 + n$ dimensions has the following form:

$$ds^2 = \delta^{-n}(t)\, ds_E^2 + \delta^2(t)\, dH_n^2\,, \tag{4.65}$$

where ds_E^2 is the four-dimensional cosmological spacetime that will result in Einstein frame after the reduction, dH_n is the n-dimensional hyperbolic space, and $\delta(t)$ is the *warp factor*, which will act as a time-dependent 'volume' of the internal space. The dimension n of the internal space is left arbitrary, but for string/M-theory we need $n = 6, 7$. I will not write down the actual solutions for $\delta(t)$ and ds_E^2, for I want to stress the qualitative information. The $(4+n)$-dimensional Ansatz is itself flat, i.e. it is Minkowski spacetime with some identifications that do not affect curvature. However, the reduction Ansatz we have chosen, yields a non-trivial four-dimensional spacetime with interpolating behavior. In the early universe it has $a \sim t^{1/3}$; in the future it has $a \sim t^{n/(n+2)}$; and in between it has an epoch of transient acceleration. This is in principle what we are looking for, as this scenario has its own mechanism to begin and end inflation. Unfortunately, the acceleration period generated by this scheme is not long enough

[2]An Einstein space is a manifold with a metric that solves the Einstein equations *in vacuo* or in the presence of a cosmological constant. As a consequence of that, it has the highest possible degree of symmetry.

[3]You may wonder what I mean by 'volume' if the internal space is hyperbolic. In this case one must always make the space compact by topological identifications. Otherwise, one must face the undesirable physical consequences of a so-called *non-compactification*.

to yield the so much needed 60 e-foldings of inflation. But the result may still apply to current acceleration.

The Townsend-Wohlfarth solution turns out to be a special case of a larger class of supergravity solutions called *S-branes*, found in [84]. These are essentially solutions of supergravity, which look like p-brane solutions, except that time is *transverse* to their world-volume as opposed to being in it. These solutions are sourced by the dilaton and some antisymmetric tensor, just like p-brane solutions:

$$\mathcal{L} = R - \frac{1}{2}\,(\partial\phi)^2 - \frac{1}{2\,(p+2)!}\,e^{a_p\,\phi}\,F_{p+2}^2\,, \tag{4.66}$$

where F_{p+2} is the field-strength, and a_p is determined by the supergravity in question. This differs from the previous Ansatz in that the latter was a solution to Einstein's equation *in vacuo*, whereas the S-brane is carried by the dilaton and has a flux from the $p + 2$-form field-strength. The Ansatz for the metric is similar to the previous one, except that the internal space no longer needs to be hyperbolic; it can be flat or spherical. The Ansatz for an SD2-branes looks roughly as follows:

$$ds^2 = -f(t)^2\,dt^2 + g(t)^2\,d\mathbf{x}_3^2 + h(t)^2\,d\Sigma_{k,6}^2\,, \tag{4.67}$$

where the three boldfaced spatial coordinates correspond to our space, and to the world-volume of the SD2-brane, the six-dimensional internal space can now be positively curved, flat, or negatively curved ($k = 1, 0, -1$ respectively), and $f(t), g(t)$ and $h(t)$ are determined by the equations of motion. This solution is no longer flat in ten dimensions, since it now solves the Einstein equations with RR flux turned on, but interpolates in time between a flat metric and a horizon-like geometry. In four dimensions, it yields an interpolating solution with a transient accelerating epoch regardless of the kind of internal space we pick (i.e. $k = 1, 0, -1$). For a more detailed review on the subject of S-branes and their status, the reader is referred to [85].

The schemes I have mentioned so far are all based on the assumption that the supergravity approximation is a valid one, allowing one to treat string theory as field theory. However, this assumption is not necessarily justified. One uses it only because it is very difficult to deal with the full string theory. For instance, in a scenario where the dilaton grows large over time, string perturbation theory will break down. That is why attempts are being made to take non-perturbative string theory effects into account in compactification schemes. Another problem posed by these compactifications is that the volume and shape of the internal space, being dynamical by construction, are not always stable. For instance, in many solutions, the volume will tend to blow up in time. This is known as *spontaneous decompactification*. If one takes such models seriously, then one should expect to be able to observe these extra dimensions in the present, or assume that we live in a special moment in the history of the universe, when the extra dimensions happen to be small. These compactification schemes should, therefore, not be regarded as phenomenologically realistic models, but merely as evidence that demonstrates that it is possible to circumvent the Maldacena-Nuñez no-go theorem [82].

Currently, string theorists are trying to create realistic models that can stabilize all of the *moduli* of the internal compactification manifold. A couple of years ago, the authors of [33] came up with a string compactification scenario that exploits non-perturbative string theory effects to stabilize the internal moduli. The idea relies on non-perturbative instanton effects induced by wrapping a Euclidean D3-brane around a 4-cycle of the internal Calabi-Yau space.

The authors of the paper, however, did not find an explicit choice for the required Calabi-Yau space to carry out this idea. String theorists have only recently been able to write down concrete realizations of this scenario. For instance, while this thesis was being written, an article was published [86], in which not only the moduli stabilization problem was dealt with, but also the problem of breaking supersymmetry softly for particle phenomenological purposes.

All of the schemes to obtain acceleration from string/M-theory that I have mentioned so far have one thing in common: from the four-dimensional point of view, they all reduce to an effective field theory with Einstein gravity and scalar fields with potentials. This is true even for models that take non-perturbative string theory effects into account. Therefore, although one would like to be able to derive the ultimate string theory mechanism or scenario that leads to inflation and present day acceleration right away, it is useful and wise to also study which four-dimensional scalar models are capable of driving those two cosmological phenomena at all. After all, most of the conceivable reduction schemes will reduce to four-dimensional scalar-gravity field theories. Should one find a class of models that drive a realistic cosmology, one could then investigate how to obtain it from string theory. In the next two chapters, we will be doing a bit of both. We will study scalar-gravity models with exponential potentials in general, but will also pay attention to potentials obtained from some specific dimensional reduction schemes.

Chapter 5

Scalar Cosmologies I: A simple case

5.1 Introduction

The discovery that the universe may currently be in a phase of accelerated expansion [79, 80] has led to strong interest in finding de Sitter solutions or more general accelerating cosmologies from M-theory, see [33, 83, 87–97] and references therein.

A simple way to study accelerating cosmologies is to consider models containing just gravity and a number of scalars with a potential. This method has a long history and has resulted in models for inflation [98], describing the early universe, and for quintessence [99], describing the present universe. The potentials for the scalar fields give rise to a small effective cosmological constant. Multi-exponential potentials comprise a specific class of potentials, which have been frequently studied, and these are of interest for two reasons: first, they can arise from M-theory in many ways; e.g. via compactifications on product spaces possibly with fluxes [100–103], and second, the equations of motion can be written as an autonomous system. This approach allows for an algebraic determination of power-law and de Sitter solutions, which are viewed as critical points that can correspond to early- and late-time asymptotics of general solutions. Many authors have made use of this fact, see [95, 104–110] and references therein.

The purpose of this chapter is to investigate the possibility of transient acceleration for the class of cosmologies whose solutions are described by a metric and N scalars, with a scalar potential given by a single exponential. The consequence of this is that, effectively, the scalar potential depends on only one scalar. All other $N-1$ scalars are represented by their kinetic terms only. Since the metric cannot distinguish between these different $N-1$ scalars, there is no qualitative difference between the $N = 2$ scalar cosmology and the $N > 2$ scalar cosmologies. We therefore only consider the one-scalar ($N = 1$) and two-scalar ($N = 2$) cosmologies. We will be studying these models purely from the four-dimensional point of view, without reference to possible higher dimensional origins. This chapter can be considered as a warm up for the next chapter, where we will study scalar cosmologies with multi-exponential potentials. In that case, things will be much more complicated, as there it will no longer be possible to reduce a multi-scalar system into a 2-scalar model.

The cosmological solutions discussed in this paper have been given sometime ago [111,112].

The fact that these cosmologies, for particular cases at least, exhibit a period of acceleration, was noted recently in [83] where a specific class of solutions was obtained by compactification over a compact hyperbolic space (for earlier discussions, see [90, 101, 112–114]). The relation with S-branes was subsequently noted in [91, 92, 97] (for general literature on S-brane solutions, see [84, 94, 100–102, 114–121]).

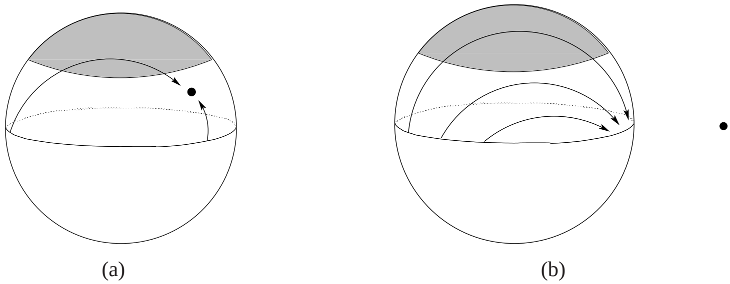


Figure 5.1: *Each cosmological solution is represented by a curve on the sphere. In figure (a) "Rome", represented by the dot, is on the sphere and each curve is directed from the equator towards "Rome", which corresponds to a power-law solution to the equations of motion. In figure (b) "Rome" is not on the sphere and each curve, again being directed towards "Rome", begins and ends on the equator. In this case "Rome" is not a solution. The accelerated expansion of the solution occurs whenever the curve lies within the "arctic circle". This region is shown by the shaded area.*

In this work we will discuss systematically the accelerating phases of all 2-scalar cosmologies with a single exponential potential by associating to each solution a trajectory on a 2-sphere. It turns out that all trajectories have the property that, when projected onto the equatorial plane, they reduce to straight lines which are directed towards a point that we will call "Rome". Depending on the specific dilaton coupling of the potential, this point can be either on the sphere or not. In the former case, it corresponds to a power-law solution for the scale factor, whereas in the latter case, it is not a solution. We find that the accelerating phase of a solution is represented by the part of the trajectory that lies within the "arctic circle" on the sphere, see figure 5.1. This enables us to calculate the expansion factors in a straightforward way for each of the solutions.

This chapter is based on a collaboration with E. Bergshoeff, U. Gran, M. Nielsen, and D. Roest, entitled *Transient quintessence from group manifold reductions or how all roads lead to Rome* [95]. It is organized as follows: in sections 5.2-5.4 we present, under the assumptions stated, the most general N-scalar accelerating cosmology in 4 dimensions. The accelerating phases of these cosmologies are discussed in section 5.5. Their equations of state and the one-scalar truncations are discussed in sections 5.6 and 5.7, respectively.

5.2 Setup: Lagrangian and Ansatz

Our starting point is gravity coupled to N scalars [122] which we denote by $(\varphi, \vec{\phi})$. We assume that the scalar potential consists of a single exponential term:

$$\mathcal{L} = \sqrt{-g}\left[R - \tfrac{1}{2}(\partial\varphi)^2 - \tfrac{1}{2}(\partial\vec{\phi})^2 - V(\varphi, \vec{\phi})\right], \qquad V(\varphi, \vec{\phi}) = \Lambda \, \exp(-\alpha\varphi - \vec{\beta}\cdot\vec{\phi}), \tag{5.1}$$

where we restrict[1] to $\Lambda > 0$. To characterize the potential we introduce the following parameter:

$$\Delta \equiv \alpha^2 + |\vec{\beta}|^2 - \frac{2\,(D-1)}{(D-2)} = \alpha^2 + |\vec{\beta}|^2 - 3 \quad \text{for} \quad D = 4. \tag{5.2}$$

This parameter, first introduced in [52], is invariant under toroidal reductions.

The kinetic terms of the dilatons are invariant under $SO(N)$-rotations of $(\varphi, \vec{\phi})$. However, in the scalar potential the coefficients α and $\vec{\beta}$ single out one direction in N-dimensional space. Therefore the Lagrangian (5.1) is only invariant under $SO(N-1)$. The remaining generators of $SO(N)$ can be used to set $\vec{\beta} = 0$, in which case only the scalar φ appears in the scalar potential. Such a choice of basis leaves Δ invariant.

Motivated by observational evidence, we choose a flat FLRW Ansatz. This basically means a spatially flat metric that can only contain time-dependent functions. One can always perform a reparametrization of time to bring the metric to the following form:

$$ds^2 = -a(u)^{2\delta}\, du^2 + a(u)^2\, dx_3^2\,, \tag{5.3}$$

for some δ. In this paper we will choose δ as follows[2]:

$$\text{Cosmic time:} \qquad \delta = 0\,, \qquad u = \tau\,, \quad \frac{da}{d\tau} = \dot{a}\,, \tag{5.4}$$

$$\text{Non-cosmic time:} \quad \delta = 3\,, \qquad u = t\,, \quad \frac{da}{dt} = a'\,. \tag{5.5}$$

As a part of the Ansatz, we also assume:

$$\varphi = \varphi(u)\,, \qquad \vec{\phi} = \vec{\phi}(u)\,. \tag{5.6}$$

For this Ansatz one can reduce the $N-1$ scalars $\vec{\phi}$ that do not appear in the potential to one scalar by using their field equations as follows:

$$\frac{d^2\vec{\phi}}{du^2} = (\delta - 3)\,\frac{d\,\log a}{du}\,\frac{d\vec{\phi}}{du} \qquad \Rightarrow \qquad \frac{d\vec{\phi}}{du} = \vec{c}\,a^{\delta-3}\,, \tag{5.7}$$

where $\vec{c}$ is some constant vector. The only influence of the $N-1$ scalars comes from their total kinetic term:

$$\left|\frac{d\vec{\phi}}{du}\right|^2 = |\vec{c}|^2\, a^{2\delta-6}\,. \tag{5.8}$$

[1] We make this choice in order to obtain dark energy and therefore accelerating solutions.

[2] The non-cosmic time corresponds to the gauge in which the lapse function $N \equiv \sqrt{-g_{tt}}$ is equal to the square root of the determinant γ of the spatial metric, i.e. $N = \sqrt{\gamma}$, whereas cosmic time corresponds to $N = 1$. We thank Marc Henneaux for a discussion on this point.

Therefore, from the metric point of view, there is no difference between $N = 2$ and $N > 2$ scalars (under the restriction of a single exponential potential). The truncation of the system (5.1) to one scalar corresponds to setting $\vec{c} = 0$.

To summarize, we will be using the following Lagrangian:

$$\mathcal{L} = \sqrt{-g}\left[R - \tfrac{1}{2}(\partial\varphi)^2 - \tfrac{1}{2}(\partial\phi)^2 - V(\varphi)\right], \qquad V(\varphi) = \Lambda\ \exp(-\alpha\varphi)\,, \tag{5.9}$$

with $\Lambda > 0$ and we choose the convention $\alpha \geq 0$. From now on we will use $\Delta = \alpha^2 - 3$ instead of α.

In the next two subsections we will first discuss the critical points corresponding to the system (5.9) and then the solutions that interpolate between these critical points. We will use cosmic time (5.4) when discussing the critical points in section 5.3 and non-cosmic time (5.5) when dealing with the interpolating solutions in section 5.4.

5.3 Critical points

It is convenient to choose a basis for the fields, such that they parametrize a 2-sphere. In this basis, we will be able to regard our system as an autonomous one, and we will find that all constant configurations (critical points) correspond to power-law solutions for the scale factor $a(\tau) \sim \tau^p$ for some p. By studying the stability of these critical points [104,107] one can deduce that there exist interpolating solutions which tend to these points in the far past or the distant future. We will actually be able to draw these interpolating solutions without having to do any stability analysis.

We begin by choosing the flat FLRW Ansatz (5.3) in cosmic time:

$$ds^2 = -d\tau^2 + a(\tau)^2\,(dx^2 + dy^2 + dz^2)\,. \tag{5.10}$$

The Einstein equations for the system (5.9) with this Ansatz become:

$$H^2 = \tfrac{1}{12}\,(\dot{\varphi}^2 + \dot{\phi}^2) + \tfrac{1}{6}\,V\,, \tag{5.11}$$

$$\dot{H} = -\tfrac{1}{4}\,(\dot{\varphi}^2 + \dot{\phi}^2)\,, \tag{5.12}$$

where $H \equiv \dot{a}/a$ is the Hubble parameter and the dot denotes differentiation w.r.t. τ. Equations (5.11) and (5.12) are usually referred to as the Friedmann equation and the acceleration equation, respectively. The scalar equations are:

$$\ddot{\varphi} = -3\,H\,\dot{\varphi} + \sqrt{\Delta + 3}\ V\,, \qquad \ddot{\phi} = -3\,H\,\dot{\phi}\,. \tag{5.13}$$

We define the following three variables:

$$x = \frac{\dot{\varphi}}{\sqrt{12}\,H}\,, \qquad y = \frac{\dot{\phi}}{\sqrt{12}\,H}\,, \qquad z = \frac{\sqrt{V}}{\sqrt{6}\,H}\,. \tag{5.14}$$

In these variables the Friedmann equation (5.11) becomes the defining equation of a 2-sphere [107, 123]:

$$x^2 + y^2 + z^2 = 1\,. \tag{5.15}$$

This means that we can think of solutions as points or trajectories on a globe. It turns out that cosmological solutions are either eternally expanding (i.e. $H > 0$) or eternally contracting ($H < 0$), but cannot have an expanding phase and then a contracting phase (or vice-versa). Since we are only interested in expanding universes, we will only be concerned with the upper hemisphere (i.e. $z > 0$). In terms of x and y the scalar equations become:

$$\frac{\dot{x}}{H} = -3\,z^2\,(x - \sqrt{1 + \Delta/3})\,, \tag{5.16}$$

$$\frac{\dot{y}}{H} = -3\,z^2\,y\,. \tag{5.17}$$

We can rewrite the acceleration equation (5.12) as follows:

$$\frac{\dot{H}}{H^2} = -3\,(x^2 + y^2)\,. \tag{5.18}$$

If we now solve for the critical points ($\dot{x} = 0, \dot{y} = 0$), we can then integrate (5.18) twice and obtain the following power-law solutions for $a(\tau)$ [122]:

$$a(\tau) \sim \tau^p\,, \quad \text{where} \quad p = \frac{1}{3\,(x_c^2 + y_c^2)}\,, \tag{5.19}$$

and the following solutions for the scalars:

$$\varphi = \sqrt{12}\,p\,x_c\,\log(\tau) + \text{constant}\,. \tag{5.20}$$

We thus find the following critical points:

- **Equator**:

$$z = 0, \quad x^2 + y^2 = 1. \tag{5.21}$$

Every point on the equator of the sphere is a critical point with power-law behaviour $a \sim \tau^{1/3}$.

- **“Rome"**:

$$x = \sqrt{1 + \Delta/3}\,, \quad y = 0\,, \quad z = \sqrt{-\Delta/3}\,. \tag{5.22}$$

This critical point yields a power-law behaviour of the form (we ignore here irrelevant constants that rescale time)

$$a \sim \tau^{1/(\Delta+3)} \text{ for } -3 < \Delta < 0\,, \qquad a \sim e^{\tau} \text{ for } \Delta = -3. \tag{5.23}$$

Note that the greater Δ is, the further “Rome" gets pushed towards the equator, and for $\Delta = 0$ it is on the equator.

Although the equatorial points (a.k.a. kinetic-dominated solutions) do solve (5.15)-(5.18) as critical points, they are not proper solutions of (5.11)-(5.13) in terms of the fundamental fields, since $z = 0$ would imply that $V = 0$, which is impossible for $\Lambda \neq 0$ unless φ is infinite at all times. However, these points will be interesting to us, as they will provide information about the asymptotics of the interpolating solutions.

In contrast to the equator, the "Rome" critical point is a physically acceptable solution of the system, provided it is well defined on the globe (i.e. $\Delta < 0$). In the case where $\Delta = -3$ it becomes De Sitter (i.e. $a \sim e^{\tau}$), as one would expect, since $V = \Lambda$.

Besides these critical points there are other solutions, which are not points but rather trajectories. In fact, we can already determine their shapes. Dividing (5.16) and (5.17) we obtain the following:

$$\frac{dy}{dx} = \frac{y}{x - \sqrt{1 + \Delta/3}} \,. \tag{5.24}$$

Integrating this we get the following relation between x and y:

$$y = C\,(x - \sqrt{1 + \Delta/3})\,, \tag{5.25}$$

where C is an arbitrary constant[3]. This relation tells us that if we project the upper hemisphere onto the equatorial plane, in other words, if we view the sphere from above, any solution to the equations of motion must trace out a straight line that lies within the circle defined by $x^2 + y^2 = 1$ and has a y-intercept at ($x = \sqrt{1 + \Delta/3}, y = 0$). From now on, we will refer to that point as "Rome"[4]. Notice that all lines intersect at "Rome" independently of whether it is on the globe ($\Delta < 0$), right on the equator ($\Delta = 0$) or off the globe ($\Delta > 0$). These lines can only have critical points as end-points. So each line is a solution, which interpolates between two power-law solutions. In a similar, yet physically inequivalent context, such a line was found in [105].

Now that we know the shapes of the trajectories, let us figure out their time-orientations. By looking at (5.16) we realize that the time derivative of x is positive when $x < \sqrt{1 + \Delta/3}$ and negative when $x > \sqrt{1 + \Delta/3}$. This tells us that *all roads lead to Rome*. Figure 5.2 illustrates this for the cases where "Rome" is off the globe, right on the equator or on the globe.

One can also determine the orientations of the trajectories by analysing the stability of the critical points. One will find that whenever "Rome" is on the globe (i.e. $\Delta = 0$ and $\Delta < 0$), it is stable (i.e. an attractor), and the points on the equator are all unstable (i.e. repellers), except for "Rome" when $\Delta = 0$. In the case where "Rome" is off the globe (i.e. $\Delta > 0$), the equator splits up into a repelling and an attracting region. The attracting region turns out to be the portion of the equator that can "see" "Rome". In other words, any point on the equator that can be joined to "Rome" by a straight line such that the line does not intersect the equator again before reaching "Rome" is attracting. To summarize, for $\Delta > 0$, all points on the equator with $x > \sqrt{3/(\Delta + 3)}$ are attracting, and the rest are repelling. In the first illustration of figure 5.2, the attracting portion of the equator is depicted by the thick arc.

[3] Since C is finite one might think that this excludes the line defined by $x = \sqrt{1 + \Delta/3}$. However, that line can be obtained by taking the inverse of (5.24) and solving for x as a function of y.

[4] Note that we have extended our definition of "Rome": only if "Rome" is on the globe ($\Delta < 0$) is it equal to the critical point discussed before.

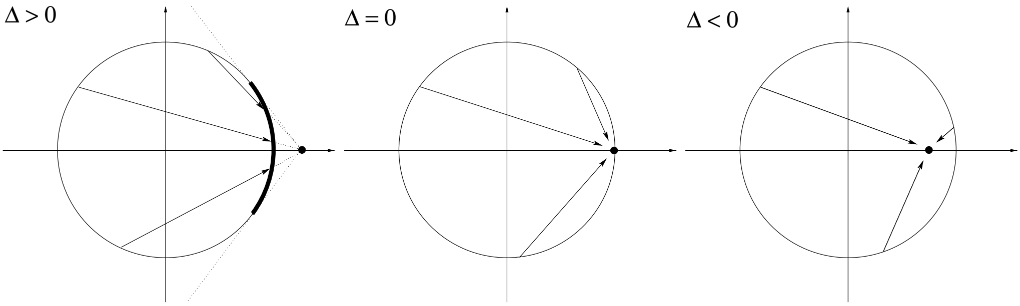


Figure 5.2: *The solutions represented as straight lines in the (x, y)-plane for $\Delta > 0$ where "Rome" is not on the sphere, $\Delta = 0$ where "Rome" is on the equator and $-3 \leq \Delta < 0$ where "Rome" is on the sphere. The thick arc in the left figure represents the attracting portion of the equator [124].*

5.4 Interpolating solutions

To solve the equations of motion, it is convenient to use the FLRW Ansatz (5.3) in non-cosmic time:

$$ds^2 = -a(t)^6\, dt^2 + a(t)^2\, dx_3^2\,. \tag{5.26}$$

Substituting this Ansatz in the Einstein equations yields

$$F^2 = \tfrac{1}{3}\, F' + \tfrac{1}{12}\, (\phi'^2 + \varphi'^2)\,, \tag{5.27}$$

$$F' = \tfrac{1}{2}\, V\, a^6\,, \tag{5.28}$$

where $F = a'/a$ is a Hubble parameter-like function, and the prime denotes differentiation w.r.t. t. The equations for the scalars are:

$$\phi'' = 0\,, \qquad \varphi'' = \sqrt{\Delta + 3}\, V\, a^6\,. \tag{5.29}$$

Combining (5.29) and (5.28) gives the following solutions for the scalars:

$$\varphi = 2\,\sqrt{\Delta + 3}\,\log(a) + a_1\, t + b_1\,, \qquad \phi = a_2\, t + b_2\,. \tag{5.30}$$

By substituting this into equation (5.25) we can deduce that the slope of the line is given by $C = a_2/a_1$. Substituting the scalars into (5.27) and (5.28) we are now left with the following two equations:

$$F' = -\Delta\, F^2 - \sqrt{\Delta + 3}\, a_1\, F - \tfrac{1}{4}\, (a_1^2 + a_2^2) \tag{5.31}$$

$$= \tfrac{1}{2}\, \Lambda\, e^{-\sqrt{\Delta+3}\,(b_1 + a_1\, t)}\, a^{-2\,\Delta}\,. \tag{5.32}$$

Keeping in mind that F' must be positive due to (5.32) we can now solve for F in the three different cases where Δ is positive, zero and negative. We can then easily find $a(t)$. We will choose b_1 (the constant part of φ) such that all solutions for $a(t)$ have a proportionality constant of 1, which does not affect the cosmological properties of the solutions. The integration constants appearing in the solutions are defined as follows:

$$c_1 = \frac{-\sqrt{\Delta+3}\,a_1}{2\,\Delta}\,, \quad c_2 = \frac{\sqrt{3\,a_1^2 - \Delta\,a_2^2}}{2}\,, \quad d_1 = -\frac{a_1^2 + a_2^2}{4\,\sqrt{3}\,a_1}\,, \quad d_2 = -\sqrt{3}\,a_1\,. \tag{5.33}$$

Below we present the solutions [111, 112] and their late- and early-time asymptotic behaviours (we give the latter without any irrelevant constants that rescale time):

1. $\underline{\Delta > 0}$:

$$a(t) = e^{c_1\,t}\,\cosh(c_2\,t)^{1/\Delta}\,, \qquad \text{for} \quad -\infty < t < +\infty\,. \tag{5.34}$$

   The positivity of F' requires a_1 to be negative, and it also imposes the following constraint:

$$\Big(\frac{a_2}{a_1}\Big)^2 < \frac{3}{\Delta}\,. \tag{5.35}$$

   This solution corresponds to a generic line on the first illustration in figure 5.2. It starts on the equator somewhere to the left of $x = \sqrt{3/(\Delta+3)}$, then moves in the direction of “Rome", but ends on the equator on the right-hand side. Note that the constraint (5.35) is simply the requirement that the slope of the line is bounded from above and from below such that the line actually intersects the sphere. We can confirm this asymptotic behaviour of the solution by converting to cosmic time (5.4) for $t \to -\infty$ and $t \to +\infty$ with the relation $a(t)^3\,dt = d\tau$:

$$\begin{aligned} &t \to -\infty\,, && \tau \to 0\,, && a \to e^t \sim \tau^{1/3}\,,\\ &t \to +\infty\,, && \tau \to +\infty\,, && a \to e^t \sim \tau^{1/3}\,. \end{aligned} \tag{5.36}$$

2. $\underline{\Delta = 0}$:

$$a(t) = e^{d_1\,t}\,\exp(e^{d_2\,t})\,, \qquad \text{for} \quad -\infty < t < +\infty\,. \tag{5.37}$$

   The positivity of F' requires a_1 to be negative. This corresponds to a line on the second illustration in figure 5.2. It starts on the equator and reaches “Rome"[5], which is also on the equator. Its asymptotic behaviour goes as follows:

$$\begin{aligned} &t \to -\infty\,, && \tau \to 0\,, && a \to e^t \sim \tau^{1/3}\,,\\ &t \to +\infty\,, && \tau \to +\infty\,, && a \to e^{e^t} \sim \tau^{1/3}\,. \end{aligned} \tag{5.38}$$

   To find the late-time behaviour of a in cosmic time one must realize the following two facts: First, $a(t) \sim \exp(e^t)$ for $t \to \infty$. Second, in this limit, $a' \sim a$ and therefore a behaves like a normal exponential.

