DISSERTAÇÃO DE MESTRADO IFT–D.001/21

Dark Energy in String Cosmology

Gabriel Lucas Andrade de Sousa

Orientador

Horatiu Stefan Nastase

Março de 2021

S725d

Sousa, Gabriel Lucas Andrade de.

Dark energy in string cosmology / Gabriel Lucas Andrade de Sousa. – São Paulo, 2021

235 f. : il.

Dissertação (mestrado) – Universidade Estadual Paulista (Unesp), Instituto de Física Teórica (IFT), São Paulo

Orientador: Horatiu Stefan Nastase

1. Constante cosmológica 2. Energia escura (Astronomia). 3. Holografia. 4. Kaluza-Klein, teorias de. 5. Teoria de cordas I. Título

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Física Teórica (IFT), São Paulo. Dados fornecidos pelo autor(a).

Dedico esta dissertação à
minha família.

Agradecimentos

Escrever esta dissertação de mestrado foi um trabalho longo e desafiador. Porém, este é somente o símbolo de mais um ciclo que se conclui na minha vida pessoal e profissional. Sem dúvidas, sem ter tido as oportunidades que tive no passado nunca conseguiria sequer chegar ao mestrado. Deste modo, começo esta lista de agradecimentos relembrando meu período pré-graduação. Procurarei não citar muitos nomes para evitar esquecer de pessoas importantes.

Aos meus amigos de ensino médio ainda na minha cidade natal, Manaus, minha profunda gratidão. Saibam que a maior parte dos amigos com os quais pretendo contar pelo resto da vida, independente da distância, ainda vem desse período. Certamente o colegial não teria o mesmo sentido sem nossa amizade e toda a alegria com que me recebem toda vez que visito Manaus até hoje. Em especial preciso citar meu braço direito Gabriel Akel, quem não só pude ter como melhor amigo durante o ensino médio mas também como colega de apartamento durante os quatro anos de graduação. Hoje, aos meus olhos, você faz parte da minha família.

Agradeço também aos meus amigos de faculdade, com os quais pude aprender muita coisa que levarei para sempre comigo. Descobri como é importante trabalhar e estudar em grupo notando que cada um tinha seus pontos fortes e fracos que quando unidos conseguíamos aprender coisas incríveis sob a constante pressão de curtos períodos de tempo da faculdade. Sou grato também, é claro, pelos momentos únicos de festas da faculdade que vivemos juntos. Sou especialmente grato a Carolina Yaly com quem, por mais de seis anos, compartilhei cada detalhe da minha vida e que certamente me fez uma pessoa melhor. Não imagino como teria sido minha vida na faculdade sem você.

Mais recentemente, tive o privilégio de poder ter colegas de mestrado que também marcaram minha vida. Foi um período em que tive que me adaptar na transição entre bacharel em física médica para mestrando em física de altas energias e ao mesmo tempo aprender muita coisa nova, o que só foi possível com a ajuda deles. Nosso convívio dentro e fora do IFT me deu novos amigos e possíveis futuros colaboradores. Por tudo isso, sou muito grato.

Fugindo da ordem cronológica, agradeço a todos os professores que já tive (antes, durante e depois da faculdade) e que ajudaram a me moldar como profissional, formando o adulto curioso e determinado que tenho orgulho de ser. Meus pais e minha tia também merecem agradecimentos aqui uma vez que, mesmo dando a vida para que eu pudesse estudar nas melhores escolas possíveis dentro da nossa realidade, sempre foram muito ativos no meu processo de aprendizagem. Gostaria de deixar meus agradecimentos em especial ao meu orientador Horatiu Nastase, que fez muito mais que apenas me orientar. Além de ter sido meu professor em duas ocasiões, sempre esteve disposto a me dar valiosas dicas profissionais.

Agradeço também a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro neste projeto de mestrado, processo número 2018/25390-0.

Finalmente, por último o mais importante, agradeço do fundo do coração à minha família e a Deus. Meu pai, Jerônimo, minha mãe, Clemília, minha tia e segunda mãe, Marlene, e minha irmã e melhor amiga, Ana, em especial. Vocês fazem com que eu me sinta a pessoa mais sortuda do mundo. Não importa quão grandes meus sonhos possam ser, sinto que com vocês ao meu lado tudo parece factível. Vocês me deram minha religião, meu time de futebol, me ensinaram a ter prioridades na vida, me ensinaram meus valores e caráter. Nunca seria capaz de expressar o quão privilegiado me enxergo por ser dessa família que fica mais feliz que eu pelas minhas conquistas e mais chateada que eu pelos meus insucessos. Amo vocês!

“Todas as escolhas implicam perdas. Ninguém é digno do fundamental se não estiver disposto a perder o que é trivial.”

Augusto Cury

“Adversity has the effect of eliciting talents which in prosperous circumstances would have lain dormant.”

Horace

“We absolutely must leave room for doubt or there is no progress and there is no learning. There is no learning without having to pose a question. And a question requires doubt. People search for certainty. But there is no certainty. People are terrified — how can you live and not know? It is not odd at all. You only think you know, as a matter of fact. And most of your actions are based on incomplete knowledge and you really don’t know what it is all about, or what the purpose of the world is, or know a great deal of other things. It is possible to live and not know.”

Richard P. Feynman

Abstract

In this master thesis, three string-inspired cosmological models are compared: string (brane) gas, holographic, and chameleon cosmologies. More precisely, dark energy features are found and the cosmological constant problem is analyzed in each scenario. Their possible solutions to this problem are quite different and common ground is hard to be found. While, within string gas cosmology, dark energy traits only appear in the primordial Hagedorn phase (analogous to inflation in standard cosmology) and have hardly anything to do with the currently observed cosmological constant, both holographic and chameleon cosmologies give us interesting insights on the dominant energy density in the universe today. On one hand, within holographic cosmology, one can map the cosmological constant problem to a renormalization group (RG) flow in the dual field theory. Then, the problem can be "holographically solved", even though the precise mechanism for the bulk solution is unknown. On the other hand, the chameleon setup encourages us to motivate a scalar potential for a modulus to create a quintessence model, hoping to explain the currently observed dark energy density. In the end, it is attempted to implement the chameleon idea in the original Kaluza-Klein theory, proposing a quintessence model from its scalar field. After fixing some problems with the initial idea, it is found that the scalar potential is still not suitable to have had relevant implications on structure formation in the universe.

Key-words: Dark energy; Cosmological constant problem; String gas; Brane gas; Holography; Chameleon scalar; Kaluza-Klein theory; Quintessence models.

Field of knowledge: Cosmology; Field theory; String theory.

Resumo

Nessa dissertação de mestrado, três modelos cosmológicos inspirados em teoria de cordas são comparados: cosmologias de gás de cordas (branas), holográfica e camaleônica. Mais precisamente, traços de energia escura são estudados e o problema da constante cosmológica é analisado em cada cenário. As possíveis soluções para este problema são bem distintas e características em comum são difíceis de serem encontradas. Enquanto, na cosmologia de gás de cordas, elementos de energia escura só aparecem na fase primordial de Hagedorn (análoga à inflação na cosmologia padrão) e pouco tem a ver com a constante cosmological observada atualmente, tanto a cosmologia holográfica quanto a camaleônica nos trazem novas informações interessantes sobre a densidade de energia dominante no universo hoje em dia. Por um lado, na cosmologia holográfica, pode-se mapear o problema da constante cosmológica ao fluxo do grupo de renormalização na teoria de campos dual. Logo, o problema pode ser "holograficamente resolvido", mesmo que o mecanismo para a solução seja desconhecido na teoria gravitacional. Por outro lado, a ideia do camaleão nos encoraja a motivar um potencial escalar para um módulo a fim de criar um modelo de quintessência, esperando explicar a densidade de energia escura observada atualmente. No final, tentamos implementar a ideia do escalar camaleão na teoria original de Kaluza-Klein, propondo um modelo de quintessência a partir do seu campo escalar. Após retificar problemas oriundos da ideia inicial, foi notado que o potencial do campo escalar continuou não tendo as características necessárias para ter tido um papel importante na formação de estruturas do universo.

Palavras-chaves: Energia escura; Problema da constante cosmológica; Gás de cordas; Gás de branas; Holografia; Escalar camaleão; Teoria de Kaluza-Klein; Modelos de quintessência.

Áreas do conhecimento: Cosmologia; Teoria de campos; Teoria de cordas.

Contents

List of Figures

Notations and Conventions

In order to avoid confusions some notations and conventions will be listed here and those will be followed throughout the whole master's thesis unless otherwise specified.

First, indices conventions. Middle-alphabet Greek letters, e.g. μ, ν, ρ, will denote spacetime indices, while middle-alphabet Latin letters, e.g. i, j, k, will stand for spatial indices only. We may also represent 4-dimensional (4D) spacetime vectors as, for example, $p^\mu = (p^0, \vec{p})$, where $\vec{p}$ is a vector in the 3-dimensional (3D) spatial subspace of the full 4D spacetime. When we consider higher dimensional spacetime, for instance in Kaluza-Klein theories, spacetime vectors will usually be denoted by capital Latin letters, e.g. $p^M = (p^\mu, p^m)$ where μ still stands for the four observed dimensions and m runs through indices of the complementary subspace. If we want to write explicitly the spinorial components of a spinor, their indices will be denoted by letters from the beginning of the Greek alphabet, e.g. α, β, γ. While the beginning of the Latin alphabet, e.g. a, b, c, will be used for indices of gauge groups in some given representation.

We will be adopting the mostly plus signature convention for Lorentzian spacetime metrics, e.g. 4-dimensional Minkowski metric is $\eta_{\mu\nu} = diag(-1, +1, +1, +1)$. As usual, spacetime indices are raised or lowered by contracting with the metric $g_{\mu\nu}$ or its inverse $g^{\mu\nu}$.

Multiplication of two objects that have the same index represents a contraction and this index is implicitly summed over: $A \cdot B \equiv A^\mu B_\mu \equiv \sum_\mu A^\mu B_\mu = \sum_\mu A^\mu B^\nu g_{\mu\nu}$. This is the Einstein summation convention.

The natural units system will be used. This means $\hbar = c = k_B = 1$. Since two of the three (considered to be) fundamental dimensional constants - $c, \hbar$ and G - are set to one, all quantities can be expressed, for example, in terms of powers of units of mass. Thus we have the the following equivalences in dimensions of quantities:

$$[\text{Mass}] = [\text{Energy}] = [\text{Length}]^{-1} = [\text{Time}]^{-1} = 1$$

The last equality is conventional, since any quantity Q has dimensions of Mass^n (for any $n \in \mathbb{Z}$) then we can represent its dimension by $[Q] = n$.

When we need to compare against experimental values expressed in some other units we will restore dimensions of time and length via dimensional analysis.

Our convention for d-dimensional Fourier transform and inverse Fourier transform will be:

$$F(k) = \int d^d x f(x) e^{ik\cdot x}\,;$$
$$f(x) = \frac{1}{(2\pi)^d}\int d^d k F(k) e^{-ik\cdot x}\,;$$

such that we have the expected integral form for the d-dimensional Dirac delta function:

$$\delta^d(x) = \frac{1}{(2\pi)^d}\int d^d k e^{ik\cdot x}$$

The symbol log will be used to represent the Napierian logarithm, i.e. logarithm to the base e, instead of $\log_{10}$ as it is sometimes used. Therefore, $\log 10 \equiv \log_e 10 \simeq 2.3 \neq 1$.

Introduction

Cosmology has recently, i.e. in the past few decades, entered a precision era that opened windows for testing fundamental theories through cosmological observations of an early hot epoch of the universe. More precisely, there was a time in the past when the universe was too hot to allow for stable atoms. As the universe cooled down, most of the free charges permeating the cosmos were put together to form electrically neutral atoms. Since then, photons were able to travel almost freely through space. Radiation of that primordial era is constantly interacting with our detectors nowadays and it earns the name cosmic microwave background radiation (CMBR). This is one of the most important pieces of data we have to infer how the universe behaved shortly after what is called the cosmological singularity, a point in time where our standard cosmological theory breaks down.

The current framework which better explains, i.e. better agrees with observations, the cosmic evolution from now back to a few minutes after the cosmological singularity is the so-called ΛCDM model or Hot Big Bang cosmology. In this model, at the largest scales, the matter content of the universe behaves like a mixture of different types of fluids whose energy densities are currently dominated by the cosmological constant Λ, hence the first letter in ΛCDM, followed by non-relativistic (dust) matter, which itself is mostly composed of cold dark matter (CDM) completing the model's name, and a small amount of radiation or relativistic matter. Back in the past both dust matter and radiation were once more dominant than the cosmological constant. Although this model is very successful in reproducing observations, some problems arise within it. These are called pre-inflationary problems since they are often considered to be solved by inflation, a period of rapid expansion of the fabric of space. Inflation plus ΛCDM cosmology are usually considered the standard model of cosmology, in analogy to the highly precise Standard Model of particle physics.

However, even the inflationary paradigm comes with its caveats. At high enough energies our best physically tested theories are supposed to give place to a yet unknown (at least not completely known) more fundamental quantum theory which includes gravity. String theory is the most developed candidate for such a quantum gravity theory. More than just a quantum theory of gravity, it includes all the other known fundamental forces of nature. For this fact, it is usually classified as a theory of everything. Nevertheless, string theory is still a work in progress. It is natural then to use these cosmological observations as an opportunity to eventually be able to get experimental hints on what string theory is since our best particle experiments on Earth are still many (~ 14) orders of magnitude away from the energy scale at which stringy effects are expected to be relevant.

In this thesis, our goal is to analyze how dark energy features arise in three different string-inspired cosmological models: string/brane gas, holographic, and chameleon cosmologies. Dark energy is just a generalization of the cosmological constant as the type of energy that is responsible for the currently observed accelerated expansion of the universe. Currently, we are not aware of what dark energy is, or what it is made of, we can only feel its gravitational effects. Another related problem to be studied in the previously mentioned models is the so-called cosmological constant problem. The issue is that the theoretically expected value for the cosmological constant is off by a $\sim 10^{-120}$ factor – in the most severe case – of its observed value. This likely indicates something is fundamentally wrong with our understanding of what the cosmological constant (or dark energy) is.

The first three chapters of this thesis are introductory ones on standard textbook subjects. Chapter 1 properly introduces the inflation plus ΛCDM model. The relevant issue of how CMBR data is compared to it is sketched at the end of the chapter. Chapter 2 is shorter than the rest but earns its privileged spot since it introduces dark energy and the main problem of the thesis, perhaps one of the main problems in theoretical physics today, the cosmological constant problem. After that, we make a digression in chapter 3 to introduce a few string theory concepts that will be needed in the last chapters.

Finally, the main focus of the thesis is in the following three chapters, one for each cosmological model. Chapter 4 introduces string gas cosmology, first proposed by Brandenberger and Vafa [1], and its more general version, brane gas cosmology [2]. We start by reviewing the model's defining properties and their consequences. Its comparison with CMBR data is also described since the main goal of any physical model must be to be compared against observations. In the end, we show how dark energy features arise in the model. Chapter 5 is on holographic cosmology, started off by Maldacena [3]. This model has increasing expectations of being a reliable and robust alternative to inflation. It is argued that, within it, one can reproduce observations with fewer assumptions than within inflation, i.e. it is a broader paradigm. Most of the chapter is concerned about the gauge/gravity duality introduced in [4, 5] by McFadden and Skenderis. At the end of the chapter, we review how the cosmological constant problem could potentially be "holographically solved" within the model [6]. Chapter 6 is on chameleon cosmology, first proposed by Khoury and Weltman [7, 8]. After the main properties of chameleon scalars are described, we review an attempt to embed this scenario in string theory and how observations constrain the model. After that, some new calculations are presented by trying to implement the chameleon idea within the Kaluza-Klein (KK) theory. We try to build a quintessence model out of the KK chameleon. Some problems are found in this attempt and solutions to them are proposed as phenomenological chameleon models which, unfortunately, end up not being suitable to have had an important role in structure formation in the past. In chapter 7 we conclude by comparing the models.

The thesis includes four appendices in which long, yet relevant (but not too relevant to be placed in the main chapters), calculations are found. The most important of them is appendix D, where the Kaluza-Klein original theory is dimensionally reduced in a considerable amount of detail. This appendix is essential to the Kaluza-Klein discussion of chapter 6.

Chapter 1

The Standard Model of Cosmology

Cosmology is the the branch of science which studies the largest scales in our universe. Actually, this does not capture the whole picture. Since the discovery of special relativity by Albert Einstein in 1905, see original papers [9, 10], physicists realized that there is a maximum speed, c, for propagation of causal influence through space, the so-called "speed of light" for historical reasons. The existence of this threshold for causal influence implies that when we are looking at the furthest points in the sky we are actually looking back in time, since it takes a finite and considerable amount of time for light to travel from a star far away from Earth until it reaches us. This phenomenon is very much like when one throw a rock at a window. The window doesn't break immediately after the rock is released and it takes more time between the release of the rock and the breaking of the window if one is further away from the window in the first place. Newton's law of gravity, for example, assumes the effect of a cause to happen instantaneously. Thus Einstein's postulate was really groundbreaking as it implies that cosmology also tells us the history of our universe.

This is a rather new branch of science, compared to other subareas of physics, as one of the first observations that shaped cosmology as it is today only happened in 1929 when Edwin Hubble noticed the universe was expanding [11]. We summarize here the two main observed facts that enable us to figure out the geometry of our universe and thus, from the

Einstein equations, its evolution.

1.1 Initial Observations

1. Hubble law:

   The first such observation was the Hubble expansion cited above. To understand the consequences of this observation we first recall the definition of redshift, from the Doppler effect in special relativity:

$$z \equiv \frac{\lambda - \lambda_0}{\lambda_0}; \tag{1.1}$$

   where λ is the observed wavelength of the light ray emitted by the star and λ_0 is the wavelength at the moment of emission of this ray. We are able to figure out λ_0, even though we don't measure it, from the atomic composition of the star.

   As we can see, this quantity earns the name redshift because when $\lambda > \lambda_0$ we have $z > 0$ and we say that the wavelength was "redshifted" (large wavelengths on the visible spectrum correspond to the color red). On the other hand, when $\lambda < \lambda_0$ we have $z < 0$ and we say the wavelength was "blueshifted" (small wavelengths on the visible sprectum correspond to the colors violet or blue).

   Hubble measured the distances of stars in the sky to us, r, and their respective redshifts. We will not go into much detail of how one measures the distance of astronomical objects here (for more details see chapter 3 of book [12]), but it can be done by measuring the brightness of some celestial bodies that work as standard candles, in the sense that we expect them all to have the same brightness at the same distances. He then was able to fit a straight line whose slope defines the famous Hubble constant, H.

$$z = Hr \tag{1.2}$$

Nowadays we know that the Hubble constant is not necessarily a constant and, in general, we have $H = H(t)$. Thus Hubble parameter is a more appropriate name and that is how we will call it from now on in this thesis.

2. Isotropy of the universe:

   The second observation comes from a set of data called the Cosmic Microwave Background Radiation (CMBR). The CMBR is so important to cosmology that it will have a section later on better discussing its features. For now it suffices to say that this is the radiation arriving at Earth from the Big Bang that was last scattered at the last moment our universe was ionized. A long time ago the gas of photons (and by that time another types of matter coupled to it) permeating our universe was so hot that stable atoms were not able to form. Every time an atom was formed it quickly interacted with some nearby photon that would ionize it. After that period, the universe cooled down allowing atoms to be stable and photons were able to travel through an electrically neutral space with low probability of interacting electromagnetically. However their gravitational interaction with massive bodies was not shut down by this cooling down process. Moreover, with the electromagnetic interaction being negligible, the gravitational interaction dominates and when we observe the CMBR nowadays we see the effect that massive bodies have on its path until it reaches us. Since we observe an isotropic distribution of this radiation coming from everywhere in the sky, up to a high precision, we conclude that photons coming from any direction were affected by objects in their paths roughly in the same manner and therefore the distribution of mass in our universe is isotropic.

3. Homogeneity of the universe:

   The last experimental fact that shapes our understanding of cosmology is the homogeneity of the universe. Although it is easy to see that on scales of the Milky Way our universe is composed of small regions with a large density (stars and planets) and basically empty space in between them, i.e. the space is inhomogeneous, at the largest scales our universe is homogeneous. On cosmological scales galaxies are just point-like objects and we can indirectly conclude, from studying peculiar velocities of galaxies (we will discuss this method in more detail when we talk about dark matter), that the distribution of mass in our universe is indeed homogeneous.

Now that we listed the important facts that enable us to guess the form of the metric of our universe, we can go to a more geometrical or general relativistic discussion of how our universe looks like at the largest, i.e. cosmological, scales and how it evolves.

1.2 Geometry of the Universe

We started this chapter by highlighting how groundbreaking special relativity was. Now we must briefly introduce its extension, the so-called general relativity (GR). Special relativity earns its name because it is a theory that reconciled the known fact that every inertial frame is equivalent, meaning that the laws of physics should be written in the same way in any such frame, with the peculiar fact that Maxwell equations implied a constant speed of light, later on Einstein figured out that this speed is not just that of light but it is an upper limit on the speed of any causal interaction. "Special" means that this equivalence is manifested only among inertial frames while "relativity" is what we call a theory that relates observations from one frame to another. GR, then, is the generalization of the equivalence of special relativity to any frame, including accelerated ones. Einstein realized that any accelerated frame is locally equivalent to a frame under the influence of gravity, this is the equivalence principle. The importance of this statement is incomplete until we realize that gravity is better described as the influence of the shape of space itself, spacetime to be more precise, on the objects moving on it rather than an actual force. Therefore, an object under the influence of just gravity is not actually under the influence of any force and a reference frame following its motion would be inertial. Since accelerated frames are locally equivalent to frames under the influence of gravity, as we said, we finally are able to relate non-inertial frames to inertial ones and can use the rules of special relativity, taking into consideration the curvature of spacetime, to any frame.

In order to describe spacetime as a dynamical background on which all the physical phenomena take place one needs to understand differential geometry. Our purpose here is not to describe GR in its full geometrical fashion but to jump as soon as possible to the equations that determine the dynamics of the fabric of spacetime, i.e. the Einstein field equations, and apply them to the cosmological scenario. Therefore we will claim that spacetime, from the perspective of GR, is a 4-dimensional differential manifold $\mathcal{M}$ carrying a torsion-free connection compatible with a (0,2)-tensor field called metric $g_{\mu\nu}$ without explaining in a mathematical rigorous manner what each term means (more details can be found in classical books, for instance [13, 14]). To gain intuition, we can picture the spacetime structure ($\mathcal{M}$, $g_{\mu\nu}$) as just a 4D version of a 2D rubber surface, like a toy balloon without its lip, that cannot have sharp bumps on it but can change its shape smoothly and can stretch or shrink. For us the important point is that the metric is a field, roughly speaking in physicists sense a field is any quantity that may assume different values at different points on the manifold, and therefore we can try to figure out an action for it and work out the equations of motion following the usual classical field theory machinery. Fortunately, around the time Einstein found his equations [15], the mathematician David Hilbert found an action that yielded the same equations of motion as the work of Einstein [16]. The so-called Einstein-Hilbert action is:

$$S_{E-H} = \frac{1}{2\kappa}\int d^4x\sqrt{-g}R\,; \tag{1.3}$$

where κ is just a proportionality constant, $g \equiv \det(g_{\mu\nu})$ is the needed factor to make $d^4x\sqrt{-g}$ invariant and R is the Ricci scalar which is defined as the trace of the Ricci curvature tensor $R \equiv g^{\mu\nu}R_{\mu\nu}$. In its turn the Ricci tensor is defined from the Riemann tensor as $R_{\mu\nu} \equiv R^{\rho}{}_{\mu\rho\nu}$. Moreover, for a torsion-free connection compatible with the metric we also can write the

Riemann tensor in terms of the Christoffel symbols:

$$R^{\rho}{}_{\mu\nu\sigma} = \partial_\nu \Gamma^{\rho}{}_{\mu\sigma} - \partial_\sigma \Gamma^{\rho}{}_{\mu\nu} + \Gamma^{\lambda}{}_{\mu\sigma}\Gamma^{\rho}{}_{\lambda\nu} - \Gamma^{\lambda}{}_{\mu\nu}\Gamma^{\rho}{}_{\lambda\sigma} \tag{1.4}$$

And the Christoffel symbols in terms of the inverse metric and derivatives of the metric:

$$\Gamma^{\rho}{}_{\mu\nu} = \frac{g^{\rho\sigma}}{2}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}) \tag{1.5}$$

As we can see from the above equations, although the action looks very simple in terms of the Ricci scalar, it actually involves first and second derivative of the metric. The important point is that R is a scalar field, so every observer, i.e. frame, will agree on its value at a specific point on the manifold.

We shall make a digression from the usual path of calculating the equations of motion for S_{E-H} for a while to comment on the role of the constant κ. First, note that the metric is dimensionless, $[g_{\mu\nu}] = 0$, since this is the structure that tells us how to measure distances in a space, $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$, and we can assume without loss of generality that $[ds^2] = 2[dx^\mu] = -2$. Actually, not all components of dx^μ need to have units of length but we can always redefine our components such that this is the case. Anyway, the important fact is $[g_{\mu\nu}] = 0$ and we will skip minor details. Thus $[R] = 2$, each term contains two derivatives, and since $[d^4x] = -4$ and an action is always dimensionless – it actually has dimensions of $\hbar$ which is dimensionless in the units we are using – we have $[\kappa] = -2$. Since we must reproduce Newton's law of gravity in the non-relativistic limit, we get $\kappa = 8\pi G$, where G is Newton's gravitational constant in 4 dimensions (see [14]). Furthermore, we can rewrite it in terms of the 4D reduced Planck mass $M_{Pl}^2 = \frac{1}{8\pi G} = \kappa^{-1}$. Note that $[\kappa] = -2 \Rightarrow [M_{Pl}] = 1\,;\,[G] = -2$ as expected. This discussion on the role of κ will be important when we talk about higher dimensions of spacetime. Now we end our digression and comeback to the equations of motion.

We get the Einstein field equations without presence of matter by requiring that $\delta S_{E-H} =$

0, from equation (1.3), when we vary the metric (or its inverse). If we want to include matter fields to this scenario we need to consider an action $S_m = \int d^4x\sqrt{-g}\mathcal{L}_m$ that is added to the Einstein-Hilbert one. Moreover, another simple addition we can make is to include a cosmological constant term $S_\Lambda = -\frac{1}{\kappa}\int d^4x\sqrt{-g}\Lambda$ to it, where Λ is just a spacetime constant. This cosmological constant term turned out to be very relevant to describe cosmology as we will see later on. All in all, we end up with $S = S_{E-H} + S_\Lambda + S_m$ and we require that $\delta S = 0$ when the metric is varied. The resulting equations of motion are, see appendix A:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}\,; \tag{1.6}$$

where $T_{\mu\nu} \equiv \frac{-2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}$ is the energy-momentum (or stress-energy) tensor that represents the content of energy and momentum of matter fields in our system, i.e. for cosmology in the whole universe.