[5] In this case, “Rome" is again attracting, however to see that, one must perform the stability analysis by going to second order perturbation. The first order vanishes, which means that the interpolating trajectory approaches “Rome" more slowly than in the cases where $\Delta < 0$.

3. <u>$-3 \le \Delta < 0$</u>:

$$a(t) = e^{c_1 t} \sinh(-c_2 t)^{1/\Delta}, \qquad \text{for} \quad -\infty < t < 0. \tag{5.39}$$

This solution corresponds to any line on the third illustration in figure 5.2. It starts at any point on the equator and ends at “Rome". This is reflected in the asymptotics as follows:

$$\begin{aligned} t &\to -\infty, & \tau &\to 0, & a &\to e^t \sim \tau^{1/3}, \\ t &\to 0, & \tau &\to +\infty, & a &\to (-t)^{1/\Delta} \sim \tau^{1/(\Delta+3)} \quad \text{for} \quad \Delta > -3, \\ & & & & &\sim e^{\tau} \quad \text{for} \quad \Delta = -3. \end{aligned} \tag{5.40}$$

There is one more solution for $-3 \le \Delta < 0$. If we set $a_1 = a_2 = 0$ we find:

$$a(t) = (-t)^{1/\Delta} \qquad \text{for} \quad -\infty < t < 0. \tag{5.41}$$

This solution corresponds to the “Rome" solution itself. For $-3 < \Delta < 0$ the conversion to cosmic time is the following:

$$a \sim \tau^{1/(\Delta+3)}. \tag{5.42}$$

Notice, however, that in the case where $\Delta = -3$, the “Rome" solution (5.41) and therefore the late-time asymptotics of (5.39) have a different conversion to cosmic time, namely:

$$a \sim (-t)^{1/\Delta} \sim e^{\tau}, \tag{5.43}$$

which we recognize as the De Sitter solution, in agreement with the fact that we have $V = \Lambda$.

The interpolating solutions above are given in non-cosmic time, which as mentioned is related to cosmic time by

$$d\tau = a(t)^3\, dt. \tag{5.44}$$

Integrating this equation yields hypergeometric functions for a generic interpolating solution, which we cannot invert to get the scale factor as a function of cosmic time. However, it is possible to get interpolating solutions in cosmic time for negative Δ when the following constraint on the constants holds:

$$\left(\frac{a_2}{a_1}\right)^2 = 12\, \frac{\Delta + \frac{9}{4}}{(2\Delta+3)^2}, \tag{5.45}$$

which can only be fulfilled for $-9/4 \le \Delta < 0$. The relation between the two time coordinates is

$$\tau = \frac{2^{-3/\Delta}}{2c_2}\, \frac{\Delta}{3+\Delta}\, (e^{2c_2 t} - 1)^{(3+\Delta)/\Delta}, \tag{5.46}$$

and the scale factor in cosmic time becomes

$$a(\tau) = \left(k_1\, \tau^{3/(3+\Delta)} + k_2\, \tau\right)^{1/3}, \tag{5.47}$$

where $k_1 = (2/c_1)^{3/\Delta}$ and $k_2 = k_1\, c_1\, (2\Delta + 3)/(18 + 6\Delta)$. From this solution, the asymptotic power-law behaviours are easily seen. The special one-scalar case, corresponding to $\Delta = -9/4$, was found in [125].

5.5 Acceleration

In this section we will investigate under which conditions “Rome" and the interpolating solutions represent an accelerating universe. This can be given a nice pictorial understanding in terms of the 2-sphere. We will show that acceleration takes place when the trajectory enters the region bounded by an “arctic circle". This is summarized in figure 5.3.

An accelerating universe is defined by $\ddot{a}/a > 0$. The existence of the “arctic circle" in connection to acceleration can now easily be determined. Assuming an expanding universe and using

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 \,, \tag{5.48}$$

as well as (5.18), we see that the condition for acceleration is equivalent to[6]

$$z^2 > \frac{2}{3} \,, \quad \text{i.e.} \quad x^2 + y^2 < \frac{1}{3} \,, \tag{5.49}$$

which exactly yields an “arctic circle" as the boundary of the region of acceleration. The straight line representing the exact solution is parametrized by the constants a_1 and a_2 as found in the previous section. From (5.49) and (5.25) it then easily follows that the condition for acceleration leads to the following restriction for the slope of the line:

$$\Big(\frac{a_2}{a_1}\Big)^2 (2 + \Delta) < 1 \,. \tag{5.50}$$

This condition is always fulfilled when $\Delta \leq -2$ and otherwise there is an interval of values for a_2^2/a_1^2 yielding an accelerating universe. This can easily be understood from figure 5.3. In general, a solution will only have transient acceleration. The only exception is when “Rome" lies within or on the “arctic circle", corresponding to $\Delta \leq -2$. Then, from the moment the line crosses the “arctic circle", there will be eternal acceleration [89] towards “Rome". When $\Delta = -2$, there will only be eternal acceleration when “Rome" is approached from the left. The possibilities of acceleration can be summarized as:

- $\Delta > -2$: A phase of transient acceleration is possible ,
- $\Delta = -2$: A phase of eternal acceleration is possible ,
- $-3 \leq \Delta < -2$: Always a phase of eternal acceleration .

The phase of eternal acceleration can also be understood from the power-law behaviour of the “Rome" solution, i.e. $a(\tau) \propto \tau^{1/(3+\Delta)}$. We have asymptotic acceleration when $1/(3+\Delta) > 1$, i.e. $\Delta < -2$. In the limiting case $\Delta = -3$, corresponding to “Rome" being on the North Pole, the interpolating solution will asymptote to De Sitter.

[6] A similar inequality was given in [105] for the one-scalar case, and in terms of the scalars and the potential in [112] for the multi-scalar case.

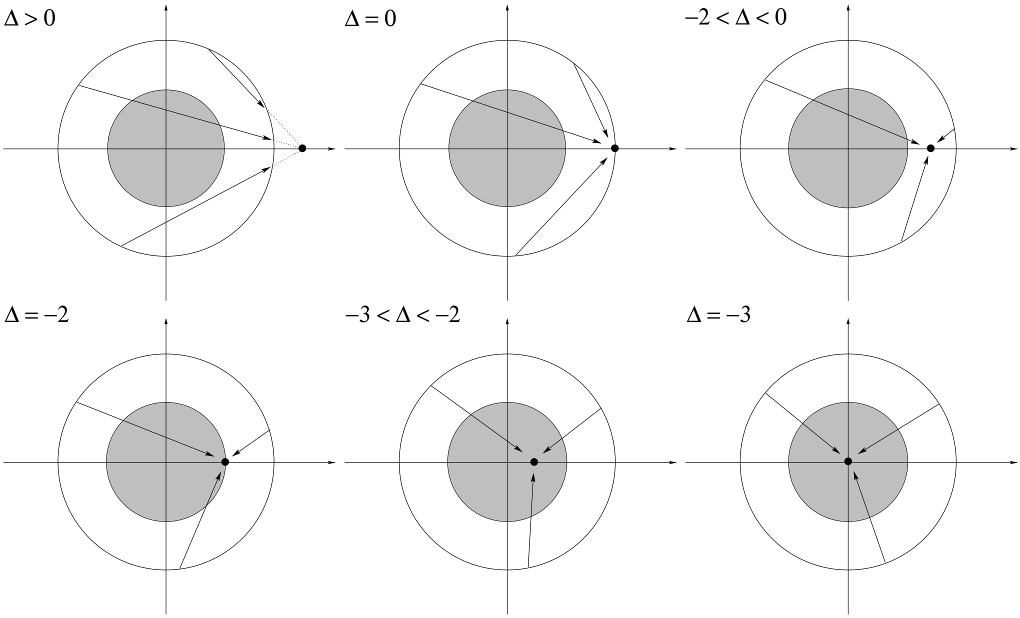


Figure 5.3: *The solutions represented as straight lines in the (x, y)-plane for $\Delta > 0$ where "Rome" is not on the sphere, $\Delta = 0$ where "Rome" is on the equator and $-3 \leq \Delta < 0$ where "Rome" is on the sphere. The inner circle corresponds to the "arctic circle", and solutions are accelerating when they enter the shaded area. The lower part of the figure corresponds to the cases where "Rome" is lying on the "arctic circle", $\Delta = -2$, inside the "arctic circle", $-3 < \Delta < -2$ and on the North Pole, $\Delta = -3$.*

5.6 Equation of state

In a cosmological setting, one often writes the matter part of the equations in terms of a perfect fluid, which is described by its pressure p and energy density ρ. These two variables are then assumed to be related via the equation of state:

$$p = \kappa \rho \,. \tag{5.51}$$

As is well known in standard cosmology, $\kappa = 0$ corresponds to the matter dominated era, $\kappa = 1/3$ to the radiation dominated era and $\kappa = -1$ to an era dominated by a pure cosmological constant. Quintessence is a generalization of the latter with $-1 \leq \kappa < -1/3$.

In our case, the matter is given by the two scalar fields, and thus p and ρ are given by the difference and sum of the kinetic terms and the potential, respectively:

$$p = \tfrac{1}{2}\,(\dot{\varphi}^2 + \dot{\phi}^2) - V\,, \quad \rho = \tfrac{1}{2}\,(\dot{\varphi}^2 + \dot{\phi}^2) + V\,. \tag{5.52}$$

Writing the above in terms of x, y and z, we see that the scalars describe a perfect fluid with an equation of state given in terms of the parameter:

$$\kappa = 1 - 2z^2\,. \tag{5.53}$$

Hence, κ varies from 1 on the equator to -1 on the North Pole, and we need $\kappa < -1/3$ for quintessence. For the interpolating solutions, which are given as curves on the sphere, κ will depend on time, but it will be constant for the critical points with the following values [89]:

- Equator : $\kappa = 1\,,$
- "Rome" : $\kappa = 1 + \frac{2}{3}\Delta\,.$ (5.54)

5.7 One-scalar truncations

The analysis has so far been done for two scalars, and as such it also contains the truncation to a system with one scalar with a potential, corresponding to $\phi = 0$. Here we will summarize the results of the previous sections in this truncation. On the sphere this yields $y = 0$, and for the solutions it corresponds to $a_2 = b_2 = 0$.

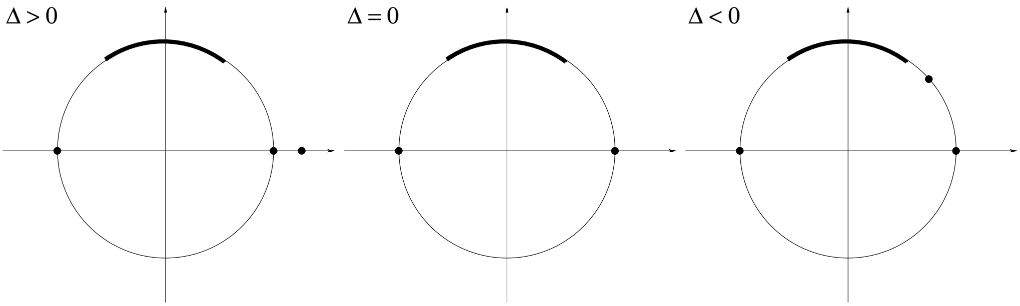


Figure 5.4: *The 2-dimensional (x, z) space and the critical points for the one-scalar truncations. The thick curve is the accelerating region. The two points on the x-axis are the equatorial critical points. The third point is "Rome". Note that in the middle illustration "Rome" coincides with the equatorial critical point $x = 1$.*

Since we only have one scalar, the Friedmann equation will define a circle when written in terms of x and z [105, 109]:

$$x^2 + z^2 = 1\,. \tag{5.55}$$

The critical points are [104]:

- Equatorial : $z = 0\,,\quad x^2 = 1\,,$
- "Rome" $(-3 \leq \Delta < 0)$: $z = \sqrt{-\Delta/3}\,,\quad x = \sqrt{1 + \Delta/3}\,.$ (5.56)

The full circle is shown in figure 5.4, including the critical points, and is just the vertical slice of the two-sphere including the North Pole. The equator therefore becomes two points, and the region bounded by the "arctic circle" now becomes the part of the circle[7] corresponding to $x^2 < 1/3$.

[7] This is equivalent to the accelerating region of [126, 127].

The exact solutions can be obtained from the previous section by setting $a_2 = b_2 = 0$. For $\Delta \geq 0$ the solutions correspond to the curves starting at $x = -1$ and ending at $x = 1$, whereas for $-3 \leq \Delta < 0$ the curves start at either one of the equatorial points and end at "Rome". In all cases where the curve starts at $x = -1$, the corresponding solution will give rise to acceleration. For this reason, interpolating solutions with $\Delta \geq 0$ will always give rise to a period of acceleration. This is in clear contrast to the two-scalar case, where it is possible to avoid acceleration (see figure 5.3). As for the 2-scalar case, if "Rome" lies in the "arctic" region, the solution will be eternally accelerating from the moment it enters this region.

One can also consider the truncation to zero scalars. However, from the scalar field equations, it is seen that this is only consistent if $\Delta = -3$, and this corresponds to the De Sitter solution with $V = \Lambda$.

A comment on the relevance of the interpolating solutions to inflation would be in order. In this context the number of e-foldings is crucial. As mentioned already, it is defined by $N_e = \log(a(\tau_2)/a(\tau_1))$ with τ_1 and τ_2 the start and end times of the accelerating period. These times can easily be found in our approach as the points where the straight lines intersect the "arctic circle". The number of e-foldings is required to be of the order of 65 to account for astronomical data. For the interpolating solutions with $\Delta > -2$, which is a necessary requirement to have a finite period of acceleration, one finds $N_e \lesssim 1$ [93,97,120] for all values of a_1, a_2 and Λ. The only exception to this behaviour is when $\Delta \to -2$, where N_e blows up. For the required 65 e-foldings one needs to take $\Delta + 2 \sim 10^{-60}$. As an example, for a compactification over an m-dimensional hyperbolic space, leading to $\Delta = -2 + 2/m$, this translates into $m \sim 10^{60}$. Thus, it seems that the e-foldings requirement for inflation cannot be met by a single exponential potential emanating from a dimensional reduction from the effective action of string/M-theory. Such a potential may, however, be relevant for describing present day acceleration. This does not exclude, however, that potential with Δ close enough to 2 for inflation might arise in a string theory scenario that takes other string theory effects into account, such as in [33].

In this chapter we introduced the scalar-gravity model in the FLRW context. We also introduced the language of autonomous systems and their fruitful application to cosmology. We learned that it is not necessary to find explicit solutions to the Einstein equations in order to get important qualitative information about our system. Although we were fortunate enough to write down the solutions explicitly, just by reasoning in terms of critical points and stability, we realized that power-law and de Sitter solutions are not the only kind of cosmology. We found solutions that interpolate between those two basic cases, some of which showed periods of transient acceleration. Transient acceleration is phenomenologically more interesting because a realistic model of cosmology should dynamically bring inflation to an end. It is also useful to consider scenarios with transient acceleration simply because we do not know whether present day acceleration will last forever.

So far, we have specialized in the case where the scalar potential consists of one exponential term. This lead to the huge simplification of being able to redefine our fields such that only one scalar appears in the exponent, no matter how many scalars were present in it to begin with. This was a particularly simple prelude to what we are about to do in the next chapter, where we will deal with the *most* general multi-exponential potential. We will rely entirely on the language of dynamical systems, as explicit interpolating solutions will become virtually impossible to find. By looking for critical points in such systems we will discover that intricate multi-exponential potentials do not simply accumulate the effects of a single exponential potential, but actually lead new de Sitter solutions that had not been seen before.

Chapter 6

Scalar Cosmologies II: A not so simple case

6.1 Introduction

In chapter 5, we studied scalar cosmology models, specializing in the single exponential potential case. We wrote down the field equations in the form of an autonomous system and studied its critical points and its interpolating solutions. However, the assumption of a single exponential in the potential lead to a great simplification, namely the amount of scalar fields was effectively reduced to two. In this chapter, we will drop this simplifying assumption and look at multi-exponential potentials.

The understanding of multi-exponential potentials has gradually evolved over the years. In the early days, the single exponential was studied in the context of inflation, where it was discovered that this potential allowed for a power-law solution [104]. Later on, the effect of adding exponential terms, each carrying a different scalar, was studied. This model is called "assisted inflation" [108]. The outcome is that the scalars 'assist' each other in the sense that each term contributes in the same way to the power-law behaviour of the scale factor and all the contributions are added. Later on, the effect of a cross-coupling between scalars was searched for, resulting in a model called "generalized assisted inflation" [122]. It was shown that these multi-exponential potentials also allowed for power-law solutions. However, the understanding of multi-exponential potentials wasn't complete. The class of potentials described in [122] does not cover all possible multi-exponential potentials. There is a strong restriction on the scalar couplings in that model, such that it only allows for power-law solutions. In other words, the potentials do not have any extrema. However, nowadays, a considerable amount of models that are inspired by string theory seem to be multi-exponentials with extrema (which allow for de Sitter solutions). Hence, they do not fall in the class of generalized assisted inflation.

The goal of this chapter is to study the most general multi-exponential potential. This is done using the elegant formalism of autonomous dynamical systems. We will construct all possible power-law and de Sitter solutions by finding the critical points to which they correspond in this formalism. We point out that this will uncover many new power-law and de Sitter solutions corresponding to so-called non-proper critical points that cannot be found in the case of generalized assisted inflation. To illustrate this, we will consider the special cases of double and triple exponential potentials with one or two scalars. For certain values of the scalar couplings, these can arise from M-theory, and their interpolating solutions correspond to the reduction of S-branes [117] and so-called exotic S-branes [95], respectively. For the exotic S-branes we derive the phases of accelerated expansion and find special cases where the number of such phases can be arbitrarily high. This can be useful for solving the cosmological coincidence problem, since oscillating dark energy could explain why we see a recent take over of dark energy in our present universe. It would simply be an event that occurs many times during the evolution of the universe.

The chapter is based on a collaboration with M. Nielsen and T. Van Riet, entitled *Scalar cosmology with multi-exponential potentials* [128]. It is organized as follows: in section 2, we present the system consisting of gravity and scalars with a potential. In section 3, we perform the general analysis of critical points. In section 4, we consider the special cases of double exponentials. In section 5, we present cases that can be obtained from the reduction over a three-dimensional group manifold. Finally, we end with a discussion of our results in section 6.

6.2 Scalar gravity with multi-exponential potentials

We consider 4-dimensional spatially flat FLRW gravity with N scalars ϕ_I which only depend on (cosmic) time τ. The scalars have a potential which is of the most general exponential form:

$$V(\vec{\phi}) = \sum_{i=1}^{m} \Lambda_i \, e^{-\vec{\alpha}_i \cdot \vec{\phi}} \, . \tag{6.1}$$

Thus, the scalar potential is characterized by m vectors $\vec{\alpha}_i$ and m constants Λ_i which can have positive or negative signs. The $\vec{\alpha}_i$ vectors form an $m \times N$ matrix α_{iI}, where the indices $i = 1, \ldots, m$ parametrize the exponential terms in the potential and the indices $I = 1, \ldots, N$ parametrize the different scalars. The Lagrangian for the system then reads[1]:

$$\mathcal{L} = \sqrt{-g}\left(R - \tfrac{1}{2}\,(\partial\vec{\phi})^2 - V(\vec{\phi})\right). \tag{6.2}$$

The equations of motion derived from the Lagrangian are

$$\begin{aligned} &\ddot{\phi}_I + 3H\dot{\phi}_I + \frac{\partial V}{\partial \phi_I} = 0 \, , \\ &H^2 = \tfrac{1}{12}\,(\dot{\vec{\phi}} \cdot \dot{\vec{\phi}}) + \tfrac{1}{6}\, V \, , \\ &\dot{H} = -\tfrac{1}{4}\,(\dot{\vec{\phi}} \cdot \dot{\vec{\phi}}) \, , \end{aligned} \tag{6.3}$$

where the dot is differentiation w.r.t. cosmic time. We refer to the equations as the scalar equations, the Friedmann equation and the acceleration equation, respectively. The Hubble constant H is defined as $H = \dot{a}/a$ where $a(\tau)$ is the scale factor appearing in the flat FLRW metric:

$$ds^2 = -d\tau^2 + a(\tau)^2 \, dx_3^2 \, . \tag{6.4}$$

[1]We use the convention for the metric with mostly plus signature.

There are $N + 1$ degrees of freedom, namely, the scale factor and the N scalars (and accordingly only $N + 1$ equations of motion are independent). For example, the acceleration equation can be obtained from the Friedmann equation and the scalar equations). There exist 2 types of solutions:

- Critical points: These solutions correspond to stationary solutions defined in terms of certain dimensionless variables, which will be introduced in the next section. The critical points can be obtained explicitly and they correspond to power-law solutions ($a(\tau) \sim \tau^p$) or de Sitter solutions[2] ($a(\tau) \sim e^\tau$). The solutions can be attractors, repellers or saddle points. In the former two cases they correspond to the asymptotic behaviour of more general solutions, whereas a saddle point just corresponds to an intermediate regime.

- Interpolating solutions: These are the non-stationary solutions and in general they will interpolate between the critical points. Often they cannot be found explicitly, but a numerical analysis can reveal most of their properties.

6.3 The critical points

Critical points (also known as fixed points or equilibrium points) are solutions of differential equations in the context of autonomous dynamical systems. An autonomous system is defined as a system described by n variables, say $\vec{z}$, that depend on one variable t, whose dynamical equations are of the form:

$$\frac{d\vec{z}}{dt} = \vec{f}(\vec{z}) \,, \tag{6.5}$$

where $\vec{f} : \mathbb{R}^n \to \mathbb{R}^n$ is interpreted as a vector field on $\mathbb{R}^n$. A key feature of autonomous systems is the absence of the independent variable t on the right-hand-side of the dynamical equation (6.5). Solutions for $\vec{z}(t)$ are then integral curves to the vector field $\vec{f}$; i.e. $\vec{f}$ is everywhere tangent to all possible curves $\vec{z}(t)$. The critical points of an autonomous system are defined as those points $\vec{z}_0$ obeying $\vec{f}(\vec{z}_0) = 0$. These points are always exact constant solutions since $d\vec{z}_0(t)/dt = 0$. The interesting property about these systems is that the critical points are often the end points (and initial points) of the orbits and therefore describe the asymptotic behaviour. If the solutions interpolate between critical points, they can be divided into two classes:

- Heteroclinic orbit: This is an orbit connecting two different critical points.

- Homoclinic orbit: This is an orbit connecting a critical point to itself.

Most of the examples we have found are of the first type and we will focus on those. More on the theory of dynamical systems in cosmology can be found in [129, 130].

An useful property of multi-scalar cosmology with exponential potentials is that they allow for a description in terms of variables that make the system autonomous [104, 106, 107, 110]. With an arbitrary multi-exponential potential, the variables are defined as follows:

$$x_I = \frac{\dot{\phi}_I}{\sqrt{12}\, H} \,, \quad y_i = \sqrt{\frac{\Lambda_i \, e^{-\vec{\alpha}_i \cdot \vec{\phi}}}{6H^2}} \,. \tag{6.6}$$

[2] Anti-de Sitter solutions are not possible since a flat FLRW metric doesn't support them.

In this notation, there are $N + m$ variables. Note that y_i will be imaginary when $\Lambda_i < 0$, but this is not a problem since only y_i^2 appears in the equations of motion. Rewriting the equations of motion with these variables yields

$$\frac{\dot{x}_I}{H} = -3\,y^2\,x_I + \sqrt{3}\sum_{i=1}^{m}\alpha_{iI}y_i^2\,, \tag{6.7}$$

$$x^2 + y^2 = 1\,, \tag{6.8}$$

$$\frac{\dot{H}}{H^2} = -3\,x^2\,, \tag{6.9}$$

where we have used the shorthand notation $x^2 = \sum_{I=1}^{N} x_I^2$ and $y^2 = \sum_{i=1}^{m} y_i^2$. An interesting consequence of the choice of variables is that the Friedmann equation (6.8) becomes the defining equation of an $(N + m - 1)$-sphere for $\Lambda_i > 0$ (otherwise it will be a generalized hyperboloid). Furthermore, from the acceleration equation, the condition for accelerated expansion translates into the following simple constraint:

$$\ddot{a} > 0 \qquad \Leftrightarrow \qquad x^2 < \frac{1}{3}\,. \tag{6.10}$$

The above condition allows us to visualize the region of acceleration for the specific examples in section 4 and 5.

It turns out that we also need the derivatives of the y-variables:

$$\frac{\dot{y}_i}{H} = \sqrt{3}\,(\sqrt{3}\,x^2 - \vec{\alpha}_i\cdot\vec{x})\,y_i\,. \tag{6.11}$$

We can also use $\ln(a)$ as evolution parameter[3] instead of cosmic time, which simplifies the equations since H drops out in the scalar equations of motion and the equations for $\dot{y}_i$, giving

$$\boxed{x_I' = -3\,y^2\,x_I + \sqrt{3}\sum_{i=1}^{m}\alpha_{iI}y_i^2\,,\quad y_i' = \sqrt{3}\,(\sqrt{3}\,x^2 - \vec{\alpha}_i\cdot\vec{x})\,y_i\,,} \tag{6.12}$$

where the prime indicates differentiation w.r.t. $\ln(a)$. The above is clearly of the form (6.5), and the critical points can therefore be calculated as $x_I' = y_i' = 0$ (or equivalently as $\dot{x}_I = \dot{y}_i = 0$). It is easy to prove that the system will obey the Friedmann constraint (6.8) at all times as long as it does so initially. Hence, if we impose (6.8) on the initial conditions, then (6.12) contains all the information about the subsequent evolution.

Integrating the acceleration equation (6.9) for a critical point yields power-law solutions if $x^2 \neq 0$

$$a(\tau) \sim \tau^p, \qquad p = \frac{1}{3\,x^2}\,. \tag{6.13}$$

If on the other hand, if $x^2 = 0$, then the critical point is an extremum of the potential with a de Sitter expansion

$$a(\tau) \sim \exp(\sqrt{\tfrac{1}{6}V(\phi_c)}\,\tau)\,. \tag{6.14}$$

[3] However, one must be careful if the scale factor is not strictly monotonic.

The equations (6.12) determining the critical points are

$$(\sqrt{3}\,x^2 - \vec{\alpha}_i \cdot \vec{x})\,y_i = 0\,, \tag{6.15}$$

$$-3\,y^2\,x_I + \sqrt{3}\,\sum_{i=1}^{m} \alpha_{iI} y_i^2 = 0\,. \tag{6.16}$$

There are two different kinds of critical points:

- Proper solutions: $\nexists i : y_i = 0\,,$
- Non-proper solutions: $\exists i : y_i = 0\,.$

We can single out special non-proper solutions, which always exist, namely, the case where all y's vanish. From the Friedmann equation it follows that these solutions have $x^2 = 1$ and for this reason we refer to them as "the equator". The solutions with some y's vanishing have infinite scalars and therefore, are not proper solutions of the equations of motion. They are, however, very important, since they correspond to the asymptotic behaviour of interpolating solutions. From this classification, we see that there are a maximum of 2^m types of critical point solutions [110]. Below we will give these solutions for the most general exponential potential by analysing (6.15) and (6.16).

The rank R of the matrix α_{iI}, i.e. the number of independent $\vec{\alpha}_i$-vectors, plays a central role in this discussion. In fact, the discussion of the general potential naturally splits up into two cases: $R = m$ and $R < m$.

The rank R gives the effective number of scalars appearing in the potential, corresponding to the part of the scalar space that is projected on the $\vec{\alpha}_i$-vectors. It is therefore always possible to perform a field redefinition, such that only R scalars appear in the potential. The part of the scalar space perpendicular to the $\vec{\alpha}_i$-vectors only appears in the kinetic term of the Lagrangian and is $(N - R)$-dimensional. Therefore, these scalars decouple from the rest. All systems with $N > R$ have decoupled scalars and this is necessarily the case when $N > m$. Systems with $N \leq m$ only have decoupled scalars if the vectors $\vec{\alpha}_i$ are linearly dependent in such a way that $N > R$.

The field redefinition yielding R scalars in the potential can be performed by an $SO(N)$ rotation (which leaves the kinetic term invariant) such that $\vec{\phi}$ changes into $\vec{\phi}'$ and $\alpha'_{iR+1} = \alpha'_{iR+2} = \ldots = \alpha'_{iN} = 0$ for all i. We then notice from (6.16) that, for critical points, all x's corresponding to decoupled scalars are zero, $x_{R+1} = x_{R+2} = \ldots = x_N = 0$. Therefore, in the rest of this section, the indices I now run from 1 to R. In the case $R = m$, this makes α_{iI} a square matrix.

We have seen that the discussion of the system can be split up into two cases, depending on the rank of α_{iI}. Alternatively, we can formulate this in terms of the following matrix, which is quadratic in the α's

$$A_{ij} = \vec{\alpha}_i \cdot \vec{\alpha}_j\,. \tag{6.17}$$

The separation of the general exponential potential into two classes can then be characterized by the determinant of A:

$$R = m: \qquad \det(A) \neq 0\,, \tag{6.18}$$

$$R < m: \qquad \det(A) = 0\,.$$

The first class corresponds to an invertible A-matrix and this is exactly what is termed generalized assisted inflation [122]; whereas the second class, to our knowledge, has not been treated in generality in the literature.

We will extend the existing results by also treating the case of non-invertible A in generality and providing the non proper critical points of both classes. Special examples can be obtained by performing compactifications over certain three-dimensional unimodular group manifolds corresponding to class A in the Bianchi classification, see e.g. [95].

There is a subtlety about the description in terms of the (x_I, y_i)-variables, namely if $R < m$ then the y-variables are not necessarily independent. We will comment on this in section 3.2.

6.3.1 The $R = m$ case

This case has the simplifying feature that $\dot{x}_I = 0$ implies $\dot{y}_i = 0$. This can be seen in the following way: first we differentiate (6.7) and use $\dot{x}_I = d(y^2)/d\tau = 0$. Multiplying with α_{jI} and summing over I we get $\sum_j (A_{ij})\, d(y_j^2)/d\tau = 0$, and since $\det(A) \neq 0$ we know that the only solution is $d(y_i^2)/d\tau = 0$.

We will now solve for the critical points:

- Proper critical points:

  From (6.15) and (6.16) we get:

$$\sum_i (A_{ij})\, y_i^2 = 3y^2\, x^2\, e_j\,, \tag{6.19}$$

  where e_j is an m-dimensional vector with all components equal to 1. Inverting this relation and using (6.16) yields the values of y_i and x_I for the proper critical point

$$y_i^2 = \frac{3p-1}{3p^2} \sum_{j=1}^{m} (A^{-1})_{ij}\,, \qquad x_I = \frac{\sqrt{3}\,p}{3p-1} \sum_i^{m} \alpha_{iI}\, y_i^2\,, \tag{6.20}$$

  where p is the exponent given in (6.13). The result for x_I can also be given in the rotated basis where α_{iI} is a square matrix

$$x_I = \frac{1}{\sqrt{3}\,p} \sum_{i=1}^{m} (\alpha^{-1})_{iI}\,. \tag{6.21}$$

  Note that by construction, the A_{ij}-matrix (6.17) is $SO(N)$-invariant and accordingly, all quantities containing only this matrix can be calculated in any basis. We notice from our formula above that there is a unique proper critical point. However, it only exists when y_i^2, as determined from (6.20), has the same sign as Λ_i, which serves as a consistency check of definition (6.6). Thus, this critical point only exists for certain values of the α-vectors.

  Using (6.13), we get the exponent for the power-law that reproduces the result found in [122, 131]:

$$p = \sum_{i,j=1}^{m} (A^{-1})_{ij}\,. \tag{6.22}$$

By integration we can go back to the ϕ_I, H variables where the solution becomes:

$$H = \frac{p}{\tau}\,, \qquad \phi_I = \sqrt{12}\, p\, x_I\, \ln(\tau) + c_I\,, \qquad y_i^2 = \frac{k_i}{\tau^2}\,, \tag{6.23}$$

where c_I and k_i are integration constants. In fact, in [103, 122], (6.23) was used as an Ansatz to find power-law solutions.

- **Non-proper critical points**:

  These correspond to some y's being equal to zero. Parametrising the subset of nonzero y's with the indices $a, b, c, \ldots$, the equations become:

$$\sqrt{3}x^2 - \vec{\alpha}_a \cdot \vec{x} = 0\,, \qquad \sum_a \alpha_{Ia}(y_a)^2 - (1-x^2)\sqrt{3}x_I = 0\,, \tag{6.24}$$

  from which we deduce:

$$\sum_b (A_{ab}) y_b^2 = 3\, y^2\, x^2\, e_a\,. \tag{6.25}$$

  The $\vec{\alpha}_a$-vectors are of course also linearly independent and accordingly, the sub-matrix A_{ab} has non-zero determinant and is therefore invertible. Inverting relation (6.25) and using (6.16), we find a unique solution

$$y_a^2 = \frac{3p-1}{3p^2} \sum_b (A^{-1})_{ab}\,, \qquad x_I = \frac{\sqrt{3}\,p}{3p-1} \sum_a^m \alpha_{aI}\, y_a^2. \tag{6.26}$$

  The power-law is again given by (6.22) but now with the inverse of the sub-matrix A_{ab}. Just as for the proper solution, the above is only well-defined when y_a^2 has the same sign as Λ_a. Note that all the above formulae for the non-proper critical points are similar to those for the proper ones. This is due to the fact that vanishing y's just yield a truncated potential.

Note that since the solution for the proper critical point is unique and has power-law behaviour for the scale factor, there are no de Sitter solutions. This can also be seen from (6.19). Since A has maximal rank, this matrix only has the trivial nullspace, i.e $y_i = 0$, which is not consistent with the Friedmann equation, since $x = 0$ for the de Sitter solutions. We can conclude that potentials with linearly independent $\vec{\alpha}_i$-vectors generically have power-law solutions and no de Sitter solutions. This conclusion was also reached in [132] where special cases were considered.

The special case where α_{iI} is diagonalisable by an $SO(N)$-rotation is equivalent to the case where just one scalar appears in each exponential, thus yielding the model which has been called assisted inflation [108].

6.3.2 The $R < m$ case

Since $\det(A) = 0$ we will have to use another approach to determine the critical points. And the $R < m$ case will also be more difficult to treat in full generality because the y's are not necessarily independent.

The number of independent y's is always smaller than or equal to $R+1$, as we will now illustrate. After possible field redefinitions, the y-coordinates are given in terms of $R+1$ fields, namely the scalars and the Hubble parameter. Therefore, among the m coordinates, at most $R+1$ are independent, e.g. $y_1, \ldots, y_{R+1}$. This leaves us with $m-R-1$ relations for the rest of the y's. From the definitions of the y's, we can express ϕ_I and H in terms of the first $R+1$ y's

$$e^{\phi_I} = \prod_{i=1}^{R} \Big(\frac{y_i^2\,\Lambda_{i+1}}{y_{i+1}^2\,\Lambda_i}\Big)^{(\beta^{-1})_{Ii}} , \qquad H = \frac{\Lambda_{R+1}}{6}\, e^{-\vec{\alpha}_{R+1}\cdot\vec{\phi}}\, y_{R+1}^{-2} , \tag{6.27}$$

where the following square matrix has been defined

$$\beta_{iJ} = \alpha_{i+1,J} - \alpha_{iJ} , \quad i, J \in \{1, \ldots, R\} . \tag{6.28}$$

We can then express the remaining y's in terms of the first $R+1$ as follows

$$y_i^2 = y_{R+1}^2 \,\frac{\Lambda_i}{\Lambda_{R+1}}\, \frac{\prod_{j,K=1}^{R}\Big(\frac{y_j^2\,\Lambda_{j+1}}{y_{j+1}^2\,\Lambda_j}\Big)^{\alpha_{iK}(\beta^{-1})_{Kj}}}{\prod_{l,M=1}^{R}\Big(\frac{y_l^2\,\Lambda_{l+1}}{y_{l+1}^2\,\Lambda_l}\Big)^{\alpha_{R+1,M}(\beta^{-1})_{Ml}}} , \qquad i = R+2, \ldots, m . \tag{6.29}$$

Thus, the maximal number of independent y's is $R+1$. It is possible to prove that the dynamical system (6.12) will obey the above relations for the y_i's at all times if it do so initially. So again we can use (6.12) as equations that govern the whole system, as long as we pick our initial conditions consistently. With this in mind we will look for critical points.

Until now we have denoted the row vectors of the α-matrix with $\vec{\alpha}_i$ and A_{ij} was defined as the matrix with the inner products of these row vectors as entries: $A_{ij} = \vec{\alpha}_i \cdot \vec{\alpha}_j$. In this section we will also need the column vectors which we will denote by $\vec{\alpha}_I$ and we will need to define the following matrix

$$B_{IJ} = \vec{\alpha}_I \cdot \vec{\alpha}_J . \tag{6.30}$$

The R column vectors $\vec{\alpha}_I$ are all linearly independent because the rank of α equals R and, consequently, B is invertible (remember that I now runs from 1 to R). It is this property that we will use to find the solutions.

- **Proper power-law critical points**:

  Looking for the solution(s), with $y_i \neq 0$, we find from (6.15)

$$\sum_{I=1}^{R} B_{IJ} x_I = \sqrt{3} x^2 F_J , \tag{6.31}$$

  where $F_J = \sum_{i=1}^{m} \alpha_{iJ}$. Thus, we can solve for x_I:

$$x_I = \frac{1}{\sqrt{3}\,p} \sum_{J=1}^{R} (B^{-1})_{IJ}\, F_J . \tag{6.32}$$

Hence we find the extension of the power-law formula to the case where $R < m$:

$$p = |B^{-1} \cdot \vec{F}|^2 \,. \tag{6.33}$$

One can prove that this formula reduces to (6.22) if $R = m$. Since the rank of α_{iI} is R, it is enough to use R independent equations among the m equations of (6.15) to obtain x_I. This result, of course, has to be consistent with the remaining $m - R$ equations, and this puts strong restrictions on the allowed dilaton couplings as we will now show. Let $\{\vec{\alpha}_a\}_{a=1}^R$ be linearly independent. It is possible to solve (6.15) simultaneously for these vectors. The rest of the vectors can be written as linear combinations and are only guaranteed to solve (6.15) if the linear combinations are convex [4]

$$\vec{\alpha}_i = \sum_{a=1}^{R} c_{ia}\vec{\alpha}_a \,, \qquad \sum_{a=1}^{R} c_{ia} = 1 \,, \qquad i = R+1, \ldots, m \,. \tag{6.34}$$

We will give a specific example with an M-theory origin, where this is realized. A special case is $R = 1$, where after field redefinitions only one scalar appears in the potential. In this case, the above solution will never exist, since (6.15) becomes m equations with one variable (or equivalently, the requirement of convexity here would imply $m = 1$, which is not the case under consideration).

An important difference between this and the previous case is the question of the uniqueness of the solution. We cannot obtain the y-values with this procedure, and in particular we cannot determine whether they are unique. In fact, it is easy to give an example where they are not: when at least one $\Lambda_i < 0$, we have the following possibility, since A has a non-trivial kernel

$$\begin{aligned} &y^2 = 0 \,, \quad y_i^2 \in \mathrm{Ker}(A) \,, \\ &x^2 = 1 \,, \quad \vec{\alpha}_i \cdot \vec{x} = \sqrt{3} \,, \quad \text{for} \quad y_i \neq 0 \,. \end{aligned} \tag{6.35}$$

In particular, this includes a proper critical point of the form (6.32) when all $y_i \neq 0$ and where furthermore

$$|B^{-1} \cdot \vec{F}|^2 = \frac{1}{3} \,. \tag{6.36}$$

- **De Sitter solutions**:

We have seen in the previous subsection, that de Sitter solutions do not exist for $R = m$, because the matrix A has a trivial kernel. In the present case, since A has a non-trivial kernel, making a de Sitter solution is possible, we have the following:

$$x = 0 \,, \quad y^2 = 1 \,, \quad y_i^2 \in \mathrm{Ker}(A) \,. \tag{6.37}$$

Again, this solution is only well-defined when y_i^2 has the same sign as Λ_i. We can conclude that potentials with $R < m$ show the opposite behaviour of $R = m$ potentials. Here, (proper) power-law solutions are rare (only possible for certain couplings (6.34)), whereas de Sitter solutions are quite generic. Again, a similar observation was made in [132], but for specific couplings (which did not allow power-law behaviour).

[4] Of course, there are many ways to number the vectors; it suffices to find one that obeys these relations.

- **Non-proper critical points**:

  Looking for these solutions, we again put a subset of the y's to zero. This corresponds to some terms in the potential being absent. Therefore, we can analyse the new system as before but with a "truncated" potential. A subtlety appears whenever $R < m$, namely the y's are dependent on each other, and therefore only certain subsets of the y's can be zero simultaneously.

The findings of this section are summarized in the table below. The asterisk in the lower left corner stands for a truncated system, which can belong to either of the two cases ($R < m$ or $R = m$).

	$R < m\,, \quad \det(A) = 0$	$R = m\,, \quad \det(A) \neq 0$
Proper	Power-law (convex combinations)	Power-law
	de Sitter	No de Sitter
Non-proper	$*$	Power-law
		No de Sitter

Table 6.1: *The critical points for multi-exponential potentials.*

As mentioned before, the critical points give rise to the asymptotic behaviour of the general solutions. By performing stability analysis it is possible to determine the nature of the critical points, i.e. whether they are attractors, repellers, or saddle points. This can be done by linearizing the system around the critical points, $\vec{x}' = \mathbf{M} \cdot \vec{x}$, and determining the eigenvalues of the matrix $\mathbf{M}$. If the real part of all eigenvalues is negative, the critical point is an attractor; if the real part of all eigenvalues is positive, the critical point is a repeller; and in the mixed case it is a saddle point. It is easy to perform the stability analysis in the simple cases considered in the following sections, and the result is confirmed by the interpolating solutions, which are calculated numerically.

6.4 Double and triple exponential potentials

In this section we will consider some specific examples of double and triple exponential potentials with one or two scalars, i.e. $m = 2, 3$ and $N = 1, 2$. These examples serve as an illustration of the formal framework in the previous section.

As mentioned before, the critical points reveal the asymptotic behaviour of more general solutions. In some cases it has been possible to obtain these solutions exactly. For single exponential potentials, this was done for arbitrary dilaton couplings and the result can be pictured as straight lines in the space defined by the x's [95]. For double exponential potentials, exact solutions were obtained for special values of the dilaton couplings, corresponding to the reduction of S-brane solutions to 4D, see e.g. [91, 92, 94] and references therein. Ideally, we would like to obtain exact results for the general case. However, this is a highly non-trivial task, and we therefore turn to numerical methods, which can still show the qualitative behaviour of the solutions. To this end, it is convenient to use $\ln(a)$ as a time parameter. For an eternally expanding universe where a increases from 0 to ∞, our time coordinate ranges from $-\infty$ to ∞.

In general, an S-brane can be obtained as a time-dependent solution to the following system containing gravity, an antisymmetric tensor, and possibly a dilaton:

$$S = \int d^{4+d}x \, \sqrt{-\hat{g}} \left(\hat{R} - \tfrac{1}{2} (\partial\hat{\phi})^2 - \tfrac{1}{2d!} \, e^{-b\hat{\phi}} \, \hat{F}_d^2 \right) , \tag{6.38}$$

where the hats indicate that the fields live in $4 + d$ dimensions and where the dilaton coupling for maximal supergravities is given by

$$b = \sqrt{\frac{14 - 2d}{d + 2}} \, . \tag{6.39}$$

Reducing over a d-dimensional maximally symmetric space with curvature k and flux f yields the following potential [133]

$$V(\phi, \varphi) = f^2 \, e^{-b\,\phi - 3\sqrt{\frac{d}{d+2}}\,\varphi} - k\, e^{-\sqrt{\frac{d+2}{d}}\,\varphi} \, , \tag{6.40}$$

where φ is the Kaluza-Klein scalar. S2-brane solutions have been found in six to eleven dimensions, corresponding to $d = 2, \ldots, 7$. In five dimensions, an S2-brane has a 1-form field strength. The corresponding four-dimensional cosmological solution with single exponential potential was found in [95]. As explained in that paper, a general twisted reduction leads to triple exponential potentials, which could have corresponding exotic S-brane solutions in five dimensions.

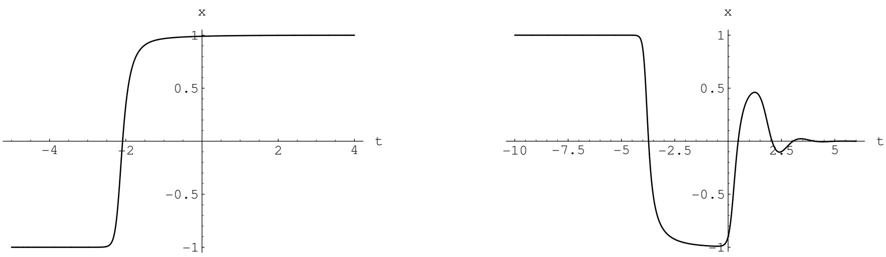
6.4.1 Double exponential potentials, one scalar

The simplest case is $m = 2$ and $N = 1$. The corresponding potential is described in terms of two dilaton couplings α_1 and α_2. We can always choose e.g. α_1 to be positive and in this example we will start by considering positive Λ_i. Since $R = 1$, we have 2 independent y's. The Friedmann equation defines a 2-sphere, but the allowed solutions can only lie on the part corresponding to non-negative y's. Using the machinery from the previous section, we find the following critical points

$$\begin{aligned}
&(i) && y_1 = y_2 = 0\,, \quad x^2 = 1\,, \\
&(ii) && y_1 = \sqrt{1 - \frac{\alpha_1^2}{3}}, \quad y_2 = 0\,, \quad x = \frac{\alpha_1}{\sqrt{3}}\,, && \text{for} \quad \alpha_1^2 < 3\,, \\
&(iii) && y_1 = 0\,, \quad y_2 = \sqrt{1 - \frac{\alpha_2^2}{3}}\,, \quad x = \frac{\alpha_2}{\sqrt{3}}\,, && \text{for} \quad \alpha_2^2 < 3\,, \\
&(iv) && y_1 = (1 - \frac{\alpha_1}{\alpha_2})^{-1/2}\,, \quad y_2 = (1 - \frac{\alpha_2}{\alpha_1})^{-1/2}\,, \quad x = 0\,, && \text{for} \quad \alpha_2 < 0\,.
\end{aligned} \tag{6.41}$$

The first, (i) corresponds to the "equatorial" points $x = \pm 1$. In an (y_1, y_2, x)-plot these become the North and South Pole. The next two, (ii) and (iii), are the non-proper critical points. The last one, (*iv*), is the proper solution, which only exists for $\alpha_2 < 0$ and corresponds to a de Sitter solution. The stability of the different points is best illustrated by considering the different possible cases[5].

- $\alpha_1, \alpha_2 > \sqrt{3}$: Only the North and South Pole are critical points; the former is attracting and the latter repelling. Any interpolating solution will be a curve between them, and these can be found numerically. An example is illustrated in figure 6.1(a).
- $\alpha_1 < \alpha_2 < \sqrt{3}$: The critical points (i)-(iii) exist. The poles are repellers and (iii) is attracting.
- $\alpha_1 < \sqrt{3}, \alpha_2 < -\sqrt{3}$: Apart from the poles, we have the two critical points, corresponding to (iii) and (iv) in (6.41). The North Pole is repelling and the de Sitter solution is attracting; this is shown in figure 6.1(b).

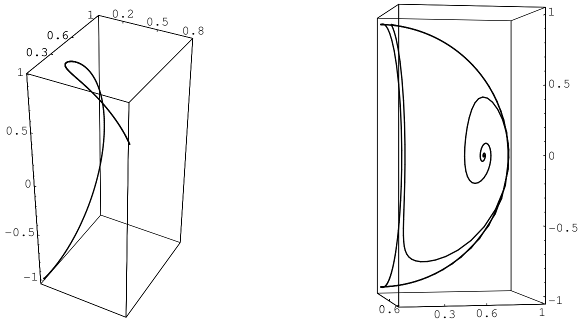


(a) (b)

Figure 6.1: *Plot (a) shows $x(t)$ in the case $(\alpha_1, \alpha_2) = (3, 2)$, where the solution interpolates between the North and South Pole. Plot (b) is for the case $(\alpha_1, \alpha_2) = (1, -2)$, yielding a solution interpolating between the North Pole and a de Sitter solution.*

- $\alpha_1 > \sqrt{3}, -\sqrt{3} < \alpha_2 < 0$: This is similar to the previous case, except that the critical point (iii) is interchanged with (ii), and the early asymptotics will be the South Pole.
- $\alpha_1 > \sqrt{3}, 0 < \alpha_2 < \sqrt{3}$: In addition to the North and South Pole, there is the non-proper critical point (ii), which is an attractor. The South Pole is repelling. An interpolating solution is shown in figure 6.2(a).
- $\alpha_1 > \sqrt{3}, \alpha_2 < -\sqrt{3}$: The critical points are the poles and the de Sitter solution, and the latter is an attractor. It turns out that the poles are saddle points; hence, they do not give rise to the early asymptotics of the solution. Instead, this will be an infinite cycle, moving closer and closer to the boundary of the space (given by $y_1 = 0$ or $y_2 = 0$), as time goes to minus infinity. This is illustrated in figure 6.2(b).
- $\alpha_1 < \sqrt{3}, -\sqrt{3} < \alpha_2 < 0$: The late-time asymptotics are similar to those of the previous case. The early-time asymptotics are different due to the fact that all the critical points (i)-(iv) are realized. Both of the poles will be repelling, and depending on initial conditions, either of these can give rise to the early-time asymptotics.

[5]If we do not explicitly classify the stability of a critical point, it will be a saddle point.