Equation (1.6) indicates how matter, through the energy-momentum tensor in the right hand side of the equation, influences the shape of the fabric of spacetime, i.e. shape of the universe in our cosmological analysis since our physical system is the whole universe. Futhermore, since $S_m[\psi_i]$ contains a factor of the metric in it – at least through $\sqrt{-g}$ but generally not just this factor – if we calculate the equations of motion with respect to the matter fields ψ_i they will depend on how the universe is curved. In conclusion, GR tells us that gravity is not a force in the usual classical sense. It is actually geometry, i.e. the effects we think of being caused by a force are a reflect of the curvature of spacetime in our region. Both the matter content of the universe influence its shape and this shape influence the path of particles traveling through spacetime.

We end this section by noticing that the term $\Lambda g_{\mu\nu}$ coming from the action S_Λ can either be interpreted as a piece of the gravitational action and therefore lying in the gravitational side of the last paragraph's analysis or it can be put in the right hand side (RHS) of the equation (1.6) and compose another type of matter content of the universe, such that the new energy momentum tensor would be $\tilde{T}_{\mu\nu} = T_{\mu\nu} - \frac{\Lambda}{8\pi G}g_{\mu\nu}$. Both interpretations, of course, generate the same experimental predictions thus it is a matter of convenience interpreting the cosmological constant as a component of the matter content or not. The observations made in this paragraph will be important later on when we discuss the types of matter composing our model of the universe.

1.2.1 Friedmann-Lemaître-Robertson-Walker (FLRW) Metric

Now that we briefly summarized the main features of GR needed for us, we will look for a specif realization of the Einstein equations to our system or interest, the whole universe. The first thing to be noted is that the Ricci scalar and tensor in equation (1.6) have up to second derivatives of the metric in all sorts of complicated ways. Therefore, the Einstein equations are a set of highly nonlinear equations that are nearly impossible to be solved analytically without symmetry restrictions to the metric. That is where the observational facts discussed in the first section of this chapter comes in handy. Spatial homogeneity and isotropy at the largest scales mean that the spacetime is composed of spatial slices that lie on top of each other in a time ordered manner. The metric needs to have the form:

$$ds^2 = -dt^2 + dl^2_{phys}\,; \tag{1.7}$$

where dl^2_{phys} represents the spatial subspace that we are going to analyze in detail now. We must notice that isotropy forbids g_{0i} components to assume a non-zero value and that, in principle, we could have written $g_{00}d\tilde{t}^2$ instead of just dt^2, however we could always redefine $dt = \sqrt{g_{00}}d\tilde{t}$ in order to write the metric in the form of equation (1.7).

The observation of the Hubble expansion implies that the physical spatial slices change in side with time, therefore we can define comoving coordinates dl, i.e. coordinates that follow the expansion of the universe, such that $dl_{phys} = a(t)dl$. The function dictating the expansion of the universe $a(t)$ is called the scale factor. If we want dl to have units of length the scale factor is dimensionless. Nevertheless, for now, it is useful to attribute the length dimension to $a(t)$ and work with a dimensionless dl.

We must now figure out the possible forms of dl^2. Again, homogeneity and isotropy play an essential role. We must only consider maximally symmetric 3D spaces. There are three possibilities:

1. a spatially flat universe: dl^2 is the line element of a 3-dimensional Euclidean space.

$$dl^2 = \delta_{ij} dx^i dx^j = d\vec{x}^2 \tag{1.8}$$

2. a positively curved spatial subspace: a sphere S^3.

   The 3D sphere's line element can be defined from its embedding in a 4-dimensional Euclidean space as:

$$dl^2 = \delta_{ij} dx^i dx^j + dy^2 = d\vec{x}^2 + dy^2 \ ; \ \text{where: } \delta_{ij} x^i x^j + y^2 = \vec{x}^2 + y^2 = R^2 \tag{1.9}$$

   Note that by a redefinition of x^i and y we can rewrite the constraint as $\vec{x}^2 + y^2 = 1$ without loss of generality. Therefore, by taking differentials of the new constraint we get:

$$dy^2 = \frac{(\vec{x} \cdot d\vec{x})^2}{1 - \vec{x}^2} \ \Rightarrow \ dl^2 = d\vec{x}^2 + \frac{(\vec{x} \cdot d\vec{x})^2}{1 - \vec{x}^2} \tag{1.10}$$

3. a negatively curved spatial subspace: a 3D Lobachevsky space or a hyperboloid H^3.

   This space's line element can be defined from its embedding in a 4D Minkowski space as:

$$dl^2 = \delta_{ij} dx^i dx^j - dy^2 = d\vec{x}^2 - dy^2 \ ; \ \text{where: } \delta_{ij} x^i x^j - y^2 = \vec{x}^2 - y^2 = -R^2 \tag{1.11}$$

   Again we can get rid of the radius R for the same argument as in the previous case and work with the constraint $\vec{x}^2 - y^2 = -1$. Again we can calculate the differential dy in terms of the other coordinates and plug it back into equation (1.11).

$$dy^2 = \frac{(\vec{x} \cdot d\vec{x})^2}{1 + \vec{x}^2} \ \Rightarrow \ dl^2 = d\vec{x}^2 - \frac{(\vec{x} \cdot d\vec{x})^2}{1 + \vec{x}^2} \tag{1.12}$$

It is easy to see that we can write all three cases into one equation for the spacetime line element by introducing a constant k.

$$ds^2 = -dt^2 + a^2(t)\left[d\vec{x}^2 + k\frac{(\vec{x}\cdot d\vec{x})^2}{1-k\vec{x}^2}\right] \text{ ; where: } k \equiv \begin{cases} 0 & \text{flat space;} \\ +1 & \text{positively curved;} \\ -1 & \text{negatively curved.} \end{cases} \tag{1.13}$$

That is the famous Friedmann-Lemaître-Robertson-Walker (FLRW) metric. Some books may call it just Robertson-Walker (RW) or Friedmann-Robertson-Walker (FRW) metric. Since the name does not change the physics, it is just important to notice they are talking about the same metric.

Now, we may want to rewrite equation (1.13) such that the scale factor becomes dimensionless and dl recovers dimension of length. We are able to do that since the FLRW metric has a rescaling symmetry, meaning that if we change $a \to a/\lambda$, $x_i \to \lambda x_i$ and $k \to k/\lambda^2$ the metric remains the same. Thus, we can not only choose λ such that $[\lambda] = -1$ leading to a dimensionless $a(t)$ but also set $a(t_0) = 1$ where t_0 is the cosmological time (we will shortly define precisely what it means) of today. From now on in this thesis, we will consider a dimensionless scale factor such that its value is equal to one at the present time.

A perhaps more commonly found, for instance [12, 17], form of this metric is the realization of equation (1.13) in spherical coordinates. In these coordinates we have $d\vec{x}^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 = dr^2 + r^2 d\Omega^2$, $\vec{x}^2 = r^2$ and $\vec{x}\cdot d\vec{x} = rdr$, leading to:

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right) \tag{1.14}$$

One may also define a conformal time η such that $d\eta = dt/a(t)$. Then, equation (1.14) becomes:

$$ds^2 = a^2(\eta)\left(-d\eta^2 + \frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right) \text{ ;} \tag{1.15}$$

where $a(\eta)$ is the same $a(t)$ when we plug in the expression $t = t(\eta)$ that we can calculate by integrating $d\eta = dt/a(t)$. Moreover, notice that the reason for the name "conformal" time is that in equation (1.15) the scale factor works as a conformal factor of a conformal transformation, i.e. $\tilde{g}_{\mu\nu} = a^2(\eta) g_{\mu\nu}$.

Since we introduced the scale factor, we are now able to define the Hubble parameter from it. As we discussed in the beginning of the chapter, the Hubble parameter, first thought to be a constant, quantified the rate of expansion of the universe, therefore it is natural to express it in terms of the change (derivative) of $a(t)$. More precisely, from equation (1.1) and the fact that the speed of light is constant in the vacuum ($c = 1$):

$$z \equiv \frac{\lambda - \lambda_0}{\lambda_0} = \frac{1/\nu - 1/\nu_0}{1/\nu_0} = \frac{\nu_0 - \nu}{\nu}; \tag{1.16}$$

where ν_0 is the emitted frequency and ν is the observed one.

From the expression of the relativistic Doppler effect we can rewrite ν_0 in terms of ν as:

$$\nu_0 = \nu \sqrt{\frac{1+v}{1-v}} = \nu(1+v) + \mathcal{O}(v^2); \tag{1.17}$$

where v is the velocity of the source of light (in our case, a bright star in the sky) with respect to the detector (us on Earth). In our convention v is positive if the source is moving away from the detector. Thus, up to first order, $z = v$ and Hubble actually found a relation between the velocity of a star moving away from us with the distance to us.

Now, coming back to the implications the scale factor has on cosmological scales, recall that any physical distance can be written as the scale factor times a comoving distance, $\vec{r}_{phys} = a(t)\vec{r}$, and notice that we can calculate the physical velocity of any object moving with respect to us as:

$$\vec{v}_{phys} \equiv \frac{d\vec{r}_{phys}}{dt} = \frac{da}{dt}\vec{r} + a\frac{d\vec{r}}{dt} \equiv \frac{\dot{a}}{a}\vec{r}_{phys} + \vec{v}_{pec}\,; \tag{1.18}$$

where we defined the peculiar velocity $\vec{v}_{pec}$ as the velocity that is not due to the change of

$a(t)$ and the other component of $\vec{v}_{phys}$, $(\dot{a}/a)\vec{r}_{phys}$, is called the Hubble flow. If an observer follows the Hubble flow, i.e. if it is a comoving observer, then the scale factor would be just a constant from his or her perspective, therefore the observer would only measure the peculiar velocity.

Equation (1.18) shows us that the angular coefficient of the linear relation $z = v = Hr$ measured by Hubble is actually equal to:

$$H \equiv \frac{1}{a}\frac{da}{dt} \equiv \frac{\dot{a}}{a}\,; \tag{1.19}$$

and we arrive at our expression for the Hubble parameter in terms of the scale factor as promised.

One last comment is in order. In the beginning of the chapter, we skipped the fact that the universe is essentially isotropic in a special frame, the CMBR (or cosmic) frame, not in any frame. The CMBR frame is defined as the one in which this fluid that is composed of photons traveling around from the Big Bang is homogeneous and isotropic. Observations made on Earth point out that the CMBR has a dipole (see [12]). However, it is actually associated to the motion of Earth with respect to the CMBR frame. Therefore, since we constructed the FLRW metric from observations valid in the cosmic frame, the cosmological time which appears in equation (1.13) is defined in this special frame. It is with respect to this time that we compute the age of the universe, for example. At first sight, on special relativistic grounds, it seems weird that we can compute the age of the universe and usually do not say with respect to which observer. Nonetheless, we can do such a thing because we have a special cosmic frame in this case and it is perfectly fine according to GR.

1.3 Matter Content and the Friedmann Equations

We are almost ready to apply the Einstein equations, equation (1.6), to the universe as a whole. We already know the form of the metric which is going to be used, however we still have to describe the energy-momentum tensor of our system. First, we must recall the discussion of the first section, where we pointed out that, at the largest scales, galaxies and even clusters of galaxies are point-like objects. Therefore, if we could zoom out our perspective to these scales we would see a bunch of points potentially interacting with each other which could be heated up (or cooled down) if we find a mechanism to supply the system with (or extract from it) some energy. In conclusion, the components of the universe behave as a fluid. It is usual to model them as a perfect fluid - in this background cosmology description, on which we will consider perturbations later on - and we will restrict ourselves to this standard analysis in this thesis. A fluid is described by its pressure P, its energy density ρ and its comoving 4-velocity u^μ (or, equivalently, its dual covetor $u_\mu \equiv g_{\mu\nu}u^\nu$). The stress-energy tensor, i.e. the energy-momentum tensor, of a perfect fluid assumes the form:

$$T_{\mu\nu} = \rho u_\mu u_\nu + P(g_{\mu\nu} + u_\mu u_\nu) \tag{1.20}$$

It is important to notice that P and ρ are not two independent variables. In fact, we call the equation of state of an energy content in our universe the relation between its pressure and its energy density, $P = P(\rho)$. We will shortly see that we can classify all energy content in our universe basically in three categories. Heavy matter, with respect to the average temperature of the CMBR, is called non-relativistic matter, baryonic matter or just matter. Matter which is light with respect to the CMBR is called relativistic matter or just radiation. Finally, we also have an exotic type of matter called dark energy. Even though the relation between P and ρ for any type of energy doesn't have to be linear, the equation of state of these three energy contents can be written as $P = w\rho$, where w is a different constant for each type of matter. We will describe exactly what is the value of w for each energy content but first, for the sake of simplicity, let us assume only one energy content described by equation (1.20). We finally are able to work out the equations of motion of cosmology, the Friedmann equations, from the Einstein equations, equation (1.6), see appendix B for the derivation. The first Friedmann equation is:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{1.21}$$

And the second Friedmann equation is:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{1.22}$$

Since equations (1.21) and (1.22) tell us the evolution of the scale factor, they actually dictate how our universe behaves.

Another important equation which will be analyzed in more detail now is the continuity equation, see appendix B equation (B.26). Each matter component obey this equation.