(a) (b)

Figure 6.2: *The figure shows* (y_1, y_2, x)*-plots for two cases. Figure (a) with* $(\alpha_1, \alpha_2) = (2, 1)$*, shows a solution interpolating between the South Pole and the critical point (ii). Plot (b), with* $(\alpha_1, \alpha_2) = (3, -2)$ *shows a solution spiralling towards the de Sitter point.*

For all the cases above, the solutions enter a phase of acceleration when $x^2 < 1/3$. The cases with $|\alpha_1|, |\alpha_2| > 1$ give rise to one period of transient acceleration, otherwise the solution will end up in a phase of eternal acceleration, which, as mentioned before, is an asymptotic de Sitter phase when $\alpha_2 < 0$. In the case $\alpha_1 > \sqrt{3}, \alpha_2 < -\sqrt{3}$ the phase of late-time acceleration is preceded by an infinite cycle, alternating between acceleration and deceleration.

The case with $\Lambda_2 < 0$ can be analysed in a similar way, but the interpolating solutions will now be given by curves on a hyperboloid. The critical point (*iii*) will only exist for $\alpha_2^2 > 3$, since this yields $y_2^2 < 0$. By the same token, the de Sitter critical point (*iv*) only exists for $\alpha_2 > \alpha_1 > 0$.

The S-brane case corresponding to $\hat{\phi} = 0$ in (6.38), gives the following dilaton couplings

$$\alpha_1 = 3\sqrt{\frac{d}{d+2}}, \quad \alpha_2 = \sqrt{\frac{d+2}{d}}. \tag{6.42}$$

This system, which can be obtained from eleven dimensions where it will give rise to SM2-brane solutions, was analysed in [96], where curvature of the external space is also included. One can show that only the critical points (*i*) and (*iii*) exist for $\Lambda_2 > 0$; and (*i*) and (*ii*) exist for $\Lambda_2 < 0$, with the latter being attracting. However, In the latter case, we also have a de Sitter critical point, which is not an attractor.

6.4.2 Double exponential potential, two scalars

Let us now study the case with double exponential potentials and two scalars, i.e. $m = 2$ and $N = 2$. Considering the two α-vectors to be independent, we get $R = 2$. The critical points can be obtained as a special case of the general analysis from the previous section and consist of the equator, $x^2 = 1, y_i = 0$; the proper critical point, $y_i \neq 0$; and two non-proper critical points with $y_i = 0, y_j \neq 0, i \neq j$.

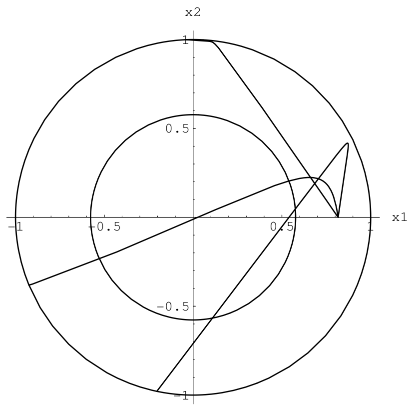


Figure 6.3: *Three interpolating solutions, corresponding to S2-branes reduced to four dimensions, projected on the (x_1, x_2)-plane. The inner circle is the boundary of the accelerating region.*

Specialising to the reduction of S-branes, we get the following dilaton couplings:

$$\vec{\alpha}_1 = (3\sqrt{\frac{d}{d+2}}, \sqrt{\frac{14-2d}{d+2}}), \quad \vec{\alpha}_2 = (\sqrt{\frac{d+2}{d}}, 0)\,. \tag{6.43}$$

For these α-couplings we get $\det(A) = 14/d - 2$; therefore, the matrix A is invertible for $d < 7$. However, using (6.20), y_1^2 we see that is negative, and, since $\Lambda_1 = f^2 > 0$, the proper critical point does not exist. Apart from the equator, there is another critical point, which has $y_1 = 0$ and $y_2 \neq 0$ and corresponds to a power-law behaviour with exponent $p = d/(d+2)$. This critical point is an attractor and the equator is a repeller. Thus, an S2-brane reduced to four dimensions corresponds to a solution interpolating between the equator and the attracting critical point. This is similar to the behaviour of the solution found in [83], which is the fluxless limit of a reduced S2-brane [91]. Indeed, the attracting power-law solution is the same with or without flux. Examples of interpolating solutions, projected on the (x_1, x_2)-plane, are shown in figure 6.3, in the case of $d = 2$. One can see that they indeed interpolate between the equator and the attracting critical point, which, according to (6.26), has the coordinates $(x_1, x_2) = (\sqrt{2/3}, 0)$. For $d = 7$, there is a possibility of a de Sitter solution, since $\det(A) = 0$, see (6.37). However, it does not exist because y_1^2 is negative.

6.4.3 Triple exponential potential, one scalar

This example is the simplest case where the y-variables are not all independent, and this subsection serves as an illustration. The potential is described in terms of three dilaton couplings α_1, α_2, and α_3. For simplicity, we take $\Lambda_i > 0$; the case with negative Λ_i can be analysed in a similar way. We can choose $\alpha_3 > 0$. Since $R = 1$, we have two independent y's, leaving one relation, which reads

$$(\Lambda_2)^{\alpha_1-\alpha_3}\,(y_2^2)^{\alpha_3-\alpha_1} = (\Lambda_1)^{\alpha_2-\alpha_3}\,(\Lambda_3)^{\alpha_1-\alpha_2}\,(y_1^2)^{\alpha_3-\alpha_2}\,(y_3^2)^{\alpha_2-\alpha_1}\,. \tag{6.44}$$

The analysis of critical points is analogous to the previous case, except for the extra feature of the relation above. There are three kinds of critical points (for the moment we leave aside the y-dependence)

$$
\begin{aligned}
&(i) \qquad y_i = 0\,, \quad x^2 = 1\,, \\
&(ii) \qquad y_i = y_j = 0\,,\ y_k = \sqrt{1-\frac{\alpha_k^2}{3}}\,,\ x = \frac{\alpha_k}{\sqrt{3}}\,, \quad i,j,k \quad \text{different} \\
&(iii) \qquad x = 0\,.
\end{aligned} \tag{6.45}
$$

A necessary condition for its existence is $\alpha_k^2 < 3$. However, this is not sufficient, since (6.44) only allows certain y's to be non-zero while the others are zero. For instance, with $\alpha_3 > \alpha_2 > \alpha_1$, having $y_1 = 0$ or $y_3 = 0$ implies $y_2 = 0$.

The third type of critical point is a de Sitter solution given by the following equations:

$$\alpha_1 y_1^2 + \alpha_2 y_2^2 + \alpha_3 y_3^2 = 0\,, \qquad y_1^2 + y_2^2 + y_3^2 = 1\,, \tag{6.46}$$

which can be rewritten as

$$\frac{\alpha_3-\alpha_1}{\alpha_3}\,y_1^2 + \frac{\alpha_3-\alpha_2}{\alpha_3}\,y_2^2 = 1\,. \tag{6.47}$$

This defines an ellipse for $\alpha_3 > \alpha_1\,,\alpha_2$. When substituting y_3, (6.44) also gives a curve in the (y_1, y_2)-plane, and the critical point is given as the intersection between these two curves.[6] For example, for $(\alpha_1,\alpha_2,\alpha_3) = (-1/2, 1/2, 3/2)$ and $\Lambda_i = 1$, the de Sitter critical point becomes $(y_1 = 0.78, y_2 = 0.52, y_3 = 0.34)$. Figure 6.4 shows the time-development of an interpolating solution for this case. One can see that the late-time behaviour indeed corresponds to the de Sitter critical point above.

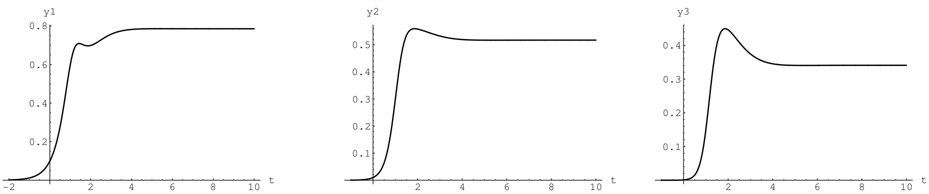


Figure 6.4: *The plots show $y_1(t)$, $y_2(t)$ and $y_3(t)$, respectively. As t increases, they tend towards the de Sitter critical point.*

6.5 Multi-exponential potentials from group manifolds

In this section we will consider specific cases that can be obtained by reducing pure gravity in seven dimensions over a three-dimensional group manifold. See appendix C for a basic

[6] However, it is only possible to give algebraic expressions of the solution for special values of the dilaton couplings.

definition of group manifolds. Since pure gravity in 7D can be embedded in 11D, the solutions have an M-theory origin. We will focus on the triple-exponential case.

Double exponential potentials can be obtained for certain truncations of reductions over type VIII and IX group manifolds [95]. This is equivalent to a trivial reduction over a circle followed by a reduction over a maximally symmetric 2D space with flux. The resulting potential is given by (6.40), with $d = 2$, and interpolating solutions correspond to reductions of S2-branes from six dimensions.

A triple exponential potential can be obtained from type VI_0 and VII_0 group manifolds and the result is [95]

$$V = \tfrac{1}{8}\, e^{-\sqrt{3}\varphi}\,(e^{\phi} \pm e^{-\phi})^2\,, \tag{6.48}$$

where the plus sign occurs for type VI_0 and the minus sign for type VII_0. We therefore have an example with $m = 3$ and $N = 2$, and the three dilaton couplings are

$$\vec{\alpha}_1 = (\sqrt{3}, 2)\,, \qquad \vec{\alpha}_2 = (\sqrt{3}, -2)\,, \qquad \vec{\alpha}_3 = (\sqrt{3}, 0)\,. \tag{6.49}$$

Note that only two of these are independent, hence this case falls into the $R < m$ class, and, more interestingly, we find the convex combination $\frac{1}{2}\vec{\alpha}_1 + \frac{1}{2}\vec{\alpha}_2 = \vec{\alpha}_3$, Therefore, a proper critical point with power-law behaviour is possible. The fact that we have linearly dependent $\vec{\alpha}_i$-vectors ($R < m$) is actually the case for most Bianchi class A types. For the present example, there will be two independent y-variables plus the relation $y_3 = \pm 2\, y_1 y_2$, but only y_1 and y_2 are needed.

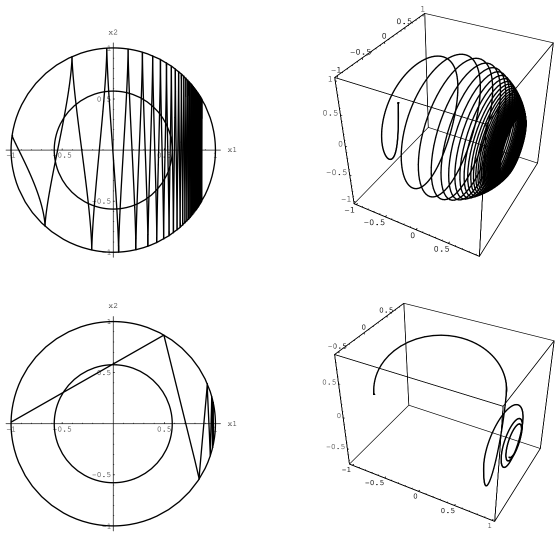


Figure 6.5: *Type VII_0. On the left-hand-side are shown the projection of two interpolating solutions on the (x_1, x_2)-plane. On the right-hand-side, the curves are shown on the 2-sphere defined by the Friedmann equation; the vertical axis is given by $y_1 - y_2$. The fact that the curves do not reach the attractor is due to the finite computation time.*

For the sake of illustration, $y_1 \pm y_2$ can be used as a variable, such that any solution will be given by points or curves on a 2-sphere (the upper half in the case of the plus sign, since the y's are positive). Interestingly, the dilaton couplings are such that most of the critical points from the previous section do not exist. In fact, for type VI$_0$, we are just left with the equator solutions and for type VII$_0$ we have the equator and an infinite set of proper solutions:

$$\begin{array}{lll} x^2 = 1\,, & y_1 = y_2 = 0 & \text{type} \quad \text{VI}_0\,, \\ \left.\begin{array}{l} x^2 = 1\,, \\ (x_1, x_2) = (1, 0)\,, \end{array}\right. & \left.\begin{array}{l} y_1 = y_2 = 0 \\ y_1 = y_2 \end{array}\right\} & \text{type} \quad \text{VII}_0\,. \end{array} \tag{6.50}$$

By studying the derivatives of the coordinates, it can be shown that the following points are attractors

$$\begin{array}{lll} (x_1, x_2) = (1, 0) & y_1 = y_2 = 0 & \text{type} \quad \text{VI}_0\,, \\ (x_1, x_2) = (1, 0) & y_1 = y_2 & \text{type} \quad \text{VII}_0\,. \end{array} \tag{6.51}$$

Thus, the latter is not unique since the y_i-values are arbitrary. The solution corresponds to (6.32), which is possible because of the convex combination: $\vec{\alpha}_3 = (\vec{\alpha}_1 + \vec{\alpha}_2)/2$. In both cases any interpolating solution will end in the point $(1, 0, 0)$ on the 2-sphere. In the VII$_0$ case, the y-value will be determined by the initial conditions. The sign of $\dot{x}_1$ is always positive. When projected on the (x_1, x_2)-plane, any curve will therefore move from left to right.

A stability analysis leads to the result that only the part of the equator with $x_1 < -1/7$ is repelling. Thus, any interpolating solution can start on this part and will end in $(1, 0, 0)$.

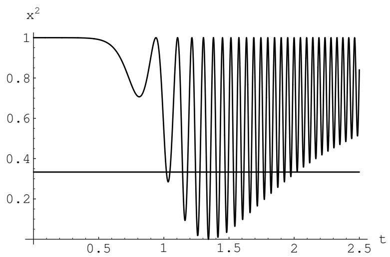


Figure 6.6: *Type VII$_0$. An example of a (t, x^2)-plot with $n < 1$.*

A couple of typical curves for interpolating solutions with different initial conditions are depicted on figure 6.5, for type VII$_0$. In this case, any curve will spiral around the 2-sphere towards the attractor. The projection on the (x_1, x_2)-plane produces a curve which bounces off the boundary of the unit circle. The inner disc, corresponding to $x^2 < 1/3$, yields phases of accelerated expansion. Depending on the initial conditions, the number of such phases can be as high or low as desired. For the two cases on figure 6.5, the numbers are 16 and 1, respectively. Even with a large number of accelerating phases, the number of e-foldings is of order 1; therefore, these models are not well suited for inflation. The numerical solutions use $t = \ln(a)$ as time parameter. The number of e-foldings is given by

$$n = \ln\Big(\frac{a(\tau_2)}{a(\tau_1)}\Big) = \ln\Big(\frac{e^{t_2}}{e^{t_1}}\Big) = \Delta t\,, \tag{6.52}$$

and its order of magnitude can easily be read off from a (t, x^2)-plot as the sum of the t-intervals where $x^2 < 1/3$. An example is given in figure 6.6.

For type VI_0, the situation is slightly different, since $y_1 + y_2$ is always positive; this confines the curves to the upper half of the 2-sphere. On figure 6.7, the curves still move towards the attractor in an oscillatory manner, but now without crossing the equator (though they can get arbitrarily close). For this case, there can only be one or no phase of accelerated expansion.

The interpolating solutions above correspond to reductions of exotic S2-branes in five dimensions, or equivalently, exotic $\mathrm{S}(D-3)$-branes in D dimensions. However, the solutions were found numerically, and we have not been able to obtain exact expressions for these exotic S-branes.

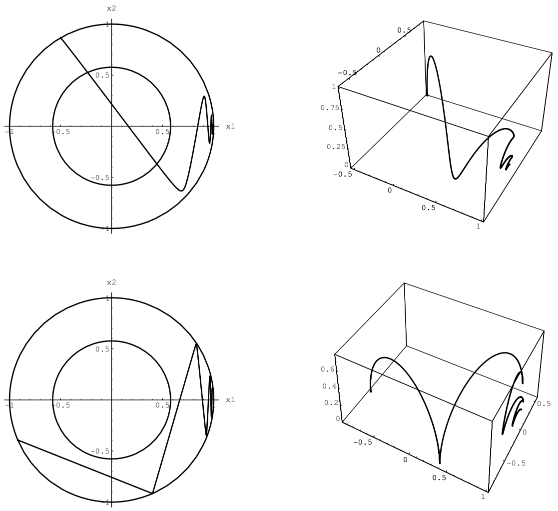


Figure 6.7: *Type VI_0. On the left-hand-side are shown the projections of two interpolating solutions on the (x_1, x_2)-plane. On the right-hand-side, the curves are shown on the 2-sphere defined by the Friedmann equation; the vertical axis is given by $y_1 + y_2$.*

6.6 Discussion

In this chapter we have considered cosmological models for an arbitrary number of scalars with arbitrary multi-exponential potentials. Using a special set of variables, the equations were written as autonomous dynamical systems, and this allowed us to determine the critical points in complete generality. We found that the nature of these critical points depends strongly on the rank R of the matrix α_{iI}. The rank also determines the number of decoupled scalars.

In the case $R = m$, both the proper and non-proper critical points are power-law solutions, and there are no de Sitter solutions. In the $R < m$ case the opposite behaviour was found. Proper power-law solutions are only possible in special cases, where the $\vec{\alpha}_i$-vectors are linearly dependent in a specific way, but the de Sitter solutions are very generic. For the non-proper solutions, this depends on whether the "truncated" potential has $R = m$ or $R < m$. We also found a new property of these systems, namely the possibility of proper critical points that are not unique. A special case was realized in section 5, where we have an infinite set of these.

It should be emphasized that the non-proper critical points are as important as the proper ones for understanding the interpolating solutions, even though they have not often been considered in the literature. In this respect, using the techniques of autonomous systems is more fruitful than simply looking for power-law solutions to the equations of motion.

It should be pointed out that in our solutions the scalars generically have run-away behaviour. The only exceptions are the de Sitter critical points, since these correspond to extrema of the potential and accordingly stabilize the values of the scalars. This is important in the context of spontaneous decompactification [134] or stabilization of dilaton and volume moduli [33]

In section 4 and 5, we provided several examples of double and triple exponential potentials. We presented the critical points and illustrated the interpolating solutions using numerical calculations. In particular, we found examples with an arbitrarily high number of phases of accelerated expansion. However, the number of e-foldings turned out to be of order one, so these models do not seem to be relevant for inflation. They might, however, be relevant for present-day acceleration and they might help solve the cosmic coincidence problem.

The numerical solutions found in section 5 for the systems obtained from reductions over group manifolds of type VI_0 and type VII_0 correspond to the reduction of exotic S2-branes in five dimensions. The two solutions belong to a set of three different solutions that can be obtained via twisted circle reductions. The third solution can be obtained from a reduction over the type II group manifold and corresponds to the reduction of a fluxless S2-brane. The existence of three classes of S-branes is similar to the cases of 7-branes in ten dimensions [42] and the non-extremal D-instantons we studied in chapter 3, and it is reminiscent of the global $SL(2,\mathbb{R})$-symmetry of the higher dimensional theory. It would be interesting to see whether it would be possible to find exact solutions for the exotic S-branes.

Recently, an elegant framework for arbitrary potentials has been developed, where the solutions correspond to geodesics in an augmented target space [135]. One of the key ingredients is the importance of systems whose late-time behaviour is governed by single exponential potentials [94]. In our analysis, these solutions asymptote to the special class of non-proper critical points where all y's but one vanish. However, we have shown that multi-exponential potentials have solutions, where the asymptotics cannot be governed by a single term in the potential. Specific examples are given by the cases of assisted inflation [108] and generalized assisted inflation [122], where each term in the potential contributes.

Comments on some possible extensions of this work would be in order. First of all, we have only considered flat universes, and it is certainly possible to extend this formalism to the spatially curved cases.

Secondly, we could also add matter, in the form of a barotropic fluid. This could play a rôle in solving the cosmic coincidence problem. The authors of [107] showed that a system with one scalar and a barotropic fluid can have attractor solutions that are neither scalar-field dominated nor matter dominated, but both at the same time. These are the so-called *scaling solutions* where dark energy and matter coexist. This may lead to a dynamical solution to the cosmic coincidence problem. In other words, a scaling solution dynamically explains why dark energy and matter have comparable energy densities, in the present universe. However, in the one-scalar system studied in that paper, the de Sitter attractor and the scaling attractor are mutually exclusive. Given a dilaton coupling, only one can exist. Although it is not known whether the universe will be eternally de Sitter, some string theory based scenarios rely upon stable de Sitter vacua that the universe 'tunnels' out of by quantum mechanical effects. Hence, it could be interesting to have a combined scaling-de Sitter attractor. By having a more complex system than the one-scalar Lagrangian, one might obtain a compromise between a pure de Sitter and a scaling solution. In [106] and [122], scalars were added to make systems with assisted and generalized assisted inflation, respectively. In [136], spatial curvature was added in the one-scalar case. The potentials we have considered here, however, allow for new de Sitter attractors. They also allow for oscillatory behavior; i.e. some of the solutions are periodic in time. An oscillatory universe might also explain cosmic coincidence. The chances of living in a period where dark energy and matter coexist are greater in a universe that forever oscillates between dark energy and matter domination. In this same spirit, we hope to extend the search for scaling solutions to the most general exponential potential with spatial curvature and a barotropic fluid and report on it in a future publication.

Thirdly, we could consider non-flat scalar manifolds. Finally, we could consider other specific numerical examples with other values of m and N and special dilaton couplings which arise from dimensional reductions of string/M-theory.

This concludes the second part of this thesis, which covered the topic of cosmological solutions. Both this and the previous part dealt with scalar-gravity solutions that depend on one parameter. D-instantons depend on one spatial direction, whereas cosmologies depend on time only. Both types of solution have the generic property of interpolating between 'trivial' configurations: The wormhole solution interpolates between two regions with flat spacetime and constant fields, and the cosmologies interpolate between power-law and/or de Sitter spacetimes. In the next and final chapter of this thesis, we will establish links between D-instantons and cosmologies. We will actually show two ways in which these objects can correspond to each other. First, we will see that they can sometimes be related to each other via Wick rotations. Then, we will show that by means of a paradigm shift, both types of solutions can be regarded as trajectories of a particle in a fictitious spacetime, a target space parametrized by the scalars in the Lagrangian.

Chapter 7

Link between D-instantons and Cosmology

Throughout this thesis, we have been studying two different kinds of scalar-gravity field configurations: D-instantons and cosmological solutions. We also briefly looked at solitons. We pointed out in chapter 2 that instantons and solitons can be equivalent, if certain conditions are met by the system they are in. In chapter 3, we studied the specific correspondence between black hole solutions and D-instantons, and the requirements for the correspondence to hold. But there is a strikingly simpler and more obvious fact that ties D-instantons, black holes, and cosmological solutions together. They all depend on one coordinate, be it space-like or time-like, and they all interpolate between 'trivial' configurations. More specifically, the wormhole geometry of the non-extremal D-instanton with $q^2 > 0$ interpolates between two flat Euclidean spaces. Cosmological solutions such as the ones we studied in chapters 5 and 6 interpolate in time between power-law regimes, or between power-law and de Sitter spaces with non-trivial behavior in between, such as transient acceleration. In this chapter, we will pursue the similarity between D-instantons and cosmological solutions carried by scalar fields in detail. We will do so in two ways. In the first section, we will relate some of these solutions to each other via the Wick rotation. In the second section, we will take a different perspective on the degrees of freedom we are studying. We will view the D-instanton and cosmological solutions as trajectories in a scalar manifold, an abstract target space, if you will. That will enable us to present these solutions in a mathematically unified way. It will even suggest a way of pasting together a cosmological solution and an instanton solution, as though they were part of the same phenomenon.

7.1 Wick rotation

WARNING: The following section contains passages with explicit Wick rotations that may not be suitable for self-respecting mathematicians. Parental discretion is advised.