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) \tag{1.23}$$

We will consider a equation of state $P = w\rho$ for a generic parameter $w \neq -1$. The specific, and relevant, case in which $w = -1$ will be described in the cosmological constant section. Soon we will derive the value of this parameter for each of the three main components of the universe and we will come back to the following general relations. From the continuity equation:

$$\frac{d\rho}{dt} = -3\rho(1+w)\frac{1}{a}\frac{da}{dt} \Rightarrow \int_{\rho^{(0)}}^{\rho(a)} \frac{d\tilde{\rho}}{\tilde{\rho}} = -3(1+w)\int_{a_0}^{a} \frac{d\tilde{a}}{\tilde{a}} \tag{1.24}$$

The subscript (or superscript) 0 indicate the value of quantities nowadays. Recall that we set $a_0 = 1$, then:

$$\rho(a) = \rho^{(0)} a^{-3(1+w)} \tag{1.25}$$

In equation (1.21), we can define a curvature energy density $\rho_k \equiv -\frac{3k}{8\pi G a^2}$ such that the equation becomes $H^2 = (8\pi G/3)(\rho + \rho_k)$. However, we will shortly discuss an experimental observation [18] which leads to $\rho^{(0)} \gg \rho_k^{(0)}$, i.e. an almost spatially flat universe, today. We will therefore drop the curvature contribution to the first Friedmann equation. In fact, we haven't showed that $\rho \gg \rho_k$ at all times and we could only drop this term if we integrate the expression between times close to today. Nevertheless we will drop this term and integrate to any time and just later we will argue that we could do that since ρ is indeed much greater than ρ_k at all times. Then, we can plug the result of equation (1.25) into equation (1.21) neglecting the curvature term to obtain:

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi G}{3}\rho^{(0)}}a^{-\frac{3}{2}(1+w)} \Rightarrow \int_0^{a(t)} da' a'^{\frac{1}{2}(1+3w)} = \sqrt{\frac{8\pi G}{3}\rho^{(0)}}\int_0^t dt'$$
$$\Rightarrow a(t) \propto t^{\frac{2}{3(1+w)}} \tag{1.26}$$

Note that in order to get to equation (1.26) we considered that $a \to 0$ as $t \to 0$. This is called the cosmological singularity. It is a natural expectation from the expansion of the universe observation. Furthermore, it would be hard to come up with a natural explanation, within this simple fluid model for matter components, for a universe that would have started off huge, then shrunk to a minimum and would be now expanding to break our assumption of a cosmological singularity.

From equation (1.26), the Hubble parameter can be related to time as:

$$a(t) = Ct^{\frac{2}{3(1+w)}} \Rightarrow \dot{a}(t) = Ct^{\frac{2}{3(1+w)}}\frac{2}{3(1+w)t} \Rightarrow H \equiv \frac{\dot{a}}{a} = \frac{2}{3(1+w)t} \propto a^{-\frac{3(1+w)}{2}} \tag{1.27}$$

From equation (1.27) it is easy to see that if $\frac{2}{3(1+w)} \sim 1$, which we will see happens for most of the history of our universe therefore is a good approximation, then $t \sim H^{-1}$ which means that the age of the universe is $t_0 \sim H_0^{-1}$ and the Hubble constant, i.e. the Hubble parameter evaluated at today's time, is $H_0 \simeq 70\,(\text{km/s})/\text{Mpc}$. Therefore, $H_0 \simeq 70 \times 3.2 \times 10^{-20} s^{-1} \simeq 7 \times 10^{-11}$ years$^{-1} \Rightarrow t_0 \sim 10$ billions years. We must point out that nowadays there are two independent ways of measuring H_0 which are not compatible, i.e. they differ from each other with a high confidence interval, see for example [18, 19, 20, 21, 22]. This incompatibility is the so-called Hubble tension and explaining what is going on is a current challenge for cosmologists. Nevertheless, for our order of magnitude calculation $H_0 \sim 100\,(\mathrm{km/s})/\mathrm{Mpc}$ was precise enough.

Finally, from equations (1.27) and (1.21) neglecting the curvature term, we find how the energy density of a type of matter whose equation of state parameter is $w \neq -1$ depends on time:

$$\rho(t) = \frac{3}{8\pi G}H^2 = \frac{3}{8\pi G}\left[\frac{2}{3(1+w)t}\right]^2 = \frac{1}{6\pi G(1+w)^2 t^2} \tag{1.28}$$

Notice that, although we were considering just one type of matter in our universe in the above analysis for simplicity, it is easy to generalize to many types of matter, each with its own w value. First, all types of matter have their own continuity equation (1.23). In the Friedmann equations (1.21) and (1.22), we just have to replace $\rho \to \sum_i \rho_i$ and $P \to \sum_i P_i$. Equation (1.25) is derived entirely from the continuity equation, thus is valid for each type of matter, taking into consideration their different values of w, of course. And equations (1.26), (1.27) and (1.28) are all valid if we have an energy density which dominate over the others such that $\sum_i \rho_i \simeq \rho_{dom}$ as long as we replace $w \to w_{dom}$.

A form of the first Friedmann equation that will be useful when we compare against observations is the following. First, let us define the critical density as:

$$\rho_c \equiv \frac{3H^2}{8\pi G} \tag{1.29}$$

And the ratios of the energy densities to the critical one as:

$$\Omega_i \equiv \frac{\rho_i}{\rho_c}\,; \tag{1.30}$$

where i stands for the different types of matter in the universe.

Considering ρ_k as we introduced before and therefore interpreting the curvature as a type of energy content, equation (1.21) can be rewritten as:

$$\frac{H^2}{H_0^2} = \frac{8\pi G}{3H_0^2}\sum_i \rho_i = \sum_i \frac{\rho_i^{(0)}}{\rho_c^{(0)}} a^{-3(1+w_i)} = \sum_i \Omega_i^{(0)} a^{-3(1+w_i)}\,; \tag{1.31}$$

where we have used the definitions (1.29) and (1.30) and the expression for $\rho = \rho(a)$ found in equation (1.25). The subscript 0 and the superscript (0) stand for quantities evaluated at today's time, as usual. Equation (1.31) is useful since CMBR data [18] are expressed in terms of $\Omega_i^{(0)}$ and H_0.

Now we should derive the equation of state parameter for relativistic, i.e. radiation, and non-relativistic matter, i.e. dust matter. First, it is important to point out that the matter contents of our universe are believed to have begun as a hot fluid in an equilibrium state, see book [12] chapter 5. This quickly reached equilibrium state allow us to describe the fluids composing the universe as gases within equilibrium thermodynamics. Remember that, from statistical mechanics, we get the Fermi-Dirac and Bose-Einstein distributions calculating the average number of particles with energy E for gases of fermions and bosons, respectively.

$$f(E) = \frac{1}{e^{\beta(E-\mu)} \pm 1}\,; \tag{1.32}$$

where $\beta = 1/T$, T is the temperature of the gas and μ is its chemical potential. The plus sign stands for Fermi-Dirac while the minus sign indicates a Bose-Einstein distribution.

For $T \ll E - \mu$, we get the Maxwell-Boltzmann distribution, f_{MB}:

$$\begin{aligned} f(E) &= \frac{1}{e^{\beta(E-\mu)}[1 \pm e^{-\beta(E-\mu)}]} = e^{-\beta(E-\mu)}[1 \mp e^{-\beta(E-\mu)} + \mathcal{O}\left((e^{-\beta(E-\mu)})^2\right)] \\ &= e^{-\beta(E-\mu)} + \mathcal{O}\left((e^{-\beta(E-\mu)})^2\right) \\ &\Rightarrow f_{MB}(E) = e^{-\beta(E-\mu)} \end{aligned} \tag{1.33}$$

For a general distribution $f(E)$, i.e. either Fermi-Dirac, Bose-Einstein or Maxwell-

Boltzmann, integral expressions for the number density (n), energy density (ρ) and pressure (P) of the gas can be written. Here the expressions are given below. For their derivation see appendix C.

$$n = \frac{g}{2\pi^2}\int_m^\infty dE\sqrt{E^2-m^2}Ef(E) \tag{1.34}$$

$$\rho = \frac{g}{2\pi^2}\int_m^\infty dE\sqrt{E^2-m^2}E^2f(E) \tag{1.35}$$

$$P = \frac{g}{6\pi^2}\int_m^\infty dE(E^2-m^2)^{3/2}f(E) \tag{1.36}$$

We are ready to use equations (1.32) and (1.33) to calculate w for the two relevant cases.

(i). Radiation ($w_r = 1/3$).

First, let us present a heuristic derivation which will add intuition to the obtained result. Imagine a box full of relativistic particles, e.g. photons. If we cannot add nor remove energy from this box, with the expansion of the universe, its energy density will drop because $\rho_r = E/V$ and the volume is increasing as $V = V_{com}a^3$, where the subscript stands for comoving. However, while that happens, the wavelengths of the photons also expand and since $E \propto \lambda^{-1} = \lambda_{com}^{-1}a^{-1}$. Adding these two contributions we have $\rho_r \propto a^{-4}$. If we compare this result with equation (1.25) it is easy to figure out that $-3(1+w_r) = -4 \therefore w_r = 1/3$.

Now, back to our thermodynamic calculations, if we take the ultrarelativistic limit, $E \gg m$, of equations (1.35) and (1.36) it is easy to see that:

$$P = \frac{\rho}{3} = \frac{g}{6\pi^2}\int_m^\infty dEE^3f(E) \Rightarrow w_r = \frac{1}{3}; \tag{1.37}$$

which confirms our heuristic derivation.

(ii). Non-relativistic/dust matter ($w_m = 0$).

Again, first we give a heuristic argument to the derivation. The only difference in our thought experiment in comparison to the previous relativistic case is that the particles in the box need not to be thought of as photons, which had an associated wavelength, therefore the only decreasing component to the energy density comes from the expanding volume of the box. Thus, $\rho_m \propto a^{-3}$ which after comparing against equation (1.25) gives $-3(1+w_m) = -3 \therefore w_m = 0$.

Back to the thermodynamic analysis, we will have to compute n, ρ and P this time. The non-relativistic limit to be considered is $T \ll \mu$ and also $m^2 \gg p^2$ such that:

$$E^2 = m^2 + p^2 \therefore E = m\sqrt{1+\frac{p^2}{m^2}} = m\left[1 + \frac{1}{2}\frac{p^2}{m^2} + \mathcal{O}\left(\frac{p^4}{m^4}\right)\right] \simeq m + \frac{p^2}{2m} \tag{1.38}$$

Note that in the non-relativistic limit we are dealing with a Maxwell-Boltzmann gas. From equation (C.2), we have:

$$n = \frac{g}{(2\pi)^3}\int_0^\infty dp 4\pi p^2 e^{-\beta(E(p)-\mu)} \simeq \frac{g}{2\pi^2}\int_0^\infty dp p^2 e^{-\beta(m-\mu)} e^{-\frac{\beta p^2}{2m}} \; ; \tag{1.39}$$

where in the second equality we used the approximation of equation (1.38).

Recall the definition of the gamma function:

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx \; ; \text{ for } \operatorname{Re}(z) > 0 \tag{1.40}$$

Then we can rewrite the integral of equation (1.39) as:

$$\begin{aligned} n &\simeq \frac{g}{2\pi^2} e^{-\frac{m-\mu}{T}} (2mT)^{3/2} \int_0^\infty dx e^{-x^2} x^2 = \frac{g}{4\pi^2} e^{-\frac{m-\mu}{T}} (2mT)^{3/2} \int_0^\infty dx e^{-x} x^{1/2} \\ &\simeq \frac{g}{4\pi^2} e^{-\frac{m-\mu}{T}} (2mT)^{3/2} \Gamma(3/2) = g\left(\frac{mT}{2\pi}\right)^{3/2} e^{-\frac{m-\mu}{T}} \end{aligned} \tag{1.41}$$

where we used the fact $\Gamma(3/2) = \sqrt{\pi}/2$.

For the energy density, from equation (C.4), we can see that if we consider only the 0th order in the expansion for the energy, namely $E(p) \simeq m$, we have:

$$\rho = \frac{g}{(2\pi)^3} \int d^3p E(p) f(E(p)) \simeq m \left[\frac{g}{(2\pi)^3} \int d^3p f(E(p)) \right] = mn \tag{1.42}$$

Finally, if we keep only $E \simeq m$ in the non-exponential piece of the integral expression for the pressure of the gas, equation (C.6), becomes:

$$\begin{aligned} P &= \frac{g}{(2\pi)^3} \int d^3p \frac{p^2}{3E} f(E(p)) \simeq \frac{g}{2\pi^2} \frac{1}{3m} \int_0^\infty dp p^4 e^{-\frac{m-\mu}{T}} e^{-\frac{p^2}{2mT}} \\ &\simeq \frac{g}{2\pi^2} \frac{1}{3m} e^{-\frac{m-\mu}{T}} (2mT)^{5/2} \int_0^\infty dx e^{-x^2} x^4 = \frac{g}{4\pi^2} \frac{1}{3m} e^{-\frac{m-\mu}{T}} (2mT)^{5/2} \int_0^\infty dx e^{-x} x^{3/2} \\ &\simeq \frac{g}{4\pi^2} \frac{1}{3m} e^{-\frac{m-\mu}{T}} (2mT)^{5/2} \Gamma(5/2) = \left[g \left(\frac{mT}{2\pi} \right)^{3/2} e^{-\frac{m-\mu}{T}} \right] T \simeq nT \end{aligned} \tag{1.43}$$

We conclude that, from equations (1.42) and (1.43), $P \simeq nT \ll nm \simeq \rho$ thus $P/\rho \to 0$ and we indeed get $w_m = 0$ as anticipated.

Now that w_r and w_m are known, note that, since $\rho_r \propto a^{-4}$ and $\rho_m \propto a^{-3}$, the total energy density, i.e. sum of all components, tends to be dominated by radiation in the early history of the universe but then, after a while, dust matter energy density overcomes the former radiation dominance. We will soon discuss dark energy and add the last relevant component in order to tell the history of the universe.

1.3.1 Dark Matter

If we want to observe the amount of energy density for each type of matter in the universe today $\rho_i^{(0)}$, or their ratios to the critical density $\Omega_i^{(0)}$, we would realize there is a discrepancy in the amount of non-relativistic matter which is measured to exist via gravitational observations and which can be seen, i.e. which interacts electromagnetically and therefore we can see their brightness through telescopes. There is much more matter than what we can see according to gravitational observations. The conclusion is that there is an exotic type of matter, dark matter, in the universe which adds to the usual baryonic matter, i.e. any matter composed of Standard Model particles and which interacts with light, increasing the expected gravitational pull in distant galaxies. The name “dark matter” was given since we cannot directly observe them, i.e. they don’t shine, thus looking to the sky we would see just darkness instead of them. Yet, from their gravitational effects, they are known to be there. Actually, ignoring gravitational lensing effects, we would see through them, since they don’t interact with light, and therefore perhaps “transparent matter” would be a better name. Nevertheless the name got stuck and we will keep calling it dark matter.

Here, I will quickly discuss a few observations that led to conclusion which was pointed out last paragraph in order to properly define the ΛCDM model later. The main goal of this thesis is not to study dark matter, even though its composition is a relevant open question, therefore to find a more comprehensive list of observations see chapter 4 of book [12].

The first evidence for dark matter was due to Zwicky and dates back to 1933 [23]. Zwicky was analyzing the Coma Cluster and realized that the visible mass was much less than it was expected to be. Since in this cluster galaxies can already be approximated as points, he modeled the cluster as a gas with gravitational interaction among components. He noticed, using the virial theorem, that the visible mass was way less than the expected to counterbalance the kinetic term. Something seemed off, nowadays we know that is because $m_{dark} \gg m_{visible}$ and he was not taking dark matter into consideration since he could not see it.

Another evidence for dark matter is the unusual feature of light being distorted in the sky by no visible object. This phenomenon is explained by the presence of an invisible matter in the path between Earth and the light ray coming to us. In GR massive objects bend the path of light and that is why we observe such a distortion of light for apparently no good reason.

The last observation we are going to discuss here which led to the conclusion that there is such an exotic type of matter in our universe is the CMBR data. The Planck satellite collected data of the radiation coming to us from the primordial times of the universe. The latest published paper with its analyzed data is [18]. We will come back to it later but for now it is enough to say they are able to figure out both $\Omega_{dm}^{(0)}$ and $\Omega_b^{(0)}$ – in our notation dm stands for dark matter and b for baryonic matter – from the fitting of the six parameters of the ΛCDM model to observations. The best fitting generates the values:

$$\Omega_{dm}^{(0)} \simeq 0.26 \text{ ; and } \Omega_b^{(0)} \simeq 0.05\,; \tag{1.44}$$

which agrees with the previous observation by Zwicky that there is much more dark matter than visible matter in the universe.

1.3.2 Cosmological Constant

As anticipated, there is yet another relevant type of matter due to the existence of a non-zero cosmological constant in equation (1.6) which dictates the behaviour of the scale factor with time through the Friedmann equations. The cosmological constant term in the Einstein equations is being interpreted here as another type of energy content and, as we have already discussed, a stress-energy tensor of the form $T_{\mu\nu}^{(\Lambda)} = -\frac{\Lambda}{8\pi G} g_{\mu\nu}$ is associated to it. It is easy to figure out the equation of state parameter for the cosmological constant as:

$$T^{(\Lambda)\mu}{}_{\nu} = -\frac{\Lambda}{8\pi G}\delta^{\mu}{}_{\nu} = \text{diag}(-\rho_\Lambda, P_\Lambda, P_\Lambda, P_\Lambda) \Rightarrow P_\Lambda = -\rho_\Lambda = -\frac{\Lambda}{8\pi G} \Rightarrow w_\Lambda = -1 \tag{1.45}$$

Therefore our analysis of equations (1.24) through (1.28) is not valid for the cosmological constant parameter. Some weird features of the cosmological constant will be pointed out now. First, from the continuity equation (1.23), we get $\dot{\rho}_\Lambda = 0$ which means the energy density is constant with time, even with the expansion of the universe. This fact seems strange when we try to think of a box containing this type of matter and expanding according to the

Hubble flow, as we did for radiation and matter. This exotic type of energy is not diluted as the volume expands, its energy has to grow $\propto a^3$ to compensate the increasing factor in the volume due to the Hubble flow. However, as we will discuss in detail in chapter 2, a $w = -1$ equation of state parameter can be associated with the vacuum energy of quantum fields, thus since when the box is expanded by the Hubble flow there is automatically an increasing volume of vacuum – the "amount of vacuum" increases – inside the box justifying the weird fact of $\dot{\rho}_\Lambda = 0$. The association of the cosmological constant with a vacuum energy is problematic and is further discussed in chapter 2. This problem is so central to this thesis that it deserves a separate chapter.

From equation (1.21), again dropping the curvature term for the reasons already mentioned, we realize that H is constant since ρ_Λ in an universe dominated by the cosmological constant energy density – like ours as we will shortly see. This implies an exponential behaviour for $a = a(t)$.

$$\int_1^{a(t)} \frac{da'}{a'} = \int_{t_0}^{t} H dt' \therefore a(t) = e^{H(t-t_0)} \tag{1.46}$$

Finally, we must note that, from the Planck data [18], $\Omega_\Lambda^{(0)} \simeq 0.69$ for the cosmological constant. Therefore our universe is indeed dominated by this type of energy today. This leads to the conclusion $\Lambda > 0$ which in turns means we live in a de Sitter space, named after Willem de Sitter. For our purposes, a de Sitter or dS (anti-de Sitter or AdS) space is a space whose size grows (shrinks) exponentially with time, as it is our case from equation (1.46). More generally, a dS (AdS) space is a maximally symmetric Lorentzian manifold with constant positive (negative) Ricci curvature. The relation between the curvature and the cosmological constant is easily obtained by taking the trace of the Einstein equations (1.6) without matter, i.e. pure gravity:

$$g^{\mu\nu} R_{\mu\nu} - \frac{1}{2} g^{\mu\nu} g_{\mu\nu} R + g^{\mu\nu} g_{\mu\nu} \Lambda = 0 \therefore R = 4\Lambda \tag{1.47}$$

1.4 The Inflationary Paradigm

Up until now, we have constructed a model which is successful in describing the early history of our universe and can be used also to predict the its future. Nonetheless, even the pre-inflationary ΛCDM model, also called Hot Big Bang cosmology, has its problems when we try to extend its predictive power close to the cosmological singularities. Three problems will be presented here to indicate where exactly this model fails. See book [12] chapter 8 for a more comprehensive list.

1. The flatness problem

   We have mentioned that the curvature energy density is negligible today compared with other types of energy density, $\rho_k^{(0)} \ll \sum_i \rho_i^{(0)}$. This is not a problem per se, however within our framework we are able to realize that something strange happens at the beginning of the universe. First note that we can rewrite the first Friedmann equation (1.21) as:

   $$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \therefore 1 = \frac{\rho}{\rho_c} - \frac{k}{a^2H^2} \therefore \Omega - 1 = \frac{k}{a^2H^2} \tag{1.48}$$

   where $\rho \equiv \sum_i \rho_i$ and similarly $\Omega \equiv \sum_i \Omega_i = \sum_i \rho_i/\rho_c$ with i running through all types of energy excluding the curvature. In the first $\therefore$ we have just divided the equation by H^2 and used the definition of critical density, $\rho_c \equiv \frac{3H^2}{8\pi G}$.

   Notice from equation (1.48) that $\Omega(t) \to 1$ means the universe is nearly flat at time t. From equation (1.26) and (1.27), valid for $w \neq 1$, it can be concluded that $\Omega - 1 \propto t^{\frac{2}{3}\frac{1+3w_{dom}}{1+w_{dom}}}$. Therefore $\Omega - 1$ increases with time (decreases going back in time) in periods of domination by a type of energy whose equation of state parameter is either $w_{dom} < -1$ or $w_{dom} > -1/3$. Although we are living in a cosmological-constant-dominated period, this dominance started just recently – compared with the age of the universe – and most of our history was dominated by either matter or radiation, both with $w > -1/3$.

Therefore the weird fact is that not only $|\Omega^{(0)} - 1| \ll 1$ today but the flatness tended to be even more severe in the beginning of the universe. We must seek for an explanation of this fine-tuning, $|\Omega(t) - 1| \lll 1$ for $t < t_0$, in the early universe.

Inflation [24, 25, 26, 27, 28] solves the problem assuming a period of rapid accelerated expansion in the beginning of the universe. This paradigm started as what we call today "old inflation", but we will be concerned only with the "new inflation" here. In order to get an accelerated expansion $\ddot{a} > 0$, from equation (1.22), we need $w_{dom} < -1/3$. It is usually considered an exponential growth of space, obtained for $w_{dom} = -1$, i.e. a cosmological-constant-dominated period. Despite being favoured by experiments [18], this is not a requirement though. One usually considers, for simplicity, a scalar field model obeying slow-roll conditions, which will be described in the next section. The important facts needed in this discussion to solve the problems are the, at least nearly, exponential growth of the scale factor and the definition of the number of e-folds N_e, which quantifies the amount of inflation that happens, as:

$$N_e \equiv \log\left[\frac{a(t_e)}{a(t_b)}\right] = H_I(t_e - t_b) \equiv H_I \Delta t\,; \tag{1.49}$$

where in the second equality we have used the fact $a(t) \propto e^{H_I t} \therefore a(t_e) = a(t_b)e^{H_I(t_e - t_b)}$ with t_e being the moment in which inflation ends and t_b its beginning. Recall that we are considering $w \simeq -1$ thus H_I is constant.

Back to the flatness problem, we saw in equation (1.48) that $\Omega(t) - 1$ decreases in the past for $w_{dom} > -1/3$, but during inflation $\Omega(t) - 1 \propto a^{-2} \propto e^{-2H_I t}$ which increases going back in time. Therefore, we get rid of the fine-tuning since $\Omega - 1$ could have assumed any value at the beginning of the universe and inflation is the mechanism which forces our universe to be almost spatially flat at the beginning of the radiation-dominated era. It is as if our universe was the surface of a deflated balloon and inflation was a person who blew it up stretching its surface and flattening it out.

Quantitatively, if we assume $\Omega(t_b) - 1 \sim \mathcal{O}(1)$, then:

$$\begin{aligned}\Omega^{(0)} - 1 &= \frac{k}{a_0^2 H_0^2} = \frac{k}{H_0^2} = \frac{k}{a_e^2 H_I^2}\frac{a_e^2 H_I^2}{H_0^2} = \frac{k}{a_b^2 H_I^2} e^{-2N_e}\frac{a_e^2 H_I^2}{H_0^2} \\ &= [\Omega(t_b) - 1] e^{-2N_e}\frac{a_e^2 H_I^2}{H_0^2} \sim e^{-2N_e}\frac{a_e^2 H_I^2}{H_0^2}\end{aligned} \tag{1.50}$$

Since $\Omega^{(0)} - 1 \ll 1$ we have the constraint:

$$e^{N_e} > \frac{a_e H_I}{H_0} \tag{1.51}$$

Based on some expectations of the energy scale in which inflation should have occurred, the usual number of e-folds needed is ~ 60, see book [12] chapter 8.3 for details.

2. The horizon problem

The CMBR observed today is basically photons coming to us from a special moment of the universe's evolution which goes by the name of last scattering surface, more details will be given shortly on CMBR's own subsection. The important fact for our discussion now is that the photons come to us from this specific and early time t_{ls}, where the subscript stands for last scattering. This radiation is isotropic to a high degree of precision, as anticipated, therefore if we look for opposite directions in the sky we would observe approximately the same distribution of energies for the photons. The problem is, within pre-inflationary cosmology, those regions in the skies are causally disconnected. Why would we observe almost the same things from causally disconnected regions? It would be only expected from regions that were once in causal contact and had the time to thermalize.

To be more concrete, we must calculate the distance d_{ls} traveled by light from the beginning of the universe, at $t = 0$, to the surface of last scattering, t_{ls}, and the distance d_0 traveled by it from t_{ls} to us today, at $t = t_0$, in order to compare the two of them. From $ds^2 = -dt^2 + a^2 dl^2$ plus the fact that light travels on lightlike trajectories, $ds^2 = 0$, the desired (comoving) distances are calculated as:

$$d_{ls} = \int_0^{t_{ls}} \frac{dt}{a(t)} \text{ ; and } d_0 = \int_{t_{ls}}^{t_0} \frac{dt}{a(t)} \tag{1.52}$$

Some additional information are in order here. Although we skipped a description of the early history of the universe, it is known that before and after the last scattering surface the universe was dominated by matter for most of the time, therefore we will use $a(t) \propto t^{2/3} \therefore a(t) = a_0(t/t_0)^{2/3} = (t/t_0)^{2/3}$ in both integrals.

$$d_{ls} = \int_0^{t_{ls}} dt \left(\frac{t}{t_0}\right)^{-2/3} = 3t_0^{2/3} t_{ls}^{1/3} \text{ ; } d_0 = \int_{t_{ls}}^{t_0} dt \left(\frac{t}{t_0}\right)^{-2/3} = 3t_0^{2/3}(t_0^{1/3} - t_{ls}^{1/3}) \tag{1.53}$$

If we want to explain how two photons from opposite regions of the sky are strongly correlated, one of these photons must have passed through the point in this comoving space where the second photon was sent to us at the last scattering surface at some point before t_{ls}. In other words, one of the photons should have travelled at least the comoving diameter of nowadays observable universe from $t = 0$ to $t = t_{ls}$ before traveling the radius of the universe until it gets to us now, $d_{ls} \geq 2d_0$, which means:

$$N \equiv \frac{2d_0}{d_{ls}} \leq 1 \text{ ; for causal connection} \tag{1.54}$$

However, its known that the surface of last scattering happened at redshift $z_{ls} \simeq 1300$. Thus:

$$N = 2\frac{t_0^{1/3} - t_{ls}^{1/3}}{t_{ls}^{1/3}} \simeq 2\left(\frac{t_0}{t_{ls}}\right)^{1/3} = 2a_{ls}^{-1/2} = 2\sqrt{1 + z_{ls}} \simeq 72\,; \tag{1.55}$$

where in first equality the results of equation (1.53) were used, in the second the fact $t_0 \gg t_{ls}$ was used, in the third we used $a(t) = a_0(t/t_0)^{2/3}$ and, finally, the last equality comes from equation (1.16) plus the fact that wavelengths are stretched by the Hubble flow as any physical length, $\lambda \propto a^{-1}$.

$$z \equiv \frac{\lambda - \lambda_0}{\lambda_0} = \frac{1}{a} - 1 \therefore \frac{1}{a} = 1 + z \tag{1.56}$$

Thus, equation (1.55) shows us we need a mechanism to increase d_{ls} (or decrease d_0) by, at least, a factor of ~ 70 if we aim to explain the similarity between two distant regions in the sky.