Let us begin by reviewing the non-extremal D-instanton solution from chapter 3, which I will rewrite here for the reader's convenience. The Euclidean Lagrangian density is the following:

$$\mathcal{L} = R - \tfrac{1}{2}\,(\partial\phi)^2 + \tfrac{1}{2}\,e^{b\,\phi}\,(\partial\chi)^2\ , \tag{7.1}$$

where, ϕ are χ scalars. The 'wrong' sign of the kinetic term for χ is explained in subsection 3.6.1. We take the Ansatz of a conformally flat metric with maximal spherical symmetry. We also assume that all fields respect the spherical symmetry. This yields the following solution:

$$ds^2 = \left(1 - \frac{\mathbf{q^2}}{r^{2\,(D-2)}}\right)^{2/(D-2)} \left(dr^2 + r^2\, d\Omega^2_{S^{D-1}}\right)\,, \tag{7.2}$$

$$e^{b\,\phi(r)} = \left(\frac{q_-}{\mathbf{q}}\ \sinh\left(H(r) + C_1\right)\right)^2\,, \tag{7.3}$$

$$\chi(r) = \frac{2}{b\,q_-}\ \left(\mathbf{q}\ \coth\left(H(r) + C_1\right) - q_3\right)\,. \tag{7.4}$$

with

$$H(r) = b\,c\,\mathrm{arctanh}\left(\frac{\mathbf{q}}{r^{D-2}}\right)\,, \tag{7.5}$$

where $\mathbf{q^2}, q_-, q_3$ and C_1 are integration constants and

$$c = \sqrt{\frac{2\,(D-1)}{(D-2)}}\,. \tag{7.6}$$

As we saw previously, $\mathbf{q^2}$ can be positive, negative or zero. Therefore, as we can see, everything depends on one coordinate r. At this point, the reader should feel the irresistible temptation to Wick rotate this solution to see if that yields a cosmological configuration. This has been explored in [137], but I will do it in my own notation here. The first step is to make 'r' timelike by letting $r \to i\,t$. This takes care of the dr^2 term in the metric, but messes up the spherical part by creating a minus sign. To fix this, let us rewrite the spherical metric as follows:

$$d\Omega^2_{S^{D-1}} = d\theta^2 + \sin^2(\theta)\, d\Omega^2_{S^{D-2}}\,. \tag{7.7}$$

The Wick rotation created an overall sign in front of this metric, so to fix it, we let $\theta \to i\,\psi$:

$$d\theta^2 + \sin^2(\theta)\, d\Omega^2_{S^{D-2}} \to -d\psi^2 - \sinh^2(\psi)\, d\Omega^2_{S^{D-2}} = -d\mathbb{H}^2_{D-1}\,, \tag{7.8}$$

where $\mathbb{H}$ stands for a hyperbolic space. The end result is the following metric:

$$ds^2 = \left(1 - \frac{\tilde{\mathbf{q}}^2}{t^{2\,(D-2)}}\right)^{2/(D-2)} \left(-dt^2 + t^2\, d\mathbb{H}^2_{D-1}\right)\,, \tag{7.9}$$

where $\tilde{\mathbf{q}}^2 = (-1)^{(D-2)}\,\mathbf{q}^2$. This is indeed a cosmological solution. Specifically, it is an FLRW metric with $k = -1$. But what about ϕ and χ? Those are less straightforward to study, but as we will see we can already gather one qualitative piece of information from them. Let us first write down the Lagrangian density for the Lorentzian system to which this cosmological solution belongs:

$$\mathcal{L} = R - \tfrac{1}{2}\,(\partial\phi)^2 - \tfrac{1}{2}\,e^{b\,\phi}\,(\partial\chi)^2\,. \tag{7.10}$$

Notice that we are now using the 'normal' sign for the kinetic term for χ. Hence, in order to establish a relation between an instanton and a cosmology, we must effectively multiply χ by i. If we take the metric (7.9) as an Ansatz for this system and fill it into the time component of the Einstein equation, we get the following:

$$R_{tt} = \frac{4\,(D-1)\,(D-2)}{r^{2\,(D-1)}\left(1 - \frac{\tilde{\mathbf{q}}^2}{t^{2\,(D-2)}}\right)^2}\,\tilde{\mathbf{q}}^2 = \tfrac{1}{2}\,\dot{\phi}^2 + \tfrac{1}{2}\,e^{b\,\phi}\,\dot{\chi}^2\,. \tag{7.11}$$

The right-hand side is positive definite, therefore, $\tilde{\mathbf{q}}^2$ *must* be positive. Hence, not all three classes of D-instantons (i.e. $\mathbf{q}^2$ positive, negative and zero) can be Wick rotated to a cosmological solution. Not even the extremal D-instanton has a cosmological partner. The only class that can be Wick rotated is the one with $\mathbf{q}^2\,(-)^{D-2} > 0$. This is obviously a dimension-dependent condition. The actual process of Wick rotating the solutions for the scalars is less obvious. The idea is that χ has to get mutliplied by an i, whereas the dilaton should remain unaffected. This is accomplished by letting $C_1 \to C_1 + i\,\pi/2$, and $q_- \to i\,q_-$. The result is the following:

$$e^{b\,\phi(r)} = \left(\frac{q_-}{\tilde{\mathbf{q}}}\,\cosh\left(H(r) + C_1\right)\right)^2\,, \tag{7.12}$$

$$\chi(r) = \frac{2}{b\,q_-}\,\left(\tilde{\mathbf{q}}\,\tanh\left(H(r) + C_1\right) - q_3\right)\,, \tag{7.13}$$

with

$$H(r) = b\,c\,\text{arctanh}\left(\frac{\tilde{\mathbf{q}}}{r^{D-2}}\right)\,. \tag{7.14}$$

7.2 Target space interpretation

In this section, we are going to investigate another parallelism between *axionic* instantons and cosmologies. The idea is to regard all fields, including the non-constant part of the metric, as coordinates in a fictitious *target space*. Because instantons and cosmologies both depend on only one parameter, they will be interpreted as trajectories of a particle in the target space.

This section is based on a collaboration with E. Bergshoeff, D. Roest, J. Russo, and P.K. Townsend, entitled *Cosmological D-instantons and Cyclic Universes*, [138]. It is organized as follows: first, we will present the general system and Ansatz we want to solve and dimensionally reduce it to one dimension. In subsection 7.2.2, we will introduce the 'Liouville' gauge, in which will allow us to view our solutions as trajectories in two-dimensional target spaces defined by the two scalar fields. In subsection 7.2.3, we will introduce the 'Milne-Rindler' gauge, which will also view the non-trivial part of the metric as a target space coordinate. In this new three-dimensional target space, we will be able to present instantons and cosmologies in a unified way, as trajectories of a particle.

7.2.1 Ansatz and reduction to one dimension

The Lagrangians we will be studying can be summarized as follows:

$$\mathcal{L} = R - \tfrac{1}{2}\,(\partial\phi)^2 + \epsilon\,\tfrac{1}{2}\,e^{b\,\phi}\,(\partial\chi)^2\,, \tag{7.15}$$

with $\epsilon = \pm 1$ for Euclidean and Lorentzian signature respectively. To investigate cosmological solutions of our model, or to find instanton solutions of its Euclidean version, we make the Ansatz

$$ds^2 = \epsilon\,(e^{\alpha\varphi} f)^2\, d\lambda^2 + e^{2\alpha\varphi/(d-1)} d\Sigma_k^2\,, \qquad \phi = \phi(\lambda)\,, \qquad \chi = \chi(\lambda)\,, \tag{7.16}$$

where f is an arbitrary function of λ, and

$$\alpha = \sqrt{\frac{d-1}{2(d-2)}}\,. \tag{7.17}$$

The $(d-1)$-metric $d\Sigma_k^2$ is (at least locally) a maximally symmetric space of positive ($k = 1$), negative ($k = -1$) or zero ($k = 0$) curvature. One can choose coordinates such that

$$d\Sigma_k^2 = (1 - kr^2)^{-1} dr^2 + r^2 d\Omega_{d-2}^2\,, \tag{7.18}$$

where $d\Omega_{d-2}^2$ is an SO$(d-1)$-invariant metric on the unit $(d-2)$-sphere. This Ansatz constitutes a consistent reduction of the original degrees of freedom to a three-dimensional subspace, the ‘augmented target space’, with coordinates (φ, ϕ, χ). The full equations of motion reduce to a set of equations that can themselves be derived by variation of the time-reparametrization invariant effective action

$$I = \tfrac{1}{2} \int d\lambda\, \left\{ f^{-1}\left(\epsilon\dot{\varphi}^2 - \epsilon\dot{\phi}^2 + e^{b\phi}\dot{\chi}^2\right) + 2k(d-1)(d-2)\,f e^{\varphi/\alpha} \right\}\,, \tag{7.19}$$

where the overdot indicates differentiation with respect to λ. For $\epsilon = -1$ we can interpret λ as a time coordinate related to the time t of FLRW cosmology in standard coordinates by

$$dt \propto e^{\alpha\varphi} f d\lambda\,. \tag{7.20}$$

For $\epsilon = 1$ the metric has Euclidean signature and we can interpret λ as imaginary time.

If we interpret all fields as being coordinates of a particle’s world-line in some target space, then we notice that the scalars ϕ and χ parametrize a two-dimensional hyperbolic space:

$$ds_T{}^2 = -\epsilon\, d\phi^2 + e^{b\,\phi}\, d\chi^2\,. \tag{7.21}$$

For $\epsilon = -1$, this is $\mathbb{H}_2$, the two-sheeted hyperboloid with Euclidean signature, in Poincaré coordinates, which are globally defined. For $\epsilon = +1$, however, this is a Lorentzian one-sheeted hyperboloid dS_2 in Poincaré coordinates, which are *not* gobally defined. They only cover half of the surface. To treat both signatures on equal footing, it is therefore convenient to switch to coordinates of the target space that are globally defined. This is done by defining new scalar field variables (ψ, θ) by

$$\begin{aligned} e^{(b/2)\phi} &= e^{\psi}\cos^2(\theta/2) - \epsilon e^{-\psi}\sin^2(\theta/2)\,, \\ e^{(b/2)\phi}\chi &= b^{-1}\left(e^{\psi} + \epsilon e^{-\psi}\right)\sin\theta\,, \end{aligned} \tag{7.22}$$

which yields the following target space metrics:

$$ds_T{}^2 = \begin{cases} -d\psi^2 + \cosh^2(\psi)\, d\theta^2 & \text{for } \epsilon = 1 \\ +d\psi^2 + \sinh^2(\psi)\, d\theta^2 & \text{for } \epsilon = -1 \end{cases}\,. \tag{7.23}$$

We can now recognize the first metric as that of a dS_2, by comparison with the dS_3 metric (4.41) we introduced in chapter 4. The second metric is the usual one for a two-dimensional hyperboloid. The new effective action is

$$I = \tfrac{1}{2} \int d\lambda \Big\{ \frac{4}{b^2} f^{-1} \left[\frac{b^2}{4} \epsilon \dot{\varphi}^2 - \epsilon \dot{\psi}^2 + \tfrac{1}{4} \left(e^{\psi} + \epsilon e^{-\psi} \right)^2 \dot{\theta}^2 \right] + + 2k(d-1)(d-2) f e^{\varphi/\alpha} \Big\} . \tag{7.24}$$

Introducing the new scale-factor variable η by

$$\eta^{\gamma} = 2\gamma(d-1)\, e^{\varphi/(2\alpha)} , \tag{7.25}$$

where

$$\gamma = 1/(b\alpha) , \tag{7.26}$$

we arrive at the action

$$I = \tfrac{1}{2} \int d\lambda \left\{ \frac{4}{b^2} f^{-1} \left[\epsilon (\dot{\eta}/\eta)^2 - \epsilon \dot{\psi}^2 + \tfrac{1}{4} \left(e^{\psi} + \epsilon e^{-\psi} \right)^2 \dot{\theta}^2 \right] + \frac{b^2}{4} k f \eta^{2\gamma} \right\} . \tag{7.27}$$

We remark, for future reference, that the Ansatz (7.16) leads to $\gamma = 2/3$ for $d = 10$ IIB supergravity.

Because of the time-reparametrization invariance, we are free to choose the function f; each choice of f corresponds to some choice of time parameter. There are two choices that are particularly convenient, and we now consider them in turn.

7.2.2 The 'Liouville' gauge

The simplest way to proceed for general b is to make the gauge choice

$$f = 4/b^2 . \tag{7.28}$$

From (7.27) one sees that the effective Lagrangian in this gauge is

$$L = \tfrac{1}{2} \left[-\epsilon \dot{\psi}^2 + \tfrac{1}{4} \left(e^{\psi} + \epsilon e^{-\psi} \right)^2 \dot{\theta}^2 \right] + \tfrac{1}{2} \left[\epsilon (\dot{\eta}/\eta)^2 + k \eta^{2\gamma} \right] . \tag{7.29}$$

Apart from the constraint, the dynamics of the motion on the target space, which is manifestly geodesic, is now separated from the dynamics of the scale factor, which is determined by a equation of Liouville-type; for this reason we will call this choice of gauge the "Liouville gauge".

As $SL(2;\mathbb{R})$ is the isometry group of both $\mathbb{H}_2$ (the target space of the Lorentzian action) and dS_2 (the target space of the Euclidean action), there is a conserved $SL(2;\mathbb{R})$ 'momentum' ℓ^{μ}, and the geodesics are such that

$$\dot{\psi}^2 - \epsilon \tfrac{1}{4} \left(e^{\psi} + \epsilon e^{-\psi} \right)^2 \dot{\theta}^2 = \ell^2 . \tag{7.30}$$

The constraint (f equation of motion) is

$$(\dot{\eta}/\eta)^2 = \ell^2 + k\epsilon\, \eta^{2\gamma} . \tag{7.31}$$

We now present the solutions of the equations of motion of (7.29) subject to the constraint (7.30) and (7.31), first for the target space fields and then for the scale factor.

Target space geodesics

Geodesics on the $\mathbb{H}_2$ ($\epsilon = -1$) or dS_2 ($\epsilon = 1$) target space are solutions of the field equations of (7.29) for ψ and θ subject to (7.30) and can be classified as follows, according to whether ℓ^2 is positive, negative or zero:

- $\ell^2 > 0$. For $\epsilon = 1$ the solution is

$$\begin{aligned} \sinh\psi &= \pm\sqrt{1+\frac{q_-^2}{\ell^2}}\,\sinh[\ell\,(\lambda-\lambda_0)] \\ \tan(\theta-\theta_0) &= \pm\frac{q_-}{\ell}\,\tanh[\ell\,(\lambda-\lambda_0)]\,, \end{aligned} \tag{7.32}$$

  for constants λ_0, θ_0 and q_- (this being the integration constant for the super-extremal D-instanton of [43]). For $\epsilon = -1$ the solution is

$$\begin{aligned} \cosh\psi &= \sqrt{1+\frac{q_-^2}{\ell^2}}\,\cosh[\ell\,(\lambda-\lambda_0)] \\ \tan(\theta-\theta_0) &= \pm\frac{q_-}{\ell}\,\coth[\ell\,(\lambda-\lambda_0)]\,. \end{aligned} \tag{7.33}$$

  In the special case that $q_- = 0$ these solutions simplify, for either choice of the sign ϵ, to

$$\psi = \pm\ell(\lambda-\lambda_0)\,, \qquad \theta = \theta_0\,, \qquad (\epsilon = \pm 1). \tag{7.34}$$

- $\ell^2 < 0$. In this case only $\epsilon = 1$ is possible, and the solution is

$$\begin{aligned} \sinh\psi &= \pm\sqrt{\frac{q_-^2}{(-\ell^2)}-1}\,\sin\left[\sqrt{-\ell^2}\,(\lambda-\lambda_0)\right] \\ \tan(\theta-\theta_0) &= \pm\frac{q_-}{\sqrt{-\ell^2}}\,\tan\left[\sqrt{-\ell^2}\,(\lambda-\lambda_0)\right]. \end{aligned} \tag{7.35}$$

- $\ell^2 = 0$. The only solution for $\epsilon = -1$ in this case is the trivial one for which both ψ and θ are constant. For $\epsilon = 1$ the solution is

$$\sinh\psi = \pm q_-\,(\lambda-\lambda_0)\,, \qquad \tan(\theta-\theta_0) = \pm q_-\,(\lambda-\lambda_0)\,. \tag{7.36}$$

It should be noted that, in each case, the $\pm$ signs for ψ and θ can be chosen independently. We should note that, for $\epsilon = 1$ (dS_2), there are three classes of solutions, with ℓ^2 positive, zero and negative, whereas for $\epsilon = -1$, there is only one class with $\ell^2 > 0$. This is because dS_2, being Minkowskian, has a light-cone structure. It admits, space-like, light-like and time-like solutions. We can interpret ℓ^2 as the momentum squared (i.e. $-m^2$), of the particle. $\mathbb{H}_2$, on the other hand, does not have a light-cone structure.

One interesting consequence of writing our instanton solutions in terms of these global coordinates, is that the singularities of the dilaton and axion that we previously encountered in chapter 3 go away. In this target space interpretation, those singularities are mere coordinate singularities, signaling that the particle's world-line has departed from the Poincaré coordinate patch. It has gone over to the half of the hyperboloid that is not covered by these coordinates. The true physical meaning of this resolution of the singularities, however, remains to be discovered.

The scale factor

We next turn to the constraint (7.31). Given ℓ^2, this determines η as follows

- $\ell^2 > 0$.

$$\eta^{2\gamma} = \eta_0^{2\gamma} \exp(\pm 2\ell\gamma\lambda)\,, \qquad (k=0)\,, \tag{7.37}$$
$$\eta^{2\gamma} = \frac{\ell^2}{\sinh^2(\ell\gamma\lambda)}\,, \qquad (k\epsilon = 1)\,, \tag{7.38}$$
$$\eta^{2\gamma} = \frac{\ell^2}{\cosh^2(\ell\gamma\lambda)}\,, \qquad (k\epsilon = -1)\,, \tag{7.39}$$

  for some constant η_0. Note that all $k = \pm 1$ trajectories are asymptotic to some $k = 0$ trajectory near $\eta = 0$, as expected since the σ-model matter satisfies the strong energy condition.

- $\ell^2 < 0$. In this case there is a solution only for $k = \epsilon = 1$:

$$\eta^{2\gamma} = \frac{-\ell^2}{\sin^2\left(\gamma\sqrt{-\ell^2}\,\lambda\right)}\,, \qquad (k = \epsilon = 1)\,. \tag{7.40}$$

- $\ell^2 = 0$. In this case there is a solution only for $k\epsilon \geq 0$. If $k = 0$ then η is constant. Otherwise

$$\eta^{2\gamma} = 1/(\gamma\lambda)^2\,, \qquad (k\epsilon = 1)\,. \tag{7.41}$$

For $\epsilon = k = 1$ these solutions yield the super-extremal ($\ell^2 > 0$), sub-extremal ($\ell^2 < 0$) and extremal ($\ell^2 = 0$) D-instantons of [43]. For $\epsilon = -1$ they yield FLRW cosmologies; from (7.20) we see that the standard FLRW time t is related to the parameter λ by

$$dt \propto \eta^{2\gamma\alpha^2}\, d\lambda\,. \tag{7.42}$$

Given one of above solutions for $\eta^{2\gamma}$ as a function of λ we can determine λ as a function of t and hence η as a function of t. Of most interest here is the behaviour near $\eta = 0$. For example, for $\ell^2 > 0$ we have

$$\eta \sim \eta_0\, e^{-\ell\lambda}\,, \tag{7.43}$$

for $\lambda \to \infty$, as $\eta \to 0$. This yields (for a choice of integration constant such that $t \to 0$ as $\lambda \to \infty$)

$$-t \propto e^{-2\gamma\alpha^2\ell\lambda}\,. \tag{7.44}$$

Given that we start with a cosmological solution for negative t, this shows that a big-crunch singularity will be approached as $t \to 0$. By considering the behaviour as $\lambda \to -\infty$ we may similarly deduce that a cosmological solution for positive t must have had a big-bang singularity at $t = 0$. In other words, cosmologies with $\ell^2 > 0$ are incomplete in the sense that they have a beginning or an end (or both) at finite t. We shall make a suggestion in the following subsection, of how they can be completed.

7.2.3 The 'Milne-Rindler' gauge

We will now upgrade the approach we took in the previous subsection by augmenting the two-dimensional target space to a three-dimensional one. We will do so by considering the single degree of freedom of the spacetime metric as another target space coordinate.

Returning to (7.27), we make the new gauge choice

$$f = \frac{4}{b^2\, \eta^2}\,. \tag{7.45}$$

As the possible choices of f are related by a redefinition of the independent variable, we will need to distinguish the independent variable in this gauge from the parameter λ previously used. Let us call the new independent variable τ; it is related to λ through the differential equation

$$d\tau = \eta^2(\lambda)\, d\lambda\,, \tag{7.46}$$

which can be solved for $\tau(\lambda)$ given any of the scale factor solutions $\eta(\lambda)$ presented above.

In the gauge (7.45) the action is

$$I = \int d\tau\, L_\tau\,, \tag{7.47}$$

where[1]

$$L_\tau = \tfrac{1}{2}\epsilon \left(\frac{d\eta}{d\tau}\right)^2 + \frac{1}{2}\eta^2 \left[-\epsilon \left(\frac{d\psi}{d\tau}\right)^2 + \tfrac{1}{4}\left(e^{\psi} + \epsilon e^{-\psi}\right)^2 \left(\frac{d\theta}{d\tau}\right)^2\right] + \frac{k}{2}\,\eta^{2\gamma-2}\,. \tag{7.48}$$

We observe that for $\epsilon = -1$ the kinetic term is that of a particle in a 3-dimensional Minkowski spacetime in Milne coordinates. However, for $\epsilon = 1$, this kinetic term is again that of a particle in 3-dimensional Minkowski spacetime, only this time in Rindler coordinates. See 4.1.3 for a discussion on those two coordinates systems for four-dimensional Minskowski. We will call this choice of gauge the 'Milne-Rindler' gauge. We can unify both actions ($\epsilon = \pm 1$), by going to Cartesian coordinates, since the latter are globally defined in Minkowski spacetime. The new field variables X_μ ($\mu = 0, 1, 2$) are

$$\begin{aligned} X_0 &= \pm\frac{1}{2}\eta\left(e^{\psi} - \epsilon e^{-\psi}\right)\,, \\ X_1 &= \pm\frac{1}{2}\eta\left(e^{\psi} + \epsilon e^{-\psi}\right)\cos\theta\,, \\ X_2 &= \pm\frac{1}{2}\eta\left(e^{\psi} + \epsilon e^{-\psi}\right)\sin\theta\,. \end{aligned} \tag{7.49}$$

[1]Note that $L_\tau d\tau = L d\lambda$, where L is the lagrangian in the gauge used previously.

Note that

$$X^2 \equiv -{X_0}^2 + {X_1}^2 + {X_2}^2 = \epsilon\eta^2 \,. \tag{7.50}$$

Since η^2 is positive, it follows that $X^2 < 0$ when $\epsilon = -1$, and $X^2 > 0$ when $\epsilon = 1$. The $X^2 < 0$ region is the Milne region of Minkowski space and cosmological solutions are trajectories in this space. Generic trajectories reach $\eta = 0$ at finite FLRW time, corresponding to a cosmological singularity. However, the hypersurface $\eta = 0$ is just the Milne horizon, and the singularity at the Milne horizon disappears in the cartesian coordinates X_μ. The trajectory can therefore be smoothly continued through the Milne horizon *in cartesian coordinates* into to the Rindler region, in which $X^2 > 0$, where we need $\epsilon = 1$. Thus, on passing through the Milne horizon, a cosmological trajectory becomes an instanton (and vice-versa).

The Milne-Rindler gauge Lagrangian L_τ in cartesian coordinates is

$$L_\tau = \tfrac{1}{2}\left[(dX/d\tau)^2 + k\left(\epsilon X^2\right)^{\gamma-1}\right] . \tag{7.51}$$

The constraint is now

$$(dX/d\tau)^2 = k\left(\epsilon X^2\right)^{\gamma-1} \,. \tag{7.52}$$

We thus have a problem analogous to that of a particle of zero energy in a central potential, with conserved SL(2; $\mathbb{R}$) "angular momentum"

$$\ell^\mu = \varepsilon^{\mu\nu\rho} X_\nu (dX_\rho/d\tau) \,. \tag{7.53}$$

The target space and the scale factor solutions given previously can now be combined into a single solution for X_μ. For example, for $\ell^2 > 0$, the solutions are

$$X_\mu = \begin{cases} \pm\eta\left(s_\mu \sinh(\ell\lambda) + c_\mu \cosh(\ell\lambda)\right) , & \epsilon = 1\,, \\ \pm\eta\left(s_\mu \cosh(\ell\lambda) + c_\mu \sinh(\ell\lambda)\right) , & \epsilon = -1\,, \end{cases} \tag{7.54}$$

where

$$\begin{aligned} s_0 &= \sqrt{1+a^2}\cosh(\ell\lambda_0)\,, \qquad a \equiv q_-/\ell\,, \\ c_0 &= -\sqrt{1+a^2}\sinh(\ell\lambda_0)\,, \\ c_1 &= \cosh(\ell\lambda_0)\cos(\theta_0) + a\sinh(\ell\lambda_0)\sin(\theta_0)\,, \\ s_1 &= -\sinh(\ell\lambda_0)\cos(\theta_0) - a\cosh(\ell\lambda_0)\sin(\theta_0)\,, \\ c_2 &= -a\sinh(\ell\lambda_0)\cos(\theta_0) + \cosh(\ell\lambda_0)\sin(\theta_0)\,, \\ s_2 &= a\cosh(\ell\lambda_0)\cos(\theta_0) - \sinh(\ell\lambda_0)\sin(\theta_0)\,. \end{aligned} \tag{7.55}$$

Note that $(c_\mu \pm s_\mu)$ is null.

In [138], this coordinate system is used to 'continue' cosmological solutions into instanton solutions by passing through the target space Milne horizon. See figure 7.1, for the case where $\gamma = 1$. In this case, the Lagrangian simplifies tremendously, as all trajectories become goedesics in the three-dimensional Minkowski spacetime. The $\gamma \neq 1$ have a central potential, which complicates the picture. An interesting idea that has not been experimented with, would be to eliminate such a potential by augmenting the target space to a four-dimensional one with non-trivial curvature. In [135, 139], this idea was applied in the context of cosmological solutions by

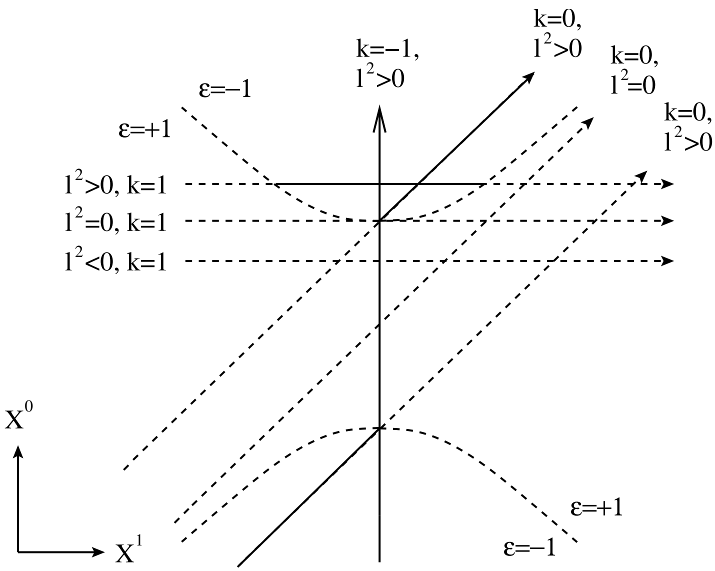


Figure 7.1: *Instantons and cosmologies as geodesics in a 2-D projection of Minkowski spacetime onto a plane, $X_2 \neq 0$, that does not pass through the origin. The solid (dashed) lines are cosmologies (instantons), which are separated by the light-cone. The dotted hyperbola with ${X_0}^2 - {X_1}^2 = {X_2}^2$ is the projection of the light-cone onto the 2-D plane.*

reinterpreting the derivatives of scalar potentials in equations of motion as Christoffel symbols of an augmented target space.

The hope behind the idea of patching cosmological solutions with instanton geometries is to find a mechanism, by means of which the Big Bang singularity can be 'smoothed out'. The Big Bang of the universe would actually be preceded by another cosmological solution that underwent a Big Crunch. The two cosmologies would be 'connected' by an instanton phase. The full solution, although singular in spacetime, is singularity-free in the target space. For details on the patching of specific cases, the reader is referred to [138].