Inflation solves this problem since the two opposite points in our universe today would have been once close together before inflation begins. After the last scattering surface, inflation is supposed to have ended thus the calculation of d_0 remains unchanged, we will just rewrite it. However, inflation plays a significant role in the determination of d_{ls} so we must recalculate it taking inflation into consideration now. From equation (1.52), considering the integral is dominated by the inflationary period that happens before t_{ls}:

$$d_{ls} \simeq \int_{t_b}^{t_e} dt \frac{e^{H_I(t_e - t)}}{a_e} = -\frac{e^{H_I t_e}}{a_e H_I}(e^{-H_I t_e} - e^{-H_I t_b}) = \frac{1}{a_e H_I}(e^{N_e} - 1) \simeq \frac{1}{a_e H_I} e^{N_e} \tag{1.57}$$

Recalling that we are considering a matter-dominated era before and after the last scattering surface, since it is the case for most of the time, then $H_0 = 2/3t_0$. Therefore, d_0 is rewritten as:

$$d_0 = 3t_0^{2/3}(t_0^{1/3} - t_{ls}^{1/3}) = 3t_0 \left[1 - \left(\frac{t_{ls}}{t_0}\right)^{1/3}\right] \simeq \frac{2}{H_0} \tag{1.58}$$

From equation (1.54), in order to the regions to be causally connected in the past it is needed that:

$$N = \frac{4}{H_0} a_e H_I e^{-N_e} \leq 1 \therefore e^{N_e} > \frac{a_e H_I}{H_0}\,; \tag{1.59}$$

where we have just plugged the results of equations (1.57) and (1.58) into the definition of N. The obtained constraint on N_e is the same as the derived one for the flatness problem.

3. The monopole/relic problem

   As the universe cooled down from its initial hot state it is believed that it could have passed through some phase transitions, for instance the period when a grand unified theory (GUT) symmetry is broken, when topological defects are expected to be created. Different kinds of relics or topological defects are expected to be created, however magnetic monopoles will be our focus here and at the end we will argue that this is the most relevant case indeed.

   By the Kibble mechanism [29, 30] we would expect one monopole per horizon as well as one nucleon per horizon created during a phase transition. Without inflation we would expect this 1 : 1 proportion to be maintained up until today. However, it is observed less than 10^{-30} monopoles per nucleon on Earth materials, see [31] chapter 4.1, so we must have some mechanism to dilute the concentration of monopoles away but don't do the same to nucleons. Inflation does exactly this job. Actually the rapid expansion of inflation dilutes away both concentrations, nevertheless inflation is followed by a transition period named reheating before the standard Hot Big Bang scenario and during reheating the scalar field responsible for inflation, namely the inflaton, decay into mostly radiation which will later create more nucleons. All in all, the density of nucleons is expected to remain the same as if there wasn't inflation, thus the overall effect of inflation is to dilute away only the monopoles. Therefore a 10^{-30} reduction on the ratio of monopoles to nucleons implies an 10^{30} expansion of the volume of the universe which, in turns, implies a growth of 10^{10} in the scale factor. That is the least amount of inflation we must have to avoid incompatibility with experiments.

$$N_e \geq \log\left(10^{10}\right) \simeq 23 \tag{1.60}$$

It should be pointed out that, in order for the inflation mechanism to solve the problem, the phase transition must have happened before it, or at least in the middle of it. If it happens in the middle, $N_e \gtrsim 23$ should be the least amount of inflation which happened after the phase transition.

Other relics give us less stringent bounds on the number of e-folds and are therefore redundant. The minimum amount of dilution needed for other relics is a 10^{-11} reduction in their number density, see [17] chapter 7.5. It should be pointed out, though, that within holographic cosmology, an alternative scenario to inflation, the origin of the monopole problem is different from other types of relics and they will be discussed separately in chapter 5.

It is clear that the above problems are basically solved by only one feature of inflation, the rapid and long expansion of space before the start of the Hot Big Bang period. A reheating consideration was needed to solve the monopole problem though. Other problems, such as why we have large entropy nowadays or how perturbations which formed galaxies today were generated, are solved in a similar fashion, see [12].

1.4.1 Slow-Roll Inflation

As anticipated, the current most accepted paradigm for inflation is the "new inflation". As it is usual practice, we will be concerned with scalar field models for inflation with an almost exponential growth of the scale factor, i.e. an approximate cosmological constant. The scalar field goes by the name of inflaton, ϕ, and it obeys some conditions, known as slow-roll conditions, in order to give us the features want for inflation. To find the slow-roll conditions we must define the two slow-roll parameters first. Notice that the horizon problem was solved because the rapid expansion of the universe made physical scales, which were once in causal contact, go outside the horizon where they got frozen, i.e. there wasn't any new quantum fluctuation to destroy the correlation, and in the radiation and matter-dominated eras, which came after inflation, these scales came back inside it. The horizon, which is understood as the region in causal contact, is quantified here by the Hubble radius $r_H \equiv H^{-1}$. This is the simplest type of horizon, but not the only, and it is approximately constant during inflation, H_I^{-1}. Therefore, we need the comoving Hubble radius, defined as the physical one divided by the scale factor as usual, to shrink during inflation.

$$\frac{d}{dt}(aH)^{-1} = -\frac{\dot{a}H + a\dot{H}}{(aH)^2} = \frac{1}{a}\left(-\frac{\dot{H}}{H^2} - 1\right) < 0 \Rightarrow \varepsilon_H \equiv -\frac{\dot{H}}{H^2} \ll 1 \tag{1.61}$$

The first slow-roll parameter is ε_H and its role, as we saw, is to ensure the shrinking of the comoving Hubble radius, among other things. Since we are considering approximate cosmological constant models $\dot{H} \to 0 \therefore \varepsilon_H \ll 1$ not only $\varepsilon_H < 1$. The shrinking of the above radius also implies an accelerated expansion:

$$\frac{d}{dt}(aH)^{-1} = \frac{d}{dt}\left(\frac{1}{\dot{a}}\right) = -\frac{\ddot{a}}{\dot{a}^2} < 0 \therefore \ddot{a} > 0 \tag{1.62}$$

The second requirement, which defines the second slow-roll parameter η_H, is that we need ε_H to be small for a sufficiently large amount of time or, equivalently, for a large number of e-folds N_e. Therefore we define:

$$\eta_H \equiv \frac{d\log\varepsilon_H}{dN_e} = \frac{1}{\varepsilon_H H}\frac{d\varepsilon_H}{dt} = \frac{\dot{\varepsilon}_H}{H\varepsilon_H}\,; \tag{1.63}$$

where it was used the fact $N_e \equiv \log(a_e/a_i) = \log e^{Ht} \therefore dN_e = Hdt$, since $H \simeq H_I = \text{constant}$, in the second equality. The second slow-roll condition is $|\eta_H| \ll 1$.

Now, it is time to associate ε_H and η_H, whose definitions are valid in general, with characteristics of the scalar field in the scalar field model for inflation. Some comments on the inflaton model are in order. Initially, the potential $V(\phi)$ was thought to have to be composed of a plateau, nearly flat and large, followed by a drop. Nowadays, it is known that many types of potential can be used in inflationary models, as long as the slow-roll conditions are satisfied. From the action:

$$S_\phi = \int d^4x\sqrt{-g}\left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right] \tag{1.64}$$

The stress-energy tensor is found by varying the action with respect to the inverse metric:

$$\delta S_\phi = \int d^4x\left\{\delta(\sqrt{-g})\left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right] - \sqrt{-g}\delta g^{\mu\nu}\frac{1}{2}\partial_\mu\phi\partial_\nu\phi\right\} \tag{1.65}$$

Recall from equation (A.5) that $\delta(\sqrt{-g}) = -(1/2)\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}$. Thus:

$$\begin{aligned}\delta S_\phi &= \int d^4x\sqrt{-g}\delta g^{\mu\nu}\left[-\frac{1}{2}\partial_\mu\phi\partial_\nu\phi + \frac{1}{4}g_{\mu\nu}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi + \frac{1}{2}g_{\mu\nu}V(\phi)\right]\\ \Rightarrow T_{\mu\nu} &\equiv -\frac{2}{\sqrt{-g}}\frac{\delta S_\phi}{\delta g^{\mu\nu}} = \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}g^{\rho\sigma}\partial_\rho\phi\partial_\sigma\phi - g_{\mu\nu}V(\phi)\end{aligned} \tag{1.66}$$

The background inflaton is expected to preserve homogeneity and isotropy we observe nowadays in the universe, therefore $\phi = \phi(t)$, leading to:

$$T^\mu{}_\nu = g^{\mu\rho}\partial_\rho\phi\partial_\nu\phi - \delta^\mu{}_\nu\left[-\frac{1}{2}\dot{\phi}^2 + V(\phi)\right] \Rightarrow \begin{cases} T^0{}_0 = -\rho_\phi = -\frac{\dot{\phi}}{2} - V(\phi)\,; \\ T^i{}_j = \delta^i{}_j P_\phi = \delta^i{}_j\left[\frac{\dot{\phi}^2}{2} - V(\phi)\right]. \end{cases} \tag{1.67}$$

where $\dot{\phi} \equiv d\phi/dt$ as usual.

From equation (1.67) it is easy to see that, in general:

$$w_\phi \equiv \frac{P_\phi}{\rho_\phi} = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} \tag{1.68}$$

The first Friedmann equation (1.21) for a universe composed solely by the inflaton becomes:

$$H^2 = \frac{8\pi G}{3}\rho_\phi - \frac{k}{a^2} = \frac{1}{3M_{Pl}^2}\left[\frac{\dot{\phi}^2}{2} + V(\phi)\right] - \frac{k}{a^2}\,; \tag{1.69}$$

where we will ignore the term k/a^2 from now on, since it goes to zero quickly because of the exponential behaviour of the scale factor.

Recall that the Friedmann equations are the equations of motion for the metric, i.e. obtained by varying the action with respect to the metric. The equation of motion for the inflaton is the Klein-Gordon equation, in FLRW spacetime though.

$$\begin{aligned}
\delta(S_{E-H} + S_\phi) = \delta S_\phi &= \int d^4x\sqrt{-g}[-g^{\mu\nu}\partial_\mu(\delta\phi)\partial_\nu\phi - V'(\phi)\delta\phi] \\
&= \int d^4x[\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi)\delta\phi - \sqrt{-g}V'(\phi)\delta\phi] \\
&= \int d^4x\sqrt{-g}[\nabla_\mu(g^{\mu\nu}\partial_\nu\phi)\delta\phi - V'(\phi)\delta\phi] \\
&= \int d^4x\sqrt{-g}\delta\phi[g^{\mu\nu}\nabla_\mu(\partial_\nu\phi) - V'(\phi)]\,;
\end{aligned} \tag{1.70}$$

where in the third equality we just partially integrated the first term, considering a boundary condition that vanishes the boundary term. In the fourth equality we used the identity $\partial_\mu(\sqrt{-g}X^\mu) = \sqrt{-g}\nabla_\mu X^\mu$ valid for any tensor field X^μ, as proved in equation (A.11). Finally, in the last equality, $\nabla_\mu g^{\mu\nu} = 0$ was used.

Thus, the equation of motion is:

$$\begin{aligned}
\frac{\delta}{\delta\phi}(S_{E-H} + S_\phi) = 0 \therefore \Box\phi - V'(\phi) = g^{\mu\nu}(\partial_\mu\partial_\nu\phi - \Gamma^\rho{}_{\mu\nu}\partial_\rho\phi) - V'(\phi) = 0 \\
-\ddot{\phi} - g^{ii}\Gamma^0{}_{ii}\dot{\phi} - V'(\phi) = 0 \\
\ddot{\phi} + \left[\frac{1-kr^2}{a^2}\left(\frac{a\dot{a}}{1-kr^2}\right) + \frac{1}{a^2r^2}(a\dot{a}r^2) + \frac{1}{a^2r^2\sin^2\theta}(a\dot{a}r^2\sin^2\theta)\right]\dot{\phi} + V'(\phi) = 0 \\
\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,;
\end{aligned} \tag{1.71}$$

where Γs and the inverse metric were taken from appendix B.