Conclusions

This thesis has covered two separate topics: instantons and cosmologies in scalar-gravity truncations of supergravity and scalar-gravity theories in general. These were shown to be related in the final chapter.

The first chapter laid out the foundations of bosonic string theory and superstring theory. We learned that in a quantum theory of relativistic strings, mass and spin are actually the quantum numbers of a particle as opposed to Casimir operators (i.e. fixed properties). However, the main message of that chapter was the field theory limit of string theory. If we assume that strings couple weakly to each other, i.e. that the string coupling given by the constant value of the dilaton is small, then we can define a two-dimensional CFT on the world-sheet of the string. Fields such as the spacetime (target space) metric are viewed as field-dependent couplings of the σ-model, however, they can be shown to be operator insertions of coherent states of the string spectrum, such as the spin-2 particle called the graviton. Imposing that the classical conformal invariance of the CFT also hold at the quantum level requires setting the β-functions for the field-dependent couplings to zero. These constraints are perturbative in α', and, in the low energy approximation, we keep only the zeroth-order terms. This leaves us with constraints that look like the equations of motion of spacetime fields (such as the Einstein equation for the spacetime metric). By encoding these spacetime equations of motion into actions we get the supergravity actions, which are the ones that were used in this thesis.

In chapter 2 the basics of instantons were explained. We started with the example of the non-relativistic quantum mechanical particle in the double-well and periodic potentials. We learned that instantons are extrema of the Euclidean action that allow us to compute tunneling amplitudes. These tunneling effects taught us that the naïve degenerate perturbative vacua of the theory are actually not stationary states, since the particle can tunnel out of them. This allowed us to define the true vacuum of the theory, which is roughly a linear combination of the naïve vacua. The true vacuum samples all of the degenerate minima of a potential, thereby spontaneously restoring the symmetry of the theory. We then moved on to the application of instantons in quantum field theory, by treating the example of the Yang-Mills instanton. The latter showed us how the principles of instantons and true vacua generalize to quantum field theories. We saw that a path integral that takes instanton effects into account, i.e. a path integral that gives the *true*-vacuum-to-*true*-vacuum amplitude, effectively gets a topological θ-term in its action. At the end of the chapter, I gave a brief explanation of how instantons in D Euclidean dimensions can sometimes correspond to solitons in $D+1$ spacetime dimensions.

In chapter 3 we put this knowledge to use in a scalar-gravity theory. First, a quick explanation of the issues of the Euclidean path integral for gravity was given. Then, we defined a theory of gravity with two scalars, which can be embedded in type IIB supergravity for certain values of the dilaton coupling. We found the solutions of this theory and were able to classify them in terms of their $SL(2,\mathbb{R})$ 'conjugacy classes'. There turned out to be three $SL(2,\mathbb{R})$-unrelated families of solutions. The instanton-soliton correspondence that was explained in general in chapter 2 was put to use, as we realized that the three families of D-instantons can be viewed as spacelike sections of superextremal, extremal, and subextremal electrically charged black holes. We studied the singularity structure of these solutions and evaluated their actions. After a comment on the tunneling interpretation of these solutions, we discussed the possibility that they might lead to non-perturbative R^8 corrections to the type IIB effective action.

Finally, I commented on some work in progress. Putting D-instantons in an AdS background can lead to interesting applications in AdS/CFT. The correspondence between the extremal D-instanton in type IIB supergravity and the self-dual instanton of $\mathcal{N}=4, d=4$ super-Yang-Mills has been known for a while. We hope to understand the field theory dual of the non-extremal D-instantons, which may be pointing us toward non-self-dual super-Yang-Mills instantons.

The next part of this thesis was concerned with another kind of scalar-gravity solution that depends on one parameter: FLRW cosmologies. Chapter 4 introduced the basics of the standard cosmology and modern cosmology. Inflation and present day acceleration are experimentally undeniable events in our universe. If string theory is the theory of everything, it must be able to derive a realistic scenario for them. At the end of the chapter, I summarized a few of the many string theory based approaches toward modern cosmology, focusing on models that reduce to theories of four-dimensional gravity with scalar fields.

In chapter 5 we studied the gravity-scalar system with a single exponential potential. First, we showed that, by a proper field redefinition, the system effectively has only one scalar in the exponent of the potential. Then, the equations of motion were rewritten in the language of autonomous systems. We saw that, in this terminology, the familiar FLRW power-law and de Sitter solutions can be thought of as critical points, and the more interesting solutions are the ones that interpolate between those two regimes. This showed us how to recognize solutions that have periods of transient acceleration, which is phenomenologically interesting for models of both inflation and present day acceleration.

In chapter 6, we dropped all simplifications by studying the *most* general multi-exponential potential for an arbitrary number of scalars. A general formula for finding critical points was derived, which unveiled de Sitter critical points that had never been discovered. The general formula was then applied to some specific cases coming from reductions of pure gravity over three-dimensional group manifolds. At the end of the chapter, comments on possible extensions of this work were made. Theses possibilities are including a barotropic fluid in the system to mimic matter, and including spatial curvature. One possible application of such an extension is the cosmic coincidence problem, which might be solved by scaling solutions.

Chapter 7 was the concluding chapter that tied D-instantons and cosmologies together. Their mathematical similarity, due to the fact that both are solutions to scalar-gravity models that depend on only one coordinate, was translated into two concrete correspondences. First, we saw that some D-instantons are related to S(-1)-branes via Wick rotation. In the second part of the chapter we developed a formalism that put both types of solutions on equal footing. By interpreting the scalar fields as coordinates of a two-dimensional target space, and subsequently performing coordinate transformations on this target space, we realized that instanton solutions can be thought of as the trajectories of particles on a dS_2 space. The three families of instantons correspond to massive, massless and tachyonic particles. The cosmologies on the other hand are interpreted as trajectories of a particle on a Euclidean $\mathbb{H}_2$ space.

The Ansätze for the spacetime metrics of both the instantons and the cosmologies are such that both metrics have only one degree of freedom. By interpreting this degree of freedom as an *extra* target space coordinate, we were able to combine both systems into the action of one particle in a three-dimensional Minkowski spacetime. In this formalism, an instanton and a cosmology are patched together, and are viewed as two portions of the trajectory of a single particle. This suggested a possible scenario to resolve cosmological singularities. For instance, in this target space language, the Big Bang is preceded by an instanton phase, which is itself preceded by a Big Crunch.

Understanding the deeper links between instantons and cosmologies may lead to interesting and unexpected results. For instance, by using AdS/CFT to further knowledge about the correspondence between gravity and gauge instantons, one might establish new cosmology/gauge correspondences in the context of dS/CFT.

Appendix A

2-D Quantum Mechanics

In this appendix, we will study the point particle in (2 + 1)-dimensional quantum mechanics. In [44], Lee introduced this example as a toy model to show how the path integral of a positive-definite action can effectively be computed by finding the saddle points of an action that is not positive-definite. This toy model will allow us to understand why we are solving a system with a negative kinetic term for the axion in (3.5).

A.1 Path integral for momentum eigenstates

Let us begin by defining the system and its path integral. We want to study quantum mechanics of a unit mass particle moving in two spatial directions, by using polar coordinates: $r(t)$ and $\theta(t)$. The partition function and path integral between position eigenstates are defined and computed as follows:

$$\langle\, r_F, \theta_F |\, e^{-HT}\, | r_I, \theta_I\, \rangle = \int_{\text{b.c.}} (\Pi_{t'}\, r(t'))\; d[r(t)]\, d[\theta(t)]\; \exp\left[-\tfrac{1}{2} \int_{t_I}^{t_F} dt\, (\dot{r}^2 + r^2\, \dot{\theta}^2) \right], \qquad \text{(A.1)}$$

where I and F stand for initial and final, respectively; $T \equiv t_F - t_I$; and 'b.c' stands for Dirichlet boundary conditions, i.e $r(t_{I,F}) = r_{I,F}$ and $\theta(t_{I,F}) = \theta_{I,F}$, respectively. The product in the integration measure is simply the Jacobian for polar coordinates. For convenience, we will omit the integration over $r(t)$ and its kinetic term, and reinsert it when it is needed. As already mentioned in chapter 2, this partition function can, but need not be thought of as an imaginary time path integral. In this appendix, we will think of t as real time.

Suppose that we want to compute the partition function between initial and final angular momentum eigenstates $|\,\ell\,\rangle$, as opposed to angular position eigenstates $|\,\theta\,\rangle$. Using the following definition for the angular momentum states

$$|\,\ell\,\rangle \equiv \int d\theta\, e^{i\,\ell\,\theta}\; |\,\theta\,\rangle\;, \qquad \text{(A.2)}$$

we see that all we have to do is Fourier transform the path integral in (A.1) with respect to its boundary conditions:

$$\langle\, r_F, \ell_F |\, e^{-HT}\, | r_I, \ell_I\,\rangle = \int d\theta_I\, d\theta_F\ \exp\left(-i\,\ell_F\,\theta_F + i\,\ell_I\,\theta_I\right) \langle\, r_F, \theta_F |\, e^{-HT}\, | r_I, \theta_I\,\rangle\ . \tag{A.3}$$

The Dirichlet path integral can be combined with the integral over boundary conditions to yield one path integral without boundary conditions:

$$\langle\, r_F, \ell_F |\, e^{-HT}\, | r_I, \ell_I\,\rangle = \int_{\text{no b.c.}} d[\theta(t)]\ \exp\left[-\tfrac{1}{2}\int_{t_I}^{t_F} dt(\dot{r}^2 + r^2\,\dot{\theta}^2) - i\,\ell_F\,\theta_F + i\,\ell_I\,\theta_I\right], \tag{A.4}$$

where this sums over all possible $\theta(t)$ with arbitrary boundary values. If we try to compute this path integral via the standard saddle point approximation, the Euler-Lagrange variation of the action w.r.t. θ will be the following:

$$\delta S\ = -\int_{t_I}^{t_F} dt\ \left[\partial_t(r^2\,\dot{\theta})\,\delta\theta\right] + (r^2\,\dot{\theta}\ - i\,\ell)\,\delta\theta\Big|_{t_I}^{t_F}\ . \tag{A.5}$$

Notice we do not throw away the total derivative, because there are no boundary conditions. Since this must vanish for arbitrary variations $\delta\theta$, both terms must vanish independently. Hence, we get the following equations:

$$\begin{aligned} \partial_t(r^2\,\dot{\theta}) &= 0\,, \\ r^2\,\dot{\theta}\Big|_{t_{I,F}} &= i\,\ell|_{t_{I,F}}\ . \end{aligned} \tag{A.6}$$

The first equation is a normal equation of motion for θ; however, the second is a constraint that is inconsistent with the assumption that θ, ℓ, and t are real. Therefore, the path integral must be computed by means of a different method. In what follows, two methods for doing this will be presented.

A.2 Computing the path integral: first method

In this section, we will present one of two methods for computing the path integral in (A.4). It involves splitting up the integration into bulk and then boundary values of θ as in (A.3). We can easily compute the bulk integral using the usual methods of Euler-Lagrange variations with Dirichlet boundary conditions. Then, by Fourier transforming the result w.r.t. the boundary conditions, we obtain the final result.

We start by evaluating the angular part of (A.1), which has Dirichlet boundary conditions for θ:

$$\langle\, r_F, \theta_F |\, e^{-HT}\, | r_I, \theta_I\,\rangle = \int_{\text{b.c.}} d[\theta(t)]\ \exp\left[-\tfrac{1}{2}\int_{t_I}^{t_F} dt\, r^2\,\dot{\theta}^2\right]. \tag{A.7}$$

This is easily done by finding a saddle point through the Euler-Lagrange variation, which yields the following equation:

$$\partial_t\left(r^2\,\dot{\theta}\right) = 0 \Rightarrow r^2\,\dot{\theta} = \ell_{cl}\,, \tag{A.8}$$

where the constant ℓ_{cl} is the classical angular momentum. The solution is the following:

$$\theta(t) = \ell_{cl} \int_{t_I}^{t} \frac{dt'}{r^2(t')} + \theta_I \qquad \text{where} \qquad \ell_{cl} = (\theta_F - \theta_I) \int_{t_I}^{t_F} \frac{dt'}{r^2(t')} \,. \tag{A.9}$$

Defining $I[r] \equiv \int_{t_I}^{t_F} dt'/r^2(t')$, and substituting the solution into the action, we obtain the following:

$$-\tfrac{1}{2} \int_{t_I}^{t_F} dt\, r^2\, \dot\theta^2 = -\tfrac{1}{2}\, \ell_{cl}^2\, I[r] = -\frac{(\theta_F - \theta_I)^2}{2\, I[r]} \,. \tag{A.10}$$

Because the action is quadratic in θ, the semiclassical approximation provides us with an exact result for the path integral. Hence, evaluating the action at this saddle point and computing the functional determinant (as we saw in chapter 2) is an an exact evaluation of this part of the path integral. The functional determinant contains $\det(\partial_t^2)$ and $1/(\Pi_t\, r(t))$, which cancels the Jacobian in the path integral over $r(t)$. Now, in order to finish the evaluation of (A.4) (or (A.3)), all we have to do is Fourier transform this result w.r.t. the boundary conditions $\theta_{I,F}$:

$$\begin{aligned}\int d\theta_I\, d\theta_F\, \exp\left(-\frac{(\theta_F - \theta_I)^2}{2\, I[r]} - i\, \ell_F\, \theta_F + i\, \ell_I\, \theta_I\right) &= \int d\theta_I d\tilde\theta \exp\left(-\frac{\tilde\theta^2}{2\, I[r]} - i\, \ell_F\, \tilde\theta + i\, \theta_I\, (\ell_I - \ell_F)\right) \\ &= \sqrt{2\,\pi\, I[r]}\, \delta(\ell_F - \ell_I)\, \exp\left(-\frac{\ell_F^2\, I[r]}{2}\right),\end{aligned} \tag{A.11}$$

where the δ-function comes from the θ_I integral, and the exponential comes from the integral over the shifted variable $\tilde\theta \equiv \theta_F - \theta_I$. The path integral enforces conservation of angular momentum. Substituting[1] this result into the full path integral, we are left with the following:

$$\langle\, r_F, \ell_F |\, e^{-HT}\, | r_I, \ell_I \,\rangle = \delta(\ell_F - \ell_I) \int \sqrt{2\,\pi\, I[r]}\, d[r(t)] \exp\left[-\tfrac{1}{2} \int_{t_I}^{t_F} dt \left(\dot r^2 + \frac{\ell_F^2}{r^2}\right)\right]. \tag{A.12}$$

Performing the saddle point approximation on the remaining integration over $r(t)$, we find the following equations of motion:

$$\ddot r + \frac{\ell_F^2}{r^3} = 0 \,. \tag{A.13}$$

However, had we derived the normal Euler-Lagrange equations from the standard path integral with Dirichlet boundary conditions (A.1), these would have had a relative minus sign between these two terms. The result in (A.13) can also *effectively* be obtained by finding the saddle point of the following pseudo-action

$$S = \tfrac{1}{2} \int_{t_I}^{t_F} dt\, (\dot r^2 - r^2\, \dot\theta^2) - \ell_F\, \theta_F + \ell_I\, \theta_I \,, \tag{A.14}$$

and evaluating the action of the solution with it. The boundary conditions are then enforced by the surface term. The wrong sign in front of the kinetic term of θ is analogous to the sign in front

[1] The fact that we have expressed ℓ_{cl} as a functional of $r(t)$ and the θ boundary conditions means that this substitution is legal. What would be wrong, would be to explicitly keep ℓ_{cl}, and subsequently treat it as a constant upon integrating over $r(t)$, which it is not.

of the axion kinetic term in (3.5). This is as though we had looked for imaginary saddle points of the action in (A.4). Some papers in the litterature have gone as far as saying that one needs to rotate the contour of integration of θ into the imaginary line in the complex plane. However, I would like to stay away from such an unnecessary and unnatural interpretation of what is really taking place in this calculation.

The method we have presented in this section makes use of the fact that we can easily express ℓ_{cl} in terms of the boundary values of θ. This is, however, due to the fact that we are doing (2 + 1)-dimensional quantum mechanics, or (0 + 1)-dimensional quantum field theory. In higher dimensions, this task becomes more difficult; and the definition of a boundary is no longer unique, which it was in this case. Therefore, we need a more covariant way to compute the path integral that can be applied to higher-dimensional field theory.

A.3 Computing the path integral: second method

The second method we will be exploring involves the concept of dualization. The basic idea behind this is the realization that, if one wants to compute a path integral with initial and final *momentum* states, one should be working with momentum variables in the first place, as opposed to position variables.

We begin by rewriting (A.4) as follows:

$$\langle\, r_F, \ell_F |\, e^{-HT}\, |\, r_I, \ell_I\,\rangle = \int_{\text{b.c.}} d[\theta]\, d[\ell]\, \exp\left[-\tfrac{1}{2}\int_{t_I}^{t_F} dt\left(\dot{r}^2 + \frac{\ell^2}{r^2} + 2\, i\, \theta\, \dot{\ell}\right)\right], \qquad \text{(A.15)}$$

where we have inserted an integration over a dummy variable $\ell(t)$. We impose Dirichlet boundary conditions on ℓ, i.e. $\ell(t_{I,F}) = \ell_{I,F}$, and no boundary conditions on θ. Let us first show how this reduces to (A.4) upon integrating ℓ out. Integrating the last term by parts

$$-\int_{t_I}^{t_F} dt i\, \theta\, \dot{\ell} = \int_{t_I}^{t_F} dt\, i\, \ell\, \dot{\theta} - i\,(\theta_F\, \ell_F - \theta_I\, \ell_I)\,. \qquad \text{(A.16)}$$

we recognize the surface term as the one in (A.4). The first term can used to complete a square in the action as follows:

$$-\frac{\ell^2}{2\, r^2} + i\, \ell\, \dot{\theta} = -\frac{1}{2\, r^2}\,(\ell - i\, r^2\, \dot{\theta})^2 - \tfrac{1}{2}\, r^2\, \dot{\theta}^2\,. \qquad \text{(A.17)}$$

The remaining integral over ℓ is easy to perform:

$$\int d[\tilde{\ell}]\, \exp\left[-\tfrac{1}{2}\int dt\left(\frac{\tilde{\ell}^2}{r^2} + r^2\, \dot{\theta}^2\right) + \text{surface term}\right], \qquad \text{(A.18)}$$

where we have used the shift invariance of the measure by setting $\tilde{\ell} = \ell - i\, r^2\, \dot{\theta}$. Because the boundary values of ℓ are fixed, they are not being integrated over, hence, they are not affected by this shift. This integral is simply a Gaussian.[2] The end result is the original path integral

[2] The result of the determinant is a factor $\Pi_t\, r(t)$, which can be absorbed with the other equal factor we saw in (A.1) in the measure $r^2\, d[r] \to d[r^3]$. This is a bijective transformation of variables, and hence yields no problems in the extremization process.

(A.4) over θ. It should not come as a big surprise that one can write a partition function in terms of a path integral over *both* a variable and its momentum conjugate. Usually, in deriving a path integral from first principles, one obtains such an integral and subsequently eliminates the momentum variable as we just did above.

Now that we have proven that the right-hand-side of (A.15) yields (A.4) upon integrating ℓ out, let us change the order of integration, and eliminate θ first. The integral over θ is simply a δ-functional:

$$\int d[\theta] \exp\left[i \int dt \theta \, \dot{\ell}\right] = \delta\left[\dot{\ell}\right] . \tag{A.19}$$

This simply imposes conservation of angular momentum. Hence, the path integral over ℓ yields the following:

$$\int d[\ell]\delta\left[\dot{\ell}\right] \exp\left[-\tfrac{1}{2} \int dt \left(\frac{\ell^2}{r^2}\right)\right] = \delta(\ell_F - \ell_I) \exp\left[-\tfrac{1}{2} \int dt \left(\frac{\ell_F^2}{r^2}\right)\right] . \tag{A.20}$$

Therefore, the final result is the following:

$$\langle\, r_F , \ell_F |\, e^{-HT} \,| r_I , \ell_I \,\rangle = \delta(\ell_F - \ell_I) \int (\Pi'_t \, r(t')) \; d[r(t)] \exp\left[-\tfrac{1}{2} \int_{t_I}^{t_F} dt \left(\dot{r}^2 + \frac{\ell_F^2}{r^2}\right)\right] , \tag{A.21}$$

which is what we obtained with the previous method.

In terms of our dilaton-axion system in chapter 3, the radial coordinate r is roughly analogous to the dilaton ϕ, and the angular coordinate θ is analogous to the axion χ. The angular momentum, which is the conjugate variable to θ, is analogous to the $(D-1)$-form field-strength. There, the restriction that $dF = 0$ implies that, locally, $F = dC$. Here, this translates to the constraint $\dot{\ell} = 0$, which implies that ℓ is a constant. The conservation of angular momentum in the two-dimensional quantum mechanical system translates to the conservation of axion charge.

This method of dualization is preferable to the previous one, because it does not require an explicit choice of parametrization of the boundary, and does not require us to split up path integrals into bulk and boundary integrals. Hence, we will use this method in chapter 3.

Appendix B

Useful formulae in Riemannian geometry

In this appendix, I will spare the reader the annoying work of calculating the curvatures of metrics with spherical or hyperbolic symmetry that are relevant in this thesis. I will first write down some basic definitions for the sake of clarity, and to establish my conventions.

The Christoffel symbols are defined as follows:

$$\Gamma^{\rho}{}_{\mu\nu} = \frac{1}{2}\, g^{\rho\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right) . \tag{B.1}$$

In this thesis, I have used the following definition for the Ricci tensor:

$$R_{\mu\nu} = \partial_\rho \Gamma^{\rho}{}_{\mu\nu} - \partial_\mu \Gamma^{\rho}{}_{\nu\rho} + \Gamma^{\rho}{}_{\mu\nu}\, \Gamma^{\sigma}{}_{\sigma\rho} - \Gamma^{\rho}{}_{\sigma\mu}\, \Gamma^{\sigma}{}_{\rho\nu} . \tag{B.2}$$

The convention for metrics with Lorentzian signature is mostly plus, i.e. $(-\,,\,+\,,\,\ldots\,,+)$. For Euclidean metrics the convention is, well ... all plus.

The general Ansatz that encompasses all instanton and cosmological metrics that have been used in this thesis can be written as follows:

$$ds^2_D = \epsilon\, e^{2\,A(r)}\, dr^2 + e^{2\,B(r)}\, r^2\, d\Sigma^2_{k\,,\,D-1} , \tag{B.3}$$

where $\epsilon = \pm 1$ depending on the desired signature, and the two functions $A(r)$ and $B(r)$ are undetermined. In the second term, $d\Sigma^2_{k\,,\,D-1}$ is the line element of a $(D-1)$-dimensional sphere, plane or hyperbolic space for $k = 1\,, 0$, and -1 respectively:

$$d\Sigma^2_{k\,,\,D-1} = \frac{d\rho^2}{1 - k\,\rho^2} + r^2\, d\Omega^2_{S^{D-2}} . \tag{B.4}$$

This can also be written as follows:

$$d\Sigma^2_{k\,,\,D-1} = d\psi^2 + f^2(\psi)\, d\Omega^2_{S^{D-2}} , \tag{B.5}$$

where

$$f(\psi) = \begin{cases} \sin(\psi) & \text{if } k = +1 \\ \psi & \text{if } k = 0 \\ \sinh(\psi) & \text{if } k = -1 \end{cases} . \tag{B.6}$$

I will now write down the radial component of the Ricci tensor as R_{rr} and will denote transverse components by $R_{\theta\theta}$:

$$R_{rr} = -(D-1)\left(B'' + B'^2 - A'\,B' + 2\,\frac{B'}{r} - \frac{A'}{r}\right), \tag{B.7}$$

$$R_{\theta\theta} = -\epsilon\, g_{\theta\theta}\; e^{-2A}\left(B'' + (D-1)\,B'^2 - A'\,B' + 2\,(D-1)\,\frac{B'}{r} - \frac{A'}{r} + \frac{(D-2)}{r^2}\right) \tag{B.8}$$

$$+\, k\, g_{\theta\theta}\, e^{-2B}\,\frac{(D-2)}{r^2}\,. \tag{B.9}$$

All other components vanish. The non-vanishing Christoffel symbols are the following:

$$\begin{aligned} \Gamma^r{}_{rr} &= A'\,, \\ \Gamma^r{}_{\theta\theta} &= -\epsilon\, g_{\theta\theta}\, e^{-2A}\left(B' + \frac{1}{r}\right), \\ \Gamma^\theta{}_{\theta r} &= B' + \frac{1}{r}\,, \end{aligned} \tag{B.10}$$

where no sum over θ is intended in the last component.

Sometimes one needs to compute the Ricci tensor of a metric that is related via a Weyl rescaling to another metric whose Ricci tensor is already known. There is a very useful identity, which can save time in this situation. Let $g_{\mu\nu}$ be the components of a metric with Ricci tensor $R_{\mu\nu}$, and let $\tilde{g}_{\mu\nu}$ be the components of a metric that is related to the first metric as follows:

$$\tilde{g}_{\mu\nu} = e^{2\,\sigma(x)}\, g_{\mu\nu}\,, \tag{B.11}$$

where $\sigma(x)$ is some function on the manifold. Define the tensor $B_\mu{}^\kappa$ as follows:

$$B_\mu{}^\kappa = -\partial_\mu\sigma\,\partial^\kappa\sigma + \tfrac{1}{2}\,(\partial\sigma)^2\,\delta_\mu{}^\kappa + \nabla_\mu(\partial^\kappa\sigma)\,, \tag{B.12}$$

where the covariant derivative is defined in terms of the metric $g_{\mu\nu}$. Then, the Ricci tensor $\tilde{R}_{\mu\nu}$ of $\tilde{g}_{\mu\nu}$ is related to $R_{\mu\nu}$ as follows:

$$\tilde{R}_{\mu\nu} = R_{\mu\nu} - g_{\mu\nu}\, B_\lambda{}^\lambda - (D-2)\, B_{\nu\mu}\,. \tag{B.13}$$

This formula is derived in [140]. Note that it applies to any metric, and does not require any Ansatz for either the metric or the function $\sigma(x)$.

Appendix C

Homogeneous spaces and group manifolds

In this appendix, I will define homogeneous spaces and isotropy, giving some examples. Then, I will define group manifolds and illustrate with one example.

C.1 Homogeneous spaces

This section is based on a section in the book by Nakahara [140]. I will assume that the reader is familiar with Lie groups.

Let us begin by defining the action of a group on a manifold.
Definition: Given a Lie group G and a differentiable manifold M, we define an *action* of G on M to be a differentiable map $\sigma : G \times M \to M$, which satisfies the following conditions:

(i) $\sigma(e, p) = p$ for any $p \in M$,

(ii) $\sigma(g_1, \sigma(g_2, p)) = \sigma(g_1\, g_2, p)$ for any $g_1, g_2 \in G$ and any $p \in M$,

where e is the identity element of the group. The first condition needs no explanation, and the second one just means that the group action has to respect the group multiplication. Notice that each group element g defines a diffeomorphism from the manifold to itself as follows:

$$\sigma(g, \cdot\,) : M \to M\,. \tag{C.1}$$

From basic Physics we already know many examples of groups acting on spaces. The classic example is SO(3) acting on $\mathbb{R}^3$ as the group of rotations. More generally, whenever a group $G \subset \mathrm{GL}(n, \mathbb{F})$ acts on an n-dimensional vector space V over some field $\mathbb{F}$, we call that specific action of G on V an n-dimensional representation of the group G.

An action σ of group G on a manifold M automatically induces an action σ^* on the tangent space T_pM of any point p on the manifold. I will not state the mathematical definition here, but I will briefly give an intuitive picture of it. Given a vector $V \in T_pM$, tangent to p, one can always draw *integral curves*, i.e. curves passing through p that are tangent to V. Take one such integral curve $c(t)$, and let $\sigma(g,\cdot)$, for some $g \in G$, act on it point by point. This will yield a curve $\tilde{c}(t) \equiv \sigma(g, c(t))$ that passes through $q \equiv \sigma(g, p)$. Define $W \equiv \sigma^*_{g,p}(V)$ as the vector in T_qM that is tangent to $\tilde{c}(t)$ at q. This defines what is called an *induced action* of G on TM.

We also need to define the following properties for group actions:
Definition: Let G be a Lie group that acts on a manifold M by $\sigma : G \times M \to M$. The action σ is said to be

(a) *transitive* if, for any $p_1, p_2 \in M$, there exists an element $g \in G$ such that $\sigma(g, p_1) = p_2$;

(b) *free* if every non-trivial element $g \neq e$ of G has no fixed points in M. In other words, given an element $g \in G$, if there exists an element $p \in M$ such that $\sigma(g, p) = p$, then g must be the identity element e.

Now we are ready to define a homogeneous space. A manifold M is said to be homogeneous, if there exists a Lie group G that acts *transitively* on M. For instance, Lie groups act transitively on themselves via the group multiplication. The n-sphere is homogeneous because its group of rotations $SO(n+1)$ acts transitively on it. It is tempting to think that one can then simply identify a manifold with the group that acts transitively on it, by choosing a base point p on the manifold, which one would identify with e, and identifying all other points with the group elements required to go from p to them. In general, however, given any two points p_1 and p_2 on a homogeneous manifold, there could be several group elements that connect them. For instance, given two points on a sphere, there are infinitely many rotations that can bring one point to the other. One can easily show that this implies that for any point, one can find rotations that leave it fixed. More generally this means that the action of the group is not *free*. This leads to the concept of *isotropy group*:
Definition: Let G be a Lie group that acts on a manifold M. The *isotropy group* of $p \in M$ is a subgroup of G defined by

$$H(p) = [g \in G | \sigma(g, p) = p]. \tag{C.2}$$

In other words, $H(p) \subset G$ is the group of elements that leave p fixed. This is also called the *little group* or *stabilizer* of p. If G acts transitively on M, one can show that the isotropy groups of all points in M are isomorphic to each other. Let us take the example of the 2-sphere. Given a point p, we see that any rotation along the axis passing through p will leave the point fixed. So the isotropy of S^2 is $SO(2)$.
There is a remarkable theorem that states that, under certain conditions, if one has a homogeneous manifold M with the group G acting on it and with isotropy group H, then the coset space G/H is a manifold (i.e. it has a differentiable structure), and it is diffeomorphic to M, i.e. $G/H \cong M$. The following are a few of the classic examples:

$$\begin{aligned} SO(n+1)/\,SO(n) &\cong S^n\,, \\ O((n+1)/\,O(n) &\cong S^n\,, \\ U(n+1)/\,U(n) &\cong S^{2n+1}\,, \\ O(n+1)/[O(1)\times O(n)] &\cong S^n/\mathbb{Z}_2 \cong \mathbb{R}P^n\,. \end{aligned} \tag{C.3}$$

As the reader may have noticed, a manifold can have more than one representation as a coset space. It is this fact that will be exploited in the next section about group manifolds.

In section 4 we saw an intuitive definition of an *isotropic* manifold. We are now ready to give a more mathematical one:
Definition: Let M be a manifold with a group G acting on it via σ (not necesarily transitively), and let $H(p)$ be the isotropy group of some point $p \in M$. M is *isotropic at* p, if for any two vectors V_1 and V_2 in T_pM, there exists an element $h \in H(p)$ such that $\sigma^*_{h,p}(V_1) = V_2$. In other words, M is isotropic at p if all tangent vectors at p can be rotated into each other by elements of the isotropy group of p. This matches our intuition that isotropy means that a space 'looks' the same in every direction, because all directions are related via a symmetry transformation. It can be shown, that if a manifold is isotropic at every point, then it is also homogeneous. Spaces that are homogeneous and isotropic are said to be *maximally symmetric*.

None of the definitions and concepts we have defined so far have required us to define a metric on the manifold in question. But when dealing with general relativity, there is always a metric at hand. So, all of these definitions must be slightly altered from the physicist's point of view. Namely, every manifold must be endowed with a metric, and every group G acting as a group of diffeomorphisms on the manifold must leave the metric invariant, i.e. it must be a group of *isometries*. This means for instance, that the n-sphere will only be considered homogeneous, if its *isometry group* acts transitively on it. If S^n is endowed with the standard metric, then it will be homogeneous, since its isometry group SO($n+1$) acts transitively on it. If, however, it is endowed with a metric that has, for instance, no isometries whatsoever, then it will not be called homogeneous. The concept of isotropy also changes in that the isotropy group has to be a subgroup of the isometry group. Again, a manifold can have a larger or smaller isotropy, depending on the metric defined on it. A manifold will be called *maximally symmetric* if it is homogeneous and isotropic under the physicist's definitions of these two concepts.

C.2 Group manifolds

In this section, we will take the *physics* definitions of homogeneity and isotropy.

A *group manifold* is a Lie group endowed with a metric that makes it homogeneous. In other words, it is a manifold that is diffeomorphic to a Lie group and it has a metric, such that its isometry group acts transitively on the manifold. One simple example is the S^3 with the standard metric

$$ds^2 = d\psi^2 + \sin^2(\psi)\, d\Omega^2_{S^2}\,. \tag{C.4}$$

This manifold is diffeomorphic to the group SU(2). Endowed with this metric, it has SO(4) isometry and SO(3) isotropy, so we can write it as the quotient SO(4)/ SO(3). It is maximally symmetric.

However, as we mentioned before, there are more ways to identify a manifold with a quotient of groups. We could endow it with a metric that has less isometries, and hence less isotropy. The manifold S^3 can be regarded as a U(1) principal bundle over S^2 known as the *Hopf fibration*. This means that *locally*

$$S^3 \cong S^2 \times S^1\,. \tag{C.5}$$

So we can in principle write down a locally defined metric for S^3 that has SO(3) × SO(2) as its isometry group, and SO(2) as its isotropy group. I will not write down the explicit formulae because they are not clarifying, but they can be found in [95]. Therefore, we can rewrite our manifold as the following coset:

$$S^3 \cong \frac{\mathrm{SO}(3) \times \mathrm{SO}(2)}{\mathrm{SO}(2)}\,, \tag{C.6}$$

where the SO(2) in the denominator is a subgroup of the SO(3) in the numerator, i.e. the quotient SO(3)/ SO(2) forms the S^2 factor of the Hopf fibration. This manifold is no longer maximally symmetric, it is *anisotropic*.

We can even go further and write down a metric with the least amount of isometry that can still act transitively on the manifold. The isometry group must then be at least three-dimensional. Such a metric will then have no isotropy group left. In that case, we will be writing our manifold as follows:

$$S^3 \cong \frac{\mathrm{SO}(3)}{\cdot}\,, \tag{C.7}$$

where the '·' represents the trivial group. This space is totally anisotropic. All of these statements are valid only locally. Globally, of course, $S^3 \cong \mathrm{SU}(2)$, and $\mathrm{SO}(3) \cong \mathrm{SU}(2)/\mathbb{Z}_2$.

To summarize, we have written our manifold as three different quotients in the order of decreasing isometry and isotropy:

$$S^3 \cong \frac{\mathrm{SO}(4)}{\mathrm{SO}(3)} \cong \frac{\mathrm{SO}(3) \times \mathrm{SO}(2)}{\mathrm{SO}(2)} \cong \frac{\mathrm{SO}(3)}{\cdot}\,. \tag{C.8}$$

The first two forms are referred to as the *round* and the *squashed* 3-sphere respectively. In general relativity, one sees group manifolds as generalizations of maximally symmetric spaces, in that they are homogeneous but potentially completely anisotropic. In the standard terminology, which I personally find confusing, one names the group manifold after its isometry group. In the case of the 3-sphere one would call the cases in (C.8) the SO(4)-manifold, the SO(3) × SO(2)-manifold, and the SO(3)-manifold respectively.

All three-dimensional group manifolds were completely classified by Bianchi [141]. In [95] they were used as internal spaces to compactify seven-dimensional pure gravity. This yielded four-dimensional theories with gravity and scalars, with interesting exponential potentials, which could be used to obtain cosmological solutions with periods of transient acceleration. In [128] some of those theories were studied as autonomous systems to find solutions with periods of acceleration.

Bibliography

[1] J. M. Maldacena, *The large N limit of superconformal field theories and supergravity*, Adv. Theor. Math. Phys. **2** (1998) 231–252, `hep-th/9711200`

[2] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*, Cambridge, UK: Univ. Pr. (1998) 402 p

[3] J. Polchinski, *String theory. Vol. 2: Superstring theory and beyond*, Cambridge, UK: Univ. Pr. (1998) 531 p

[4] L. Brink, P. Di Vecchia and P. S. Howe, *A locally supersymmetric and reparametrization invariant action for the spinning string*, Phys. Lett. **B65** (1976) 471–474

[5] S. Deser and B. Zumino, *A complete action for the spinning string*, Phys. Lett. **B65** (1976) 369–373

[6] A. M. Polyakov, *Quantum geometry of fermionic strings*, Phys. Lett. **B103** (1981) 211–213

[7] A. M. Polyakov, *Quantum geometry of bosonic strings*, Phys. Lett. **B103** (1981) 207–210

[8] J. Polchinski, *Dirichlet-Branes and Ramond-Ramond Charges*, Phys. Rev. Lett. **75** (1995) 4724–4727, `hep-th/9510017`

[9] M. B. Green, J. H. Schwarz and E. Witten, *Superstring theory. Vol. 1: Introduction*, Cambridge, Uk: Univ. Pr. (1987) 469 P. (Cambridge Monographs On Mathematical Physics)

[10] M. B. Green, J. H. Schwarz and E. Witten, *Superstring Theory. Vol. 2: Loop amplitudes, anomalies and phenomenology*, Cambridge, Uk: Univ. Pr. (1987) 596 P. (Cambridge Monographs On Mathematical Physics)

[11] V. Balasubramanian, *Accelerating universes and string theory*, Class. Quant. Grav. **21** (2004) S1337–1358, `hep-th/0404075`

[12] L. Susskind, *The anthropic landscape of string theory*, `hep-th/0302219`

[13] F. Gliozzi, J. Scherk and D. I. Olive, *Supersymmetry, supergravity theories and the dual spinor model*, Nucl. Phys. **B122** (1977) 253–290

[14] S. R. Coleman, *The uses of instantons*, Lecture delivered at 1977 Int. School of Subnuclear Physics, Erice, Italy, Jul 23-Aug 10, 1977

[15] R. Rajaraman, *Solitons and Instantons. An Introduction to solitons and instantons in quantum field theory*, Amsterdam, Netherlands: North-Holland (1982) 409p

[16] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley (1995) 842 p

[17] A. Zee, *Quantum field theory in a nutshell*, Princeton, UK: Princeton Univ. Pr. (2003) 518 p

[18] S. M. Carroll, *Lecture notes on general relativity*, `gr-qc/9712019`

[19] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity*, San Francisco, USA: Addison-Wesley (2004) 513 p

[20] P. K. Townsend, *Black holes*, `gr-qc/9707012`

[21] M. J. Duff, R. R. Khuri and J. X. Lu, *String solitons*, Phys. Rept. **259** (1995) 213–326, `hep-th/9412184`

[22] T. Ortin, *Gravity and strings*, Cambridge Unversity, Cambridge University Press, 2004

[23] (eds.), G. W. Gibbons and S. W. Hawking, *Euclidean quantum gravity*, World Scientific (1993) 586 p

[24] A. Strominger, *Vacuum topology and incoherence in quantum gravity*, Phys. Rev. Lett. **52** (1984) 1733

[25] S. B. Giddings and A. Strominger, *String wormholes*, Phys. Lett. **B230** (1989) 46

[26] S. R. Coleman, *Black holes as red herrings: Topological fluctuations and the loss of quantum coherence*, Nucl. Phys. **B307** (1988) 867

[27] S. R. Coleman, *Why there is nothing rather than something: A theory of the cosmological constant*, Nucl. Phys. **B310** (1988) 643

[28] S. B. Giddings and A. Strominger, *Axion induced topology change in quantum gravity and string theory*, Nucl. Phys. **B306** (1988) 890

[29] M. Gutperle, *Aspects of D-instantons*, `hep-th/9712156`

[30] G. W. Gibbons, M. B. Green and M. J. Perry, *Instantons and Seven-Branes in Type IIB Superstring Theory*, Phys. Lett. **B370** (1996) 37–44, `hep-th/9511080`

[31] M. B. Green and M. Gutperle, *Effects of D-instantons*, Nucl. Phys. **B498** (1997) 195–227, `hep-th/9701093`

[32] J. G. Russo, *Effects of D-instantons in string amplitudes*, Phys. Lett. **B610** (2005) 152–158, `hep-th/0411035`

[33] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, *De Sitter vacua in string theory*, Phys. Rev. **D68** (2003) 046005, `hep-th/0301240`

[34] D. H. Coule and K.-i. Maeda, *Wormholes with scalar fields*, Class. Quant. Grav. **7** (1990) 955

[35] A. A. Tseytlin, *Type IIB instanton as a wave in twelve dimensions*, Phys. Rev. Lett. **78** (1997) 1864–1867, `hep-th/9612164`

[36] J. Y. Kim, H. W. Lee and Y. S. Myung, *D-instanton and D-wormhole*, Phys. Lett. **B400** (1997) 32–36, `hep-th/9612249`

[37] M. B. Einhorn and L. A. Pando Zayas, *On seven-brane and instanton solutions of type IIB*, Nucl. Phys. **B582** (2000) 216–230, `hep-th/0003072`

[38] M. B. Einhorn, *Instanton of type IIB supergravity in ten dimensions*, Phys. Rev. **D66** (2002) 105026, `hep-th/0201244`

[39] M. Gutperle and W. Sabra, *Instantons and wormholes in Minkowski and (A)dS spaces*, Nucl. Phys. **B647** (2002) 344–356, `hep-th/0206153`

[40] J. Y. Kim, Y.-b. Kim and J. E. Hetrick, *Classical stability of stringy wormholes in flat and AdS spaces*, `hep-th/0301191`

[41] V. V. Khoze, *From branes to branes*, Nucl. Phys. **B682** (2004) 217, `hep-th/0311065`

[42] E. Bergshoeff, U. Gran and D. Roest, *Type IIB seven-brane solutions from nine-dimensional domain walls*, Class. Quant. Grav. **19** (2002) 4207–4226, `hep-th/0203202`

[43] E. Bergshoeff, A. Collinucci, U. Gran, D. Roest and S. Vandoren, *Non-extremal D-instantons*, JHEP **10** (2004) 031, `hep-th/0406038`

[44] K.-M. Lee, *Wormholes and Goldstone Bosons*, Phys. Rev. Lett. **61** (1988) 263–266

[45] S. R. Coleman and K.-M. Lee, *Wormholes made without massless matter fields*, Nucl. Phys. **B329** (1990) 387

[46] C. P. Burgess and A. Kshirsagar, *Wormholes and duality*, Nucl. Phys. **B324** (1989) 157

[47] J. D. Brown, C. P. Burgess, A. Kshirsagar, B. F. Whiting and J. York, James W., *Scalar field wormholes*, Nucl. Phys. **B328** (1989) 213

[48] B. Grinstein and J. Maharana, *Vertex operators for axionic wormholes*, Nucl. Phys. **B333** (1990) 160

[49] H. J. Boonstra, K. Skenderis and P. K. Townsend, *The domain wall/QFT correspondence*, JHEP **01** (1999) 003, `hep-th/9807137`

[50] K. Behrndt, E. Bergshoeff, R. Halbersma and J. P. van der Schaar, *On domain-wall/QFT dualities in various dimensions*, Class. Quant. Grav. **16** (1999) 3517–3552, `hep-th/9907006`

[51] E. Cremmer *et al.*, *Euclidean-signature supergravities, dualities and instantons*, Nucl. Phys. **B534** (1998) 40–82, `hep-th/9803259`

[52] H. Lü, C. N. Pope, E. Sezgin and K. S. Stelle, *Stainless super p-branes*, Nucl. Phys. **B456** (1995) 669–698, hep-th/9508042

[53] H. Lu, C. N. Pope and K. W. Xu, *Liouville and Toda Solutions of M-theory*, Mod. Phys. Lett. **A11** (1996) 1785–1796, hep-th/9604058

[54] G. W. Gibbons and K.-i. Maeda, *Black holes and membranes in higher dimensional theories with dilaton fields*, Nucl. Phys. **B298** (1988) 741

[55] G. T. Horowitz and A. Strominger, *Black strings and P-branes*, Nucl. Phys. **B360** (1991) 197–209

[56] M. J. Duff and J. X. Lu, *Black and super p-branes in diverse dimensions*, Nucl. Phys. **B416** (1994) 301–334, hep-th/9306052

[57] R. Rohm and E. Witten, *The antisymmetric tensor field in superstring theory*, Ann. Phys. **170** (1986) 454

[58] C. M. Hull and P. K. Townsend, *Unity of superstring dualities*, Nucl. Phys. **B438** (1995) 109–137, hep-th/9410167

[59] K. Becker, M. Becker and A. Strominger, *Five-branes, membranes and nonperturbative string theory*, Nucl. Phys. **B456** (1995) 130–152, hep-th/9507158

[60] U. Theis and S. Vandoren, *Instantons in the double-tensor multiplet*, JHEP **09** (2002) 059, hep-th/0208145

[61] M. Davidse, M. de Vroome, U. Theis and S. Vandoren, *Instanton solutions for the universal hypermultiplet*, Fortsch. Phys. **52** (2004) 696–701, hep-th/0309220

[62] M. Davidse, U. Theis and S. Vandoren, *Fivebrane instanton corrections to the universal hypermultiplet*, Nucl. Phys. **B697** (2004) 48–88, hep-th/0404147

[63] E. Bergshoeff, A. Collinucci, U. Gran, D. Roest and S. Vandoren, *Non-extremal instantons and wormholes in string theory*, hep-th/0412183

[64] G. W. Gibbons and S. W. Hawking, *Action integrals and partition functions in quantum gravity*, Phys. Rev. **D15** (1977) 2752–2756

[65] E. Bergshoeff, A. Collinucci, A. Ploegh, S. Vandoren and T. Van Riet, *Non-extremal D-instantons and the AdS/CFT correspondence*, work in progress

[66] M. Bianchi, M. B. Green, S. Kovacs and G. Rossi, *Instantons in supersymmetric Yang-Mills and D-instantons in IIB superstring theory*, JHEP **08** (1998) 013, hep-th/9807033

[67] C.-S. Chu, P.-M. Ho and Y.-Y. Wu, *D-instanton in AdS(5) and instanton in SYM(4)*, Nucl. Phys. **B541** (1999) 179–194, hep-th/9806103

[68] I. I. Kogan and G. Luzon, *D-instantons on the boundary*, Nucl. Phys. **B539** (1999) 121–134, hep-th/9806197

[69] T. Banks and M. B. Green, *Non-perturbative effects in AdS(5) x S**5 string theory and d = 4 SUSY Yang-Mills*, JHEP **05** (1998) 002, hep-th/9804170

[70] V. Balasubramanian, P. Kraus, A. E. Lawrence and S. P. Trivedi, *Holographic probes of anti-de Sitter space-times*, Phys. Rev. **D59** (1999) 104021, hep-th/9808017

[71] M. Trodden and S. M. Carroll, *TASI lectures: Introduction to cosmology*, astro-ph/0401547

[72] F. Quevedo, *Lectures on string / brane cosmology*, Class. Quant. Grav. **19** (2002) 5721–5779, hep-th/0210292

[73] U. H. Danielsson, *Lectures on string theory and cosmology*, hep-th/0409274

[74] S. W. Hawking and R. Penrose, *The Singularities of gravitational collapse and cosmology*, Proc. Roy. Soc. Lond. **A314** (1970) 529–548

[75] R. M. Wald, *General Relativity*, Chicago, Usa: Univ. Pr. (1984) 491p

[76] A. H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, Phys. Rev. **D23** (1981) 347–356

[77] 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

[78] A. Albrecht and P. J. Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, Phys. Rev. Lett. **48** (1982) 1220–1223

[79] **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

[80] **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

[81] G. W. Gibbons, *Aspects of supergravity theories*, Proceedings of the GIFT Seminar 1984, edited by F. Del Aguila, J. A. de Azcarraga and L. E. Ibanez, pp. 123-146 (World Scientific, 1984)

[82] J. M. Maldacena and C. Nunez, *Supergravity description of field theories on curved manifolds and a no go theorem*, Int. J. Mod. Phys. **A16** (2001) 822–855, hep-th/0007018

[83] P. K. Townsend and M. N. R. Wohlfarth, *Accelerating cosmologies from compactification*, Phys. Rev. Lett. **91** (2003) 061302, hep-th/0303097

[84] M. Kruczenski, R. C. Myers and A. W. Peet, *Supergravity S-branes*, JHEP **05** (2002) 039, hep-th/0204144

[85] N. Ohta, *Accelerating cosmologies and inflation from M / superstring theories*, Int. J. Mod. Phys. **A20** (2005) 1–40, hep-th/0411230

[86] J. P. Conlon, F. Quevedo and K. Suruliz, *Large-Volume Flux Compactifications: Moduli Spectrum and D3/D7 Soft Supersymmetry Breaking*, hep-th/0505076

[87] P. Frè, M. Trigiante and A. Van Proeyen, *Stable de Sitter vacua from N = 2 supergravity*, Class. Quant. Grav. **19** (2002) 4167–4194, hep-th/0205119

[88] M. de Roo, D. B. Westra and S. Panda, *De Sitter solutions in N = 4 matter coupled supergravity*, JHEP **02** (2003) 003, hep-th/0212216

[89] P. K. Townsend, *Quintessence from M-theory*, JHEP **11** (2001) 042, hep-th/0110072

[90] L. Cornalba and M. S. Costa, *A new cosmological scenario in string theory*, Phys. Rev. **D66** (2002) 066001, hep-th/0203031

[91] N. Ohta, *Accelerating cosmologies from S-branes*, Phys. Rev. Lett. **91** (2003) 061303, hep-th/0303238

[92] S. Roy, *Accelerating cosmologies from M/string theory compactifications*, Phys. Lett. **B567** (2003) 322–329, hep-th/0304084

[93] C.-M. Chen, P.-M. Ho, I. P. Neupane and J. E. Wang, *A note on acceleration from product space compactification*, JHEP **07** (2003) 017, hep-th/0304177

[94] M. N. R. Wohlfarth, *Inflationary cosmologies from compactification*, hep-th/0307179

[95] E. Bergshoeff, A. Collinucci, U. Gran, M. Nielsen and D. Roest, *Transient quintessence from group manifold reductions or how all roads lead to Rome*, Class. Quant. Grav. **21** (2004) 1947–1970, hep-th/0312102

[96] L. Jarv, T. Mohaupt and F. Saueressig, *Quintessence cosmologies with a double exponential potential*, hep-th/0403063

[97] N. Ohta, *A study of accelerating cosmologies from superstring / M theories*, Prog. Theor. Phys. **110** (2003) 269–283, hep-th/0304172

[98] 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

[99] R. R. Caldwell, R. Dave and P. J. Steinhardt, *Quintessential cosmology: Novel models of cosmological structure formation*, Astrophys. Space Sci. **261** (1998) 303–310

[100] V. D. Ivashchuk and V. N. Melnikov, *Perfect fluid type solution in multidimensional cosmology*, Phys. Lett. **A136** (1989) 465–467

[101] H. Lü, S. Mukherji, C. N. Pope and K. W. Xu, *Cosmological solutions in string theories*, Phys. Rev. **D55** (1997) 7926–7935, hep-th/9610107

[102] A. Lukas, B. A. Ovrut and D. Waldram, *String and M-theory cosmological solutions with Ramond forms*, Nucl. Phys. **B495** (1997) 365–399, hep-th/9610238

[103] C.-M. Chen, P.-M. Ho, I. P. Neupane, N. Ohta and J. E. Wang, *Hyperbolic space cosmologies*, JHEP **10** (2003) 058, hep-th/0306291

[104] J. J. Halliwell, *Scalar fields in cosmology with an exponential potential*, Phys. Lett. **B185** (1987) 341