Our goal now is to relate ε_H and η_H to the potential $V(\phi)$ seeking to understand how the slow-roll conditions constrain the potential. Taking the time derivative of equation (1.69) and applying the result of equation (1.71) we get:

$$2H\dot{H} = \frac{1}{3M_{Pl}^2}[\dot{\phi}\ddot{\phi} + V'(\phi)\dot{\phi}] = \frac{\dot{\phi}}{3M_{Pl}^2}(-3H\dot{\phi}) \therefore \dot{H} = -\frac{\dot{\phi}^2}{2M_{Pl}^2} \tag{1.72}$$

Now, from the first slow-roll condition, it is clear why the name "slow-roll" inflation from the following:

$$\varepsilon_H \ll 1 \therefore |\dot{H}| \ll H^2 \therefore \frac{\dot{\phi}^2}{2M_{Pl}^2} \ll \frac{\dot{\phi}^2}{6M_{Pl}^2} + \frac{V(\phi)}{3M_{Pl}^2} \therefore \dot{\phi}^2 \ll V(\phi) \tag{1.73}$$

The potential term dominates over the kinetic one, which means a slow roll – compared to the potential – due to the Hubble friction term in equation (1.71)! From equation (1.73), note that $P_\phi \simeq -\rho_\phi \simeq -V(\phi) \therefore w_\phi \simeq -1$, an approximate cosmological constant as anticipated. Moreover, from the first Friedmann equation:

$$H^2 \simeq \frac{V(\phi)}{3M_{Pl}^2} \tag{1.74}$$

Now, note that, from the second slow-roll condition:

$$\eta_H = \frac{\dot{\varepsilon}_H}{H\varepsilon_H} = \frac{1}{H}\left(-\frac{H^2}{\dot{H}}\right)\left(-\frac{\ddot{H}}{H^2} + 2\frac{\dot{H}^2}{H^3}\right) = \frac{\ddot{H}}{H\dot{H}} + 2\varepsilon_H \simeq \frac{\ddot{H}}{H\dot{H}} \ll 1\,; \tag{1.75}$$

where we considered $\varepsilon_H \ll 1$ in the last equality.

Plugging in the result of equation (1.72) into equation (1.75) we get:

$$\ddot{H} = -\frac{\dot{\phi}\ddot{\phi}}{M_{Pl}^2} \Rightarrow \eta_H \simeq \left(-\frac{\dot{\phi}\ddot{\phi}}{M_{Pl}^2}\right)\left(-\frac{2M_{Pl}^2}{\dot{\phi}^2}\right)\frac{1}{H} \ll 1 \therefore \frac{\ddot{\phi}}{\dot{\phi}}\frac{1}{H} \ll 1 \tag{1.76}$$

Therefore, the above result tells us the Klein-Gordon equation may be approximated by $3H\dot{\phi}+V'(\phi) \simeq 0$, which leads to $\dot{\phi} \simeq -V'(\phi)/3H \therefore \dot{\phi}^2 \simeq [V'(\phi)]^2/9H^2 \simeq [V'(\phi)]^2M_{Pl}^2/3V(\phi)$.

This was the relation we were looking for in order to be able to rewrite the first slow-roll parameter in terms of $V(\phi)$ as:

$$\varepsilon_H \equiv -\frac{\dot{H}}{H^2} \simeq \frac{\dot{\phi}^2}{2M_{Pl}^2}\frac{3M_{Pl}^2}{V(\phi)} \simeq \frac{3}{2V(\phi)}\frac{[V'(\phi)]^2 M_{Pl}^2}{3V(\phi)} = \frac{M_{Pl}^2}{2}\left[\frac{V'(\phi)}{V(\phi)}\right]^2 \tag{1.77}$$

We can then define a new slow-roll parameter in terms of $V(\phi)$ and its slow-roll condition, which is equivalent to $\varepsilon_H \ll 1$, as:

$$\varepsilon \equiv \frac{M_{Pl}^2}{2}\left[\frac{V'(\phi)}{V(\phi)}\right]^2 \ll 1 \tag{1.78}$$

It would be nice to also redefine η_H in terms of $V(\phi)$. From equations (1.74) and (1.77) plus the condition $\varepsilon_H \simeq \varepsilon \ll 1$ and the approximate Klein-Gordon equation already derived $\dot{\phi} \simeq -V'(\phi)/3H$, after a bit of algebra we realize that:

$$\begin{aligned}
\eta_H &= \frac{\dot{\varepsilon}_H}{H\varepsilon_H} \simeq \frac{\dot{\varepsilon}}{H\varepsilon} = M_{Pl}^2\left(\frac{V'}{V}\right)\dot{\phi}\left[\frac{V''}{V} - \left(\frac{V'}{V}\right)^2\right]\frac{1}{H}\frac{2}{M_{Pl}^2}\left(\frac{V}{V'}\right)^2 \\
&\simeq \frac{2}{H}\left(\frac{V}{V'}\right)\left(-\frac{V'}{3H}\right)\left[\frac{V''}{V} - \left(\frac{V'}{V}\right)^2\right] \simeq -\frac{2V}{3}\left(\frac{3M_{Pl}^2}{V}\right)\left[\frac{V''}{V} - \left(\frac{V'}{V}\right)^2\right] \\
&\simeq -2M_{Pl}^2\frac{V''}{V} + 4\varepsilon \simeq -2M_{Pl}^2\frac{V''(\phi)}{V(\phi)}
\end{aligned} \tag{1.79}$$

Therefore, from equation (1.79) it is clear we can define η in terms of the potential such that, from $|\eta_H| \ll 1$, we get:

$$\eta \equiv M_{Pl}^2\frac{V''(\phi)}{V(\phi)} \text{ ; with slow-roll condition: } |\eta| \ll 1 \tag{1.80}$$

The importance of redefining ε_H and η_H as ε and η depending of the potential lies in the fact that, from a $V(\phi)$ of any proposed model, we can check directly if this is an appropriate model to describe slow-roll inflation and in which regions of ϕ the potential obey the slow-roll conditions.

At last, it is important to connect the previously analyzed cosmological epochs with the inflationary one. The transition period from inflation to the pre-inflationary scenario is called reheating. As was said, at first $V(\phi)$ was thought to necessarily be composed of a plateau followed by a drop. After falling off the cliff, the potential would oscillate around a minimum where the inflaton would decay into the modes of fields we observe nowadays, starting the standard pre-inflationary eras. As mentioned, this is crucial to solve the relic problem.

1.4.2 CMBR Data: Connection to Observations

So far we have been considering a perfectly homogeneous and isotropic universe. Although this is a good approximation, some inhomogeneities of order 10^{-5} of its average value are actually measured in the temperature of the CMBR. As it can be seen, this is consistent with our first approximation of an homogeneous and isotropic universe since fluctuations are orders of magnitude smaller than the average. However, in order to form the large scale structures we observe on telescopes, those fluctuations are essential. In the usual terminology, we have been studying the background dynamics of the universe so far. Now, perturbations to the background will be considered.

Perturbations both in the metric, affecting the left-hand side of the Einstein equations (1.6), and in the energy momentum tensor, right-hand side of the equations, must be considered. At the end, a connection between observations of the CMBR spectrum and these perturbations has to be constructed. If it was intended to self-consistently connect, at any level of detail, those two ends, at least another separate chapter would be needed. Here, we will just point out the main ideas behind cosmological perturbation theory and describe which parameters CMBR experiments, more precisely the Planck satellite data [18], constrains. These ideas will be important to compare inflation against other models to solve pre-inflationary problems discussed in the last chapters of this thesis. A more detailed description can be found on books [12] (chapters 7 and 11) and [32], besides the lecture notes [33].

We should start by giving an intuitive picture of how fluctuations during inflation generate the large scale structures we observe today. First consider a perturbation $\delta\phi(t,\vec{x})$ to the inflaton background field $\bar{\phi}(t)$, which was called just $\phi(t)$ until last subsection avoiding a notation overkill. Note that the bar is now important to distinguish the background field from the total field $\phi(t,\vec{x}) = \bar{\phi}(t) + \delta\phi(t,\vec{x})$. If inflation ends at $\phi = \phi_e$ then, regions where $\delta\phi > 0$ will reach the end of inflation before regions where $\delta\phi < 0$. This will eventually lead to a perturbation $\delta\rho(t,\vec{x})$ in the background energy density of these regions, $\bar{\rho}$. This simple example illustrate why we need to consider perturbations both in the metric and in the energy-momentum tensor.

An useful procedure to make calculations easier is the decomposition of the metric and the stress-energy perturbations into scalar, vector and tensor components, the so-called SVT decomposition. This decomposition is usually done in momentum space via Fourier transform $Q_k(t,\vec{k}) = \int d^3x Q(t,\vec{x}) e^{i\vec{k}\cdot\vec{x}}$, where Q stands for any metric or stress-energy quantity, e.g. $\delta\phi$. To first order, each type of perturbation evolves independently in a SVT decomposition, thus they can be treated individually. That is the usefulness of this procedure.

Let us start by considering perturbations on the line element, or equivalently on the metric, as:

$$ds^2 = a^2(\eta)\{-(1+2A)d\eta^2 - 2B_i dx^i d\eta + (\delta_{ij} + h_{ij})dx^i dx^j\}\,; \tag{1.81}$$

where a is the scale factor as usual and A, B_i and h_{ij} are the perturbations to be SVT decomposed, in the physical space, below. Note that the conformal time η is being used instead of t.

$$B_i = \partial_i B - \hat{B}_i\,;\ \text{where: } \partial^i \hat{B}_i = 0 \tag{1.82}$$

$$h_{ij} = 2C\delta_{ij} + 2\partial_{\langle i}\partial_{j\rangle}E + 2\partial_{(i}\hat{E}_{j)} + 2\hat{E}_{ij}\,; \tag{1.83}$$

$$\text{where: } \begin{cases} \partial_{\langle i}\partial_{j\rangle}E \equiv \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E\,; \\ \partial_{(i}\hat{E}_{j)} \equiv \frac{1}{2}(\partial_i\hat{E}_j + \partial_j\hat{E}_i)\,; \\ \partial^i\hat{E}_i = 0\ ;\ \partial^i\hat{E}_{ij} = 0 \text{ and } \hat{E}^i{}_i = 0 \end{cases}$$

After the SVT decomposition, we still have 10 degrees of freedom, as expected for a 4-dimensional metric. There are four scalars (A, B, C and E) two divergenceless vectors ($\hat{B}_i$ and $\hat{E}_i$) and one traceless and divergenceless 3-dimensional symmetric tensor ($\hat{E}_{ij}$), summing up to $4 + 4 + 2 = 10$ degrees of freedom. It should be pointed out that vector perturbations quickly decay with the Hubble expansion, therefore the relevant quantities to us will be only the scalars and the tensor $\hat{E}_{ij}$.

Here comes the subtle issue that the perturbations defined above are not gauge invariant. It can be shown that a simple time coordinate transformation in the FLRW metric can lead to unphysical perturbations. Therefore we will need to search for gauge-independent quantities. Choosing the so-called Newton (or Newtonian) gauge, the line element of equation (1.81) becomes:

$$ds^2 = a^2(\eta)[-(1 + 2\Psi)d\eta^2 + (1 - 2\Phi)\delta_{ij}dx^i dx^j]\,; \tag{1.84}$$

where Ψ and Φ are two of the Bardeen variables, which are invariant under general coordinate transformations.

We may then define a gauge-independent variable from a combination of metric, Φ, and energy-momentum tensor, $\delta\rho$, perturbations:

$$\zeta \equiv -\Phi - \frac{H}{\dot{\rho}}\delta\rho \tag{1.85}$$

Again, we refer to the notes [33] appendix A to a motivation to this definition.

The power spectrum of scalar perturbations, which is going to be constrained by CMBR data, is defined from the quantity defined in equation (1.85). The goal is to link Gaussian perturbations of the CMBR to the same perturbations of the gauge-invariant scalar ζ. The variance of any quantity $q = q(\vec{x})$ can be defined as:

$$\sigma_q^2(\vec{x}) \equiv \langle q^2(\vec{x})\rangle - \langle q(\vec{x})\rangle^2 \Rightarrow \sigma_q^2(\vec{x}) = \langle q^2(\vec{x})\rangle \text{ ; for } \langle q(\vec{x})\rangle = 0\,; \tag{1.86}$$

where $\langle\rangle$ means average. In the case of the CMBR quantity to be analyzed shortly $\Theta \equiv \delta T/T$

it stands for an ensemble average, i.e. average over the sky. Note that only perturbations, i.e. quantities which by definition average to zero, will be concerning us in the following analysis, therefore $\sigma_q^2 = \langle q^2 \rangle$ will be considered.

The probability density function at $q \equiv q(\vec{x}_0)$, often introduced as the probability of measuring $q(\vec{x})$ between q and $q + dq$, is generically $\propto e^{-\frac{q^2}{2\sigma_q^2}}$, i.e. it follows a Gaussian distribution. Therefore, Gaussian perturbations are obtained by two-point functions $\langle q^2(\vec{x}) \rangle$.

The two-point correlation function of scalar perturbations in real, i.e. physical, space is:

$$\xi(r) \equiv \langle \zeta(\vec{x})\zeta(\vec{x} + \vec{r}) \rangle ; \tag{1.87}$$

where we make the assumption ξ only depends on $r \equiv |\vec{r}|$.

The two-point function in momentum space is found by taking the Fourier transform on both $\vec{x}$ and $\vec{x} + \vec{r}$.

$$\langle \zeta_k \zeta_{k'} \rangle = \int d^3x \int d^3r \xi(r) e^{i\vec{k}'\cdot\vec{x}} e^{i\vec{k}\cdot(\vec{x}+\vec{r})} = \int d^3x e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} \int d^3r \xi(r) e^{i\vec{k}\cdot\vec{r}} \tag{1.88}$$

Finally, we define the power spectrum $P_\zeta(k)$ as the Fourier transform of the two-point function $\xi(r)$:

$$P_\zeta(k) \equiv \int d^3r \xi(r) e^{i\vec{k}\cdot\vec{r}} \Rightarrow \langle \zeta_k \zeta_{k'} \rangle = \int d^3x e^{i(\vec{k}+\vec{k}')\cdot\vec{x}} P_\zeta(k) = (2\pi)^3 \delta^3(\vec{k} + \vec{k}') P_\zeta(k) ; \tag{1.89}$$

where we assumed P_ζ only depends on $k \equiv |\vec{k}|$ and after the implication arrow we plugged the definition into equation (1.88).