[105] A. A. Coley, J. Ibáñez and R. J. van den Hoogen, *Homogeneous scalar field cosmologies with an exponential potential*, J. Math. Phys. **38** (1997) 5256–5271

[106] A. A. Coley and R. J. van den Hoogen, *The dynamics of multi-scalar field cosmological models and assisted inflation*, Phys. Rev. **D62** (2000) 023517, gr-qc/9911075

[107] E. J. Copeland, A. R. Liddle and D. Wands, *Exponential potentials and cosmological scaling solutions*, Phys. Rev. **D57** (1998) 4686–4690, gr-qc/9711068

[108] A. R. Liddle, A. Mazumdar and F. E. Schunck, *Assisted inflation*, Phys. Rev. **D58** (1998) 061301, astro-ph/9804177

[109] I. P. C. Heard and D. Wands, *Cosmology with positive and negative exponential potentials*, Class. Quant. Grav. **19** (2002) 5435–5448, gr-qc/0206085

[110] Z. K. Guo, Y.-S. Piao and Y.-Z. Zhang, *Cosmological scaling solutions and multiple exponential potentials*, Phys. Lett. **B568** (2003) 1–7, hep-th/0304048

[111] V. R. Gavrilov, V. D. Ivashchuk and V. N. Melnikov, *Multidimensional cosmology with multicomponent perfect fluid and Toda lattices*, gr-qc/9407019

[112] J. M. Aguirregabiria, A. Chamorro, L. P. Chimento and N. A. Zuccalá, *Assisted inflation in Friedmann-Robertson-Walker and Bianchi spacetimes*, Phys. Rev. **D62** (2000) 084029, gr-qc/0006108

[113] R. Poppe and S. Schwager, *String Kaluza-Klein cosmologies with RR-fields*, Phys. Lett. **B393** (1997) 51–58, hep-th/9610166

[114] H. Lü, S. Mukherji and C. N. Pope, *From p-branes to cosmology*, Int. J. Mod. Phys. **A14** (1999) 4121–4142, hep-th/9612224

[115] U. Bleyer and A. Zhuk, *Multidimensional integrable cosmological models with dynamical and spontaneous compactification*, gr-qc/9405028

[116] V. D. Ivashchuk and S. W. Kim, *Solutions with intersecting p-branes related to Toda chains*, J. Math. Phys. **41** (2000) 444–460, hep-th/9907019

[117] M. Gutperle and A. Strominger, *Spacelike branes*, JHEP **04** (2002) 018, hep-th/0202210

[118] C.-M. Chen, D. V. Gal'tsov and M. Gutperle, *S-brane solutions in supergravity theories*, Phys. Rev. **D66** (2002) 024043, hep-th/0204071

[119] N. S. Deger and A. Kaya, *Intersecting S-brane solutions of $D = 11$ supergravity*, JHEP **07** (2002) 038, hep-th/0206057

[120] M. N. R. Wohlfarth, *Accelerating cosmologies and a phase transition in M- theory*, Phys. Lett. **B563** (2003) 1, hep-th/0304089

[121] R. Emparan and J. Garriga, *A note on accelerating cosmologies from compactifications and S-branes*, JHEP **05** (2003) 028, hep-th/0304124

[122] E. J. Copeland, A. Mazumdar and N. J. Nunes, *Generalized assisted inflation*, Phys. Rev. **D60** (1999) 083506, astro-ph/9904309

[123] Z.-K. Guo, Y.-S. Piao, R.-G. Cai and Y.-Z. Zhang, *Cosmological scaling solutions and cross-coupling exponential potential*, hep-th/0306245

[124] A. P. Billyard, A. A. Coley, J. E. Lidsey and U. S. Nilsson, *Dynamics of M-theory cosmology*, Phys. Rev. **D61** (2000) 043504, hep-th/9908102

[125] A. B. Burd and J. D. Barrow, *Inflationary models with exponential potentials*, Nucl. Phys. **B308** (1988) 929–945

[126] L. Cornalba and M. S. Costa, *Time-dependent orbifolds and string cosmology*, hep-th/0310099

[127] P. G. Vieira, *Late-time cosmic dynamics from M-theory*, hep-th/0311173

[128] A. Collinucci, M. Nielsen and T. Van Riet, *Scalar cosmology with multi-exponential potentials*, Class. Quant. Grav. **22** (2005) 1269–1288, hep-th/0407047

[129] G. F. R. Ellis and J. Wainwright, *Dynamical systems in cosmology*, Cambridge University Press, 1997

[130] A. A. Coley, *Dynamical systems in cosmology*, gr-qc/9910074

[131] A. M. Green and J. E. Lidsey, *Assisted dynamics of multi-scalar field cosmologies*, Phys. Rev. **D61** (2000) 067301, astro-ph/9907223

[132] V. D. Ivashchuk, V. N. Melnikov and A. B. Selivanov, *Cosmological solutions in multidimensional model with multiple exponential potential*, JHEP **09** (2003) 059, hep-th/0308113

[133] M. S. Bremer, M. J. Duff, H. Lü, C. N. Pope and K. S. Stelle, *Instanton cosmology and domain walls from M-theory and string theory*, Nucl. Phys. **B543** (1999) 321–364, hep-th/9807051

[134] S. B. Giddings and R. C. Myers, *Spontaneous decompactification*, Phys. Rev. **D70** (2004) 046005, hep-th/0404220

[135] P. K. Townsend and M. N. R. Wohlfarth, *Cosmology as geodesic motion*, Class. Quant. Grav. **21** (2004) 5375, hep-th/0404241

[136] R. J. van den Hoogen, A. A. Coley and D. Wands, *Scaling Solutions in Robertson-Walker Spacetimes*, Class. Quant. Grav. **16** (1999) 1843–1851, gr-qc/9901014

[137] J. X. Lu and S. Roy, *Static, non-SUSY p-branes in diverse dimensions*, JHEP **02** (2005) 001, hep-th/0408242

[138] E. A. Bergshoeff, A. Collinucci, D. Roest, J. G. Russo and P. K. Townsend, *Cosmological D-instantons and cyclic universes*, hep-th/0504011

[139] J. G. Russo and P. K. Townsend, *Cosmology as relativistic particle mechanics: From big crunch to big bang*, Class. Quant. Grav. **22** (2005) 737–752, hep-th/0408220

[140] M. Nakahara, *Geometry, topology and physics*, Bristol, UK: Hilger (1990) 505 p. (Graduate student series in physics)

[141] L. Bianchi, *Sugli spazi a tre dimensioni che ammettono un gruppo continuo di movimenti*, Memorie di Matematica e di Fisica della Società Italiana delle Scienze, Serie Terza **11** (1898) 267–352.,
English translation:
Robert T. Jantzen, *Editor's note on On the three-dimensional spaces which admit a continous group of motions by Luigi Bianchi*, General Relativity and Gravitation **33** (2001) 2157-2170.
Luigi Bianchi, *On the three-dimensional spaces which admit a continous group of motions*, General Relativity and Gravitation **33** (2001) 2171-2253.

Samenvatting

In dit proefschrift zijn twee onderwerpen behandeld: instantonen en kosmologieën in scalar-zwaartekracht truncaties van superzwaartekracht en in scalar-zwaartekracht theorieën in het algemeen. In het laatste hoofdstuk werd aangetoond dat deze onderwerpen gerelateerd zijn.

Het eerste hoofdstuk behandelde de fundamentele beginselen van de bosonische snaartheorie en de supersnaartheorie. We hebben geleerd dat elke toestand van een kwantummechanische, relativistische snaar als een deeltje kan worden opgevat, waarbij zijn massa en spin de kwantumgetallen van de snaartrillingstoestand zijn, in tegenstelling tot Casimir operatoren (i.e. vaste eigenschappen). De nadruk van dit hoofdstuk ligt op de veldentheorie limiet van de snaartheorie. Als we veronderstellen dat snaren zwak koppelen, i.e. dat de koppeling van snaren bepaald door de constante waarde van de dilaton klein is, dan kunnen we een twee-dimensionale CFT op het wereldoppervlak van de snaar definiëren. Velden zoals de ruimtetijd (*target space*) metriek worden gezien als veld-afhankelijke koppelingen van het σ-model. Ze kunnen ook worden gezien als operator inserties van coherente toestanden van het snarenspectrum, zoals het spin-2 toestand/deeltje dat graviton wordt genoemd. Om de klassieke conforme invariantie van de CFT ook te laten gelden op het kwantumniveau, moeten de β-functies voor de veld-afhankelijke koppelingen op nul worden gezet. Deze beperkingen zijn perturbatief in α' en, in de lage energie benadering, houden we alleen de nulde-orde termen. Dit geeft ons beperkingen die eruit zien als de bewegingsvergelijkingen van ruimtetijd velden (zoals de Einstein vergelijking voor de ruimtetijd metriek). Door deze ruimtetijd bewegingsvergelijkingen weer te geven als acties, verkrijgen wij de superzwaartekracht acties, welke gebruikt zijn in dit proefschrift.

In het tweede hoofdstuk werden de basisbeginselen van instantonen uitgelegd. We begonnen met het voorbeeld van het niet-relativistische kwantumdeeltje in de dubbele-put en de periodieke potentialen. We hebben geleerd dat instantonen extrema zijn van de Euclidische actie die gebruikt kunnen worden om tunnelingsamplitudes te berekenen. Deze tunnelingseffecten leren ons dat de naïeve ontaarde perturbatieve vacua van de theorie geen stationaire toestanden zijn omdat het deeltje ze kan verlaten. Dit stelt ons in staat het werkelijke vacuüm van de theorie te definiëren, wat ruwweg een lineaire combinatie is van de naïeve vacua. Vervolgens gingen we verder met de toepassing van instantonen in de kwantumveldentheorie, door het voorbeeld van de Yang-Mills instanton te behandelen. Dit toonde ons hoe de principes van instantonen en werkelijke vacua generaliseren tot kwantumveldentheorieën. We zagen dat een padintegraal die rekening houdt met de effecten van instantonen, i.e. een padintegraal die de *echt*-vacuüm-tot-*echt*-vacuüm amplitude geeft, in essentie een topologische θ-term in zijn actie krijgt. Aan het einde van het hoofdstuk heb ik uitgelegd hoe instantonen in D Euclidische dimensies soms overeen kunnen komen met solitonen in $D + 1$ ruimtetijd dimensies.

In het derde hoofdstuk konden we deze kennis toepassen in een scalar-zwaartekracht theorie. Allereerst hebben we een aantal subtiliteiten met betrekking tot de Euclidische padintegraal voor de zwaartekracht naar voren gebracht. Daarna definieerden we een theorie van zwaartekracht met twee scalairen, welke kunnen worden ingebed in type IIB superzwaartekracht voor bepaalde waarden van de dilaton koppeling. We vonden de oplossingen voor deze theorie en waren in staat ze te classificeren in termen van hun $SL(2, \mathbb{R})$ 'conjugatieklassen'. Er bleken drie $SL(2, \mathbb{R})$-ongerelateerde families van oplossingen te zijn. Door de in hoofdstuk 2 besproken correspondentie tussen instantonen en solitonen toe te passen, realiseerden wij ons dat de drie families van D-instantonen gezien kunnen worden als ruimtelijke gebieden van superextremale, extremale en subextremale electrisch geladen zwarte gaten. Na de tunneling interpretatie van deze oplossingen naar voren te hebben gebracht, hebben we de mogelijkheid besproken dat zij leiden tot niet-perturbatieve R^8 correcties aan de type IIB effectieve actie.

Ten slotte heb ik enkele onderwerpen besproken, die onderdeel uitmaken van mijn nog lopende onderzoek. Het plaatsen van D-instantonen in een AdS achtergrond kan leiden tot interessante toepassingen in de AdS/CFT correspondentie. De correspondentie tussen de extremale D-instanton in type IIB superzwaartekracht en de zelf-duale instanton van $\mathcal{N} = 4, d = 4$ super-Yang-Mills, is sinds enige tijd bekend. Wij hopen de veldentheorie duale van de niet-extremale D-instantonen te begrijpen, wat ons in de richting zou kunnen wijzen van niet-zelf-duale super-Yang-Mills instantonen.

Het volgende deel van dit proefschrift heeft een ander soort scalar-zwaartekracht oplossing behandeld, dat ook afhangt van één parameter: FLRW kosmologieën. Hoofdstuk 4 introduceerde de basisprincipes van de standaard kosmologie en moderne kosmologie. Inflatie en de tegenwoordige versnelling zijn experimenteel gemeten gebeurtenissen in ons universum. Als de snaartheorie daadwerkelijk de theorie van alles is, moet zij een realistisch scenario voor deze gebeurtenissen toelaten. Aan het eind van het hoofdstuk heb ik een samenvatting gegeven van enkele van de vele op de snaartheorie gebaseerde benaderingen van de moderne kosmologie, gericht op modellen die te reduceren zijn tot theorieën van vier-dimensionsele zwaartekracht met scalaire velden.

In hoofdstuk 5 bestudeerden we het scalar-zwaartekracht systeem met een enkele exponentiële potentiaal. Allereerst toonde we aan dat, door een juiste veldherdefinitie, het systeem effectief slechts één scalair veld in de exponent van de potentiaal heeft. Daarna werden de bewegingsvergelijkingen herschreven in de taal van autonome systemen. We zagen dat, bij gebruik van deze terminologie, de gebruikelijke FLRW 'machtswet' (*power-law*) en de Sitter oplossingen kunnen worden gezien als kritische punten en de meer interessante oplossingen zijn diegene die interpoleren tussen deze twee regimes. Dit toonde ons hoe wij de oplossingen kunnen herkennen die periodes met tijdelijke versnelling hebben, wat fenomenologisch interessant is voor modellen van zowel inflatie als de tegenwoordige versnelling.

In hoofdstuk 6, hebben we alle vereenvoudigingen laten vallen door de *meest* algemene multi-exponentiële potentiaal voor een willekeurig aantal scalaire velden te bestuderen. Een algemene formule voor het vinden van kritische punten werd afgeleid, welke de Sitter kritische punten blootlegde die nog nooit ontdekt waren. De algemene formule werd vervolgens toegepast op enkele specifieke gevallen afkomstig van reducties van pure zwaartekracht op driedimensionale groepvariëteiten. Tenslotte werden mogelijke uitbreidingen van genoemd werk beschreven, namelijk het toevoegen van een barotropische vloeistof in het systeem om materiaal na te bootsen en van ruimtelijke kromming. Een mogelijke toepassing van een dergelijke uitbreiding is het ‘kosmische samenvallen probleem’ (*cosmic coincidence problem*), wat mogelijk opgelost kan worden door ‘schalende oplossingen’ (*scaling solutions*).

In het afsluitende hoofdstuk 7 werden D-instantonen en kosmolgieën samengebracht. Hun wiskundige overeenkomstigheid, te wijten aan het feit dat beide oplossingen van scalar-zwaartekracht modellen zijn die afhankelijk zijn van slechts één coördinaat, werd vertaald in twee concrete correspondenties. Allereerst zagen wij dat enkele D-instantonen gerelateerd zijn aan S(-1)-branen via de Wick rotatie. In het tweede deel van het hoofdstuk ontwikkelden wij een formalisme dat beide typen oplossingen op gelijke voet brengt. Door het interpreteren van scalaire velden als coördinaten van een twee-dimensionale scalarvariëteit (*target space*) en het vervolgens uitvoeren van coördinatentransformaties op deze scalarvariëteit, realiseerden wij ons dat instanton oplossingen gezien kunnen worden als de banen van deeltjes in een dS_2 ruimte. De drie $SL(2,\mathbb{R})$ families van instantonen komen overeen met respectievelijk massieve, massaloze en tachyonische deeltjes. De kosmologieën, aan de andere kant, worden geïnterpreteerd als banen van een deeltje in een Euclidische $\mathbb{H}_2$ ruimte.

De Ansätze voor de ruimtetijd metriek van zowel de instantonen als de kosmologieën zijn zodanig dat beide metrieken slechts één vrijheidsgraad hebben. Door het interpreteren van deze vrijheidsgraad als een extra scalarvariëteit coördinaat, werden wij in staat gesteld beide systemen te combineren in de actie van een deeltje in een drie-dimensionale Minkowski ruimtetijd. In dit formalisme zijn een instanton en een kosmologie samengepakt en worden beschouwd als twee delen van de baan van een enkel deeltje. Dit suggereert een mogelijk scenario om kosmologische singulariteiten op te lossen. Bijvoorbeeld, in deze scalarvariëteit taal wordt de Big Bang voorafgegaan door een instanton fase wat in zichzelf voorafgegaan wordt door een Big Crunch.

Acknowledgments

As many of you know, these past few years have been a journey into the unknown for me, since I moved to a country where I only knew a handful of people. So, I would like to thank all those who have made this possible, and easier for me. That's it! Bye!

Jokes aside, I am grateful to Eric Bergshoeff and Mees de Roo for hiring me in the first place, especially since you knew very little about me until it was too late for you to get rid of me. I want to thank you, Mees, for putting me on the rails of my scientific career, and for always having your door open for discussions. I am also grateful to you for pushing me to go to Les Houches. That experience turned out to have great influence on my subsequent career decisions. I also thank you for helping me get a postdoc position.

I would like to thank my advisor, Eric, for always steering me in the right direction without being too controlling, for knowing exactly when to listen to me . . . and when to kick me out of his office. Thank you for teaching me how to see the positive side of adverse situations and to turn them into successes. I think, after all these years, I understand. I will do my best to put the lesson into practice, and, who knows . . . may even pass it on . . .. I am also grateful to you for all the efforts you made in my final year to help me write this thesis and find a postdoc position.

I am of course grateful to my reading committee, Tomás Ortín, Antoine Van Proeyen, and Jorge Russo, for taking on the task of carefully reading and scrutinizing this thesis.

During the four years of my Ph.D. program, I've had the chance to travel to many interesting places to attend interesting conferences. For that, I am grateful to my department. I am also grateful to the institutions that organized these activities, especially ICTP, RTN, and the Les Houches Physics School.

I would like to thank someone who did not participate in my Ph.D. formation, but certainly prepared me for it. I thank Marcus Luty, my undergraduate advisor, for introducing me to the subject of QFT and for helping me get into graduate school.

It's time to thank my colleagues in Groningen. I will do so in no particular order, neither alphabetic nor prioritized. I would like to thank Wissam Chemissany, for cheering me up with your kind words of support during difficult times, even while you were under the great stress of uncertainty. I am grateful to Diederik Roest and Ulf Gran for providing a friendly working atmosphere, even at times when we disagreed, and for a nice after-work atmosphere. I would like to thank Dennis Westra for all the nice conversations, scientific or not. Te deseo mucha suerte para el año que viene con tus decisiones. Pequeño consejo para la tesis: ¡Empieza temprano! Thanks go out to Rein Halbersma for your comforting words during my first year. I am also grateful to Jelle Hartong for the discussions and laughs we shared, and to André Ploegh for the nice ongoing collaboration. I would also like to thank . . . what's his name . . . oh yeah, Sven Kerstan ;). I thank you especially for your support during that difficult application/waiting for answers period. Ich wünsch' dir viel Glück nächtes Jahr. Mikkel Nielsen, thank you for those nice two years of collaboration/chats. I hope you are well and happy with your new life. Román Linares, está bien, no me quejo más ;). Espero que te vaya bien con el tema trabajo y que nos volvamos a ver.

In a few month's time, I will be working at the KUL in Leuven, Belgium. So, under these extraordinary circumstances, in order to avoid a diplomatic conflict, I will have to thank my colleague Thomas Van Riet. Not only for the collaboration and support, and your invaluable teachings on the subject of beer, but also for being such a tenacious, yet friendly challenge to work with.

I am very grateful to Tim de Wit for having helped me help myself out of my perpetual state of computer illiteracy so uncountably many times, and for having done it with a smile. I also thank you for those few but important instances when you gave me real words of advise that had nothing to do with computers.

Martijn Eenink, I would like to thank you first of all for the collaboration during our first year. But especially for the interesting conceptual and scientific discussions that we had after our projects, from which I have learned immensely more. I am also grateful for your support, especially during the final frantic days of thesis typesetting. I wish you the best of luck in your future, whatever path you may choose.

My thanks go to Elisabetta Pallante for elucidating discussions on QFT, but especially for helping me in the process of physically bringing this thesis into existence. I thank Arnout van Enter for all those chats about the history of physics. I am of course grateful to Iris Kwant and Ynske Joustra for all their help, but more so I am grateful to you for allowing me waste so much of your time with all those never-ending conversations about this, that, and the other. Ynske, veel success met je nieuwe baan.

I thank Thomas Cohen for one of the best courses I have ever taken, for advising me from my undergraduate and graduate studies, and for helping me stay on the path in moments of doubt. I would like to thank Michael Brin, for one of the best and most challenging math courses I have ever had (topology), and for scaring me straight in tough times. Prof. Brin, I am of course also grateful to your son for inventing Google, thereby facilitating the writing of this thesis, among countless other things.

Twice during the first two years of my Ph.D. program, I attended the AIO-School, and intensive two-week Physics course covering four subjects. It would not have been such a nice working/fun atmosphere if it wasn't for Paul Koerber, Joris Van den Bergh, Achilleas Lazopoulos, Mboyo Esole, and Stijn Nevens. Stijn, I also thank you for your friendship.

I would like to extend my gratitude to Stefan Vandoren, with whom I have collaborated for the past two years, for all those scientific discussions, which were sometimes peaceful, sometimes heated, but never boring and always friendly. I am also grateful to you Stefan, for helping me get a Postdoc position.

Thanks go out to the wonderful people I met at the Les Houches Summer School for giving me three nice weeks filled with interesting discussions of *all* kinds. I especially refer to Yacine Dolivet, Fabien Vignes, Kazuhiro Sakai, Yulia Klochko, Mauro Riccardi, Shaffique Adam, Oleg

Ruchayskiy, and Marcos Mariño. Marcos, thank you for giving the first comprehensible lectures on open/closed string duality in the context of Topological String Theory. Oleg, you may not know this, but your brief words of encouragement last year have had a tremendous impact on my decisions about pursuing Physics.

All work and no play makes makes Jack a dull boy. I am grateful to Daniele Fausti and Arien Kamp for the best Friday afternoon coffee breaks I could ever hope for.

Boy are these acknowledgments long! Bear with me, I'm halfway through.

I would now like to thank some people whom I have never met. First of all, I thank Johann Sebastian Bach, for getting me through those long nights of thesis writing. I am grateful to Conan O'Brien, the Max Weinberg 7, and the rest of the Late Night Show staff, for maintaining absurdist humor in the airwaves. I thank Dennis Owens for so many years of wise and witty quotes. I am especially grateful to Jon Stewart and the staff of the Daily Show for providing me with a happy ending after every long thesis-writing-night. It's nice to know that, although the world may be devoid of logic and justice, it will never be devoid of humor. On the same note, I must thank certain political world leaders for providing the world with years of prime comedic material. Without them, the Daily Show wouldn't be the same.

I must thank the Philly crowd for the friendship and support I go during and after my Philly period: Greg Grason, Jonathan Fisher (Sitro), Ioannis Papadimitriou, Alper Corlu (sorry, I don't know how to make the Turkish symbol). Thanks go to Rogerio de Sousa for helping me with my important decisions. Matt Stanton, thanks for bringing me back on the path, and for all those late skype chats that brought many thesis nights to a soothing end.

Of course, the AAAHT club has my full gratitude for the wonderful friendship. I especially thank Adrien Russell and Hamet Touré for staying in touch. Mauge Mascolo, thank you also for your friendship. Alejandro Argerich, you's thanked of kwoas. Ale, this is a first on paper: You:"So, was it me or was it Ziggy's theory?"Me:"Naaah! Dey weah boat flaaperonis!!!!". Top that!

I would like to thank the new doctor Patricio Mena for his companionship during all those ever so important lunch/coffee/Belgian beer-breaks. Te agradesco por tu amistad, y también por to apoyo especialmente hacia el final de este sufrimiento que es escribir la tesis. No te olvides . . . we got da brains!!

Kenneth Gustafsson, I thank you for your friendship and your advise about graduate school, and other things. But I am even more grateful to you for coming into my life, because otherwise I would have probably never met a Swedish speaking Finn. I mean c'mon . . . what are the odds?

José Torero, my mentor, who continued to mentor me years beyond his contractual requirements, thank you for all those hours in that little office at College Park. Even though I was first educated in the French system, it was you who showed me how to use logic in real life. Your words had the greatest influence (i.e. it's all your fault) that made me choose this field at that bifurcation in my life. Your not too fresh, yet not too old experience in the doctoral life helped me get through this in one peace. No te pierdas Torero, nos llamamos, ¿ta?

Bruno Cazzuli I thank for 28 years of ongoing friendship. Even when you're far away, all I need to know is that you exist, and the world makes sense. ¡Y ahora comé lo que hay y no j. . .!

I also thank Diego D'Alessandro for those free therapeutic sessions that helped me through this Ph.D. process.

In the field of physics, one of the most difficult things to live with is the pressure to distinguish oneself from the rest, to exceed mediocrity. That's why it is nice for me to know that there's a place where I can feel special. I thank my family in Geneva for that: Stephania, Cati, Marios and Costa Zourdos.

I posthumously thank my aunts Margherita and Luciana Collinucci, who, before their untimely passing, morally supported me in my pursuits. I am grateful to my father, Luigi Collinucci and my aunt Regina Collinucci for believing in my abilities and encouraging me in my studies.

I also would like to posthumously thank Clara Stromer, my grandmother, for her affection, and for believing that I can do anything.

I am grateful to Anny Steen, who helped me settle into this country and treated me like one of the family.

For the sake of the trees! I must bring this to an end.

None of this would have been possible without the help and support (moral *and* financial) of my mother, who put me through school, college and beyond. Susana Eri also largely contributed to this thesis by finding numerous mistakes, but it is mostly her scientific contribution in the form of diagrams that I am grateful for ;). Gracias por todo mami, espero que estés orgullosa. SG está presente.

Last but not least, I would like to thank Christel Steen, my partner in life. Thank you for coming back to me in Philly. I couldn't have hoped for a better change in my life. Most of this thesis was written while you were trying to sleep. Thanks for being patient with my night-owl habits. But thank you most of all for your support and your warmth. I couldn't have done this without you. For this and so much more, I love you!