Moreover, its usual to define a dimensionless power spectrum of scalar perturbations, $\Delta_s^2(k)$, as:

$$\begin{aligned} \sigma_\zeta^2(\vec{x}) = \langle \zeta^2(\vec{x}) \rangle = \xi(0) &= \frac{1}{(2\pi)^3} \int d^3k P_\zeta(k) \\ &= \frac{1}{(2\pi)^3} \int dk 4\pi k^2 P_\zeta(k) \equiv \int d\log k \Delta_s^2(k) = \int \frac{dk}{k} \Delta_s^2(k) ; \end{aligned} \tag{1.90}$$

such that $\Delta_s^2(k) \equiv \frac{k^3}{2\pi^2} P_\zeta(k)$.

The parameter which quantifies the scale-dependence of Δ_s^2 is called scalar spectral index, or equivalently scalar tilt, n_s. Furthermore, there is another parameter which tells us how the tilt varies with k and is called the running of the spectral index, α_s. They are defined as:

$$n_s - 1 \equiv \frac{d\log\Delta_s^2}{d\log k} \text{ ; and } \alpha_s \equiv \frac{dn_s}{d\log k} \tag{1.91}$$

Solving the differential equations of equation (1.91) assuming a constant α_s gives us:

$$\Delta_s^2(k) = \Delta_s^2(k_*)\left(\frac{k}{k_*}\right)^{n_s(k_*)-1+(1/2)\alpha_s\log(k/k_*)} ; \tag{1.92}$$

where k_* is a pivot scale, i.e. an arbitrary scale of reference.

The connection between CMBR spectrum and the scalar perturbation power spectrum is done by the transfer function $T_l(k)$. The anisotropy $\Theta \equiv \delta T/T$ is what is measured by CMBR observations (actually we observe intensities of waves which are proportional to Θ). This quantity depends not only on spacetime coordinates η, the conformal time of today, and $\vec{x}$, the position of the Earth or the detector, but also on the angular direction from which the photons come to us from the sky $\vec{e}$. Therefore, one may perform a multipole expansion of Θ:

$$\Theta(\eta, \vec{x}, \vec{e}) = \sum_{lm} \Theta_{lm}(\eta, \vec{x}) Y_{lm}(\vec{e}) ; \tag{1.93}$$

in order to isolate the $\vec{e}$-dependence. All in all, the precise previously mentioned connection is done by $\Theta_l(k) = T_l(k)\zeta_k$. Therefore, if we know the transfer function and the scalar perturbations power spectrum, the power spectrum of Θ_l can be obtained.

The same argument we made to define Δ_s^2 can be made to analogously define a power spectrum of tensor perturbations. These perturbations were not yet observed in the CMBR, however one may expect future observations to measure them. This detection would be relevant since it would tell us the scale in which inflation happened, which nowadays can only be estimated. Tensor perturbations arise from the tensor h_{ij} in equation (1.81). They are described by two propagating degrees of freedom, i.e. the graviton polarizations, h_+ and

$h_\times$. The power spectra (dimensional and dimensionless) of tensor perturbations are:

$$\langle h_k h_{k'} \rangle = (2\pi)^3 \delta^3(\vec{k} + \vec{k}') P_h(k) \text{ ; and } \Delta_h^2 \equiv \frac{k^3}{2\pi^2} P_h(k) \, ; \tag{1.94}$$

where h stands for any of the two polarizations. The total tensor power spectrum is defined to be the sum of the power spectrum of each degree of freedom. Assuming the power spectra of h_+ and $h_\times$ are equal, we end up with $\Delta_t^2(k) \equiv \Delta_{h_+}^2(k) + \Delta_{h_\times}^2(k) = 2\Delta_h^2(k)$.

The tensor tilt is defined analogously to its scalar counterpart but, for historical reasons, without the -1. If we consider it as a constant, then we get:

$$n_t \equiv \frac{d \log \Delta_t^2}{d \log k} \Rightarrow \Delta_t^2(k) = \Delta_t^2(k_*) \left(\frac{k}{k_*} \right)^{n_t} \; ; \tag{1.95}$$

where, again, k_* is a pivot scale.

Another useful quantity is the ratio of tensor perturbations to the scalar ones, which can be defined as $r \equiv \Delta_t^2 / \Delta_s^2$.

Although it will not be derived here, we could calculate expressions for the power spectra of tensor and scalar perturbations from a general slow-roll inflationary model. With those expression in hand, we could show the already observed scalar perturbations cannot constrain directly the scale at which inflation happened, only its ratio to the slow-roll parameter ε. On the other hand, tensor perturbations, Δ_t^2, would tell us the specific scale of inflation. Furthermore, the Planck experiment [18] gives us the best fit for the scalar tilt and a constraint to the ratio of tensor to scalar perturbations:

$$n_s = 0.965 \pm 0.004 \text{ ; and } r < 0.06 \, ; \tag{1.96}$$

which means observations are well-explained by a constant n_s leading to $\alpha_s = 0$. Moreover, since $n_s \simeq 1$, the power spectrum of scalar perturbations is approximately constant and has a value $\sim 10^{-8}$ or 10^{-9}.

To conclude, in this section perturbations to the background metric and energy momentum tensor were considered. They are essential to the formation of the large scale structure observed nowadays, e.g. galaxies. During inflation, perturbations are generated by quantum fluctuations – not shown here – then scales go outside the horizon at the moment when $k \simeq aH$ (since k is the wave number mode conjugated to a comoving scale x leading to a physical momentum k/a) where they got frozen in. Later, at radiation or matter-dominated eras, those scales come back inside the horizon and we see the effects of those perturbations on them. More specifically, their effects are analyzed from CMBR data. We reinforce that this section is only a brief contextualization of an enormous subject. To more details one must see the references cited in the beginning of the section. The content exposed here contains enough information to contextualize the general features of cosmological perturbation theory that will be important in the last chapters of the thesis.

1.5 Summary of Ideas

Nowadays, we have a standard model for cosmology which is able to tell us the history of the universe from a few seconds, or even less than that, after the cosmological singularity up until now, the so-called ΛCDM model. The name is given due to the two most dominant types of energy content in our universe today: Λ stands for the cosmological constant and CDM means cold dark matter. Dark matter is believed to be cold in the sense that its fluid is believed to have low temperature compared to the CMBR's one. This is the scenario most favored by experiments.

It is an important remark the fact that CMBR data analyzed in [18] is able to give us, with good precision, the value of a bunch of observables based on the fitting of the six parameters of the model. The precision of cosmological observations nowadays contrasts with the lack of it of a few decades ago. Cosmology is said to have entered the precision era since now we have, for instance, the age of the universe as precise as 13.801 ± 0.024 billion years – value extracted from table 1 of [18] – while about 30 years ago our uncertainty bars where in the exponent of the age of the universe's order of magnitude. The six free parameters will not all be discussed here but among them there are the ratio of baryon energy density to the critical one multiplied by a factor of $h_0 \equiv \frac{H_0\,[(\mathrm{km/s})/\mathrm{Mpc}]}{100}$ squared, its analog to cold dark matter, the scalar spectral tilt and the amplitude of scalar perturbations at some fixed scale.

The obtained constraint of the energy density of curvature perturbations, actually $\Omega_k^{(0)}$, coming from baryon acoustic oscillations observations – another important type of experiment not discussed here – plus the CMBR data is $\Omega_k^{(0)} = 0.001 \pm 0.002$. This justifies the neglecting of the curvature energy density today done this whole chapter. Furthermore, this together with the known values of Ωs today also justify the fact that this type of energy is negligible at any time in our cosmological history. Note that $\rho_k \equiv -\frac{3k}{8\pi G a^2} \propto a^{-2}$, it decreases going to the future, and since $\Omega_k^{(0)} \ll \Omega_\Lambda^{(0)} \therefore \rho_k^{(0)} \ll \rho_\Lambda^{(0)}$ plus the fact that ρ_Λ is constant it is clear that the curvature energy density is not relevant in the future. Going back in time it is true that ρ_k, at some point, could have overcome ρ_Λ, however ρ_m – energy density of total matter, i.e. dark matter plus the baryonic one – increases more rapidly going to the past than the curvature energy density and since $\Omega_k^{(0)} \ll \Omega_m^{(0)} \therefore \rho_k^{(0)} \ll \rho_m^{(0)}$ at least matter energy density will always dominate over the curvature one allowing us to keep neglecting the last. We didn't even mention the radiation energy density that increases even faster than ρ_m going to the past and is known to have dominated all other types of energy content a long time ago. All in all, it is justified a posteriori the neglecting of ρ_k at all times.

The values of some relevant physical quantities found in [18] are summarized below:

$$H_0 \simeq 67.4\,(\mathrm{km/s})/\mathrm{Mpc}\,;\; \Omega_\Lambda^{(0)} \simeq 0.69\,;\; \Omega_{dm}^{(0)} \simeq 0.26\,;\; \Omega_b^{(0)} \simeq 0.05\,;\; n_s \simeq 0.97\,; \tag{1.97}$$

where it is obtained from $\Omega^{(0)} \equiv \sum_i \Omega_i^{(0)} \simeq 1$ and $\Omega_k^{(0)} \simeq 0$ that the radiation energy density is also negligible today. Nevertheless, $\rho_r \propto a^{-4}$ increases rapidly going into the past so that we know, from the thermal history of the universe, that radiation dominated the universe energy density some time ago and we are able to estimate that nowadays $\Omega_r^{(0)} \sim 10^{-5}$.

Figure 1.1 shows clearly the fact that ρ_k is negligible and also illustrates the flatness problem, since in the past the difference between ρ_k and the sum of the other types of energy was even bigger. This graph plus the derived expressions for $a = a(t)$ summarize the picture we have of the universe in pre-inflationary cosmology. When we add inflation – even though, as we will see, there are alternative scenarios to that of inflation – to the picture, we realize our universe started off very small and passed through a period of exponential expansion. Then there was a long period of radiation domination, starting very hot, transitioning to a period of matter domination later and just recently we entered a period of cosmological constant dominance, better said dark energy dominance (see chapter 2). Many details on the thermal history of the universe were left out in order not to make this chapter longer than it already is but they can be found on books on cosmology, e.g. [12] and [17].

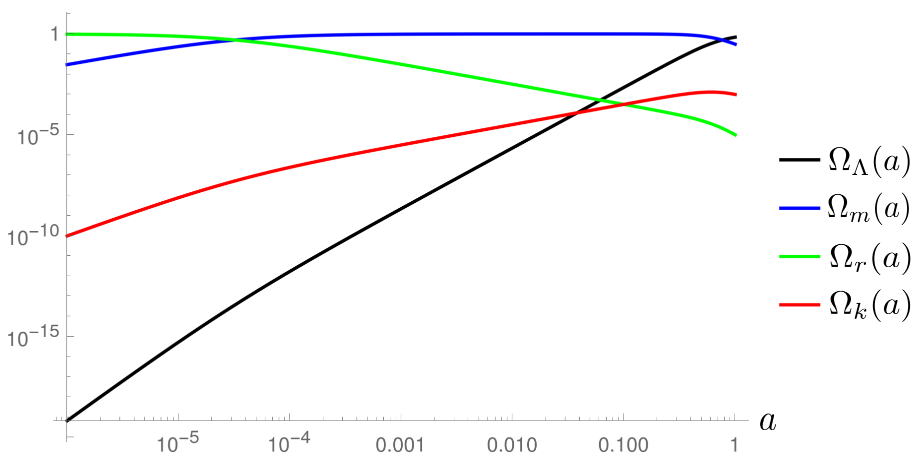

Figure 1.1: Plot of the energy densities, more precisely their ratios to the critical density, as the scale factor evolves until today. The horizontal axis is represented in log scale to highlight the early radiation-dominated era, while the y-axis is also in log scale to illustrate the flatness problem. The following current values were used to generate the plot: $\Omega_\Lambda^{(0)} = 0.69$, $\Omega_m^{(0)} \equiv \Omega_{dm}^{(0)} + \Omega_b^{(0)} = 0.31$, $\Omega_r^{(0)} = 10^{-5}$ and $\Omega_k^{(0)} = 10^{-3}$.

Finally, at the end of the chapter, slow-roll inflation was introduced as the most popular incarnation of inflationary models. It was shown that, from the slow-roll conditions $\varepsilon \ll 1$ and $|\eta| \ll 1$, one can constrain scalar field models of inflation. Furthermore, the importance of scalar and, perhaps, tensor perturbations to the formation of galaxies was explained. It was introduced the concept of power spectrum of perturbations, from which one can test or constrain inflationary models from CMBR observations.

Chapter 2

The Cosmological Constant Problem

This chapter contains the main problem which motivated the project. It could easily have been explained in the middle of chapter 1, but its relevance deserved a separate chapter. We shall not make it too long though.

The cosmological constant problem is a famous disagreement between the observed value of cosmological constant energy density, more precisely $\Omega_{\Lambda}^{(0)} \simeq 0.69$, as already discussed and the energy density coming from vacuum energy. Recall that, from quantum mechanics, the vacuum is not an empty and stationary state as it was thought to be within classical physics. It is actually a rather dynamical state in which particles can pop up out of pure energy for a small period of time and then come back to annihilate among themselves. Furthermore, the important point is that one can attribute a non-zero energy to this vacuum state. In usual quantum field theory, in Minkowski spacetime, one gets rid of the influence of this vacuum energy since only difference in energies are measurable. However, when gravity is taken into consideration the scenario change. From the Einstein field equations (1.6), any contribution to the total stress-energy tensor, i.e. sum of all components, is expected to influence the left-hand gravitational side of the equation. In other words, energy gravitates and this, a priori, should be valid for any type of energy content, including vacuum energy.

It must be shown that such an energy content coming from the vacuum indeed has equation of state $P_{vac} = -\rho_{vac}$ so that it can be identified as a cosmological constant contribution. This is done here by performing a usual GR trick: One starts by analyzing a problem in flat spacetime and then writes its covariant counterpart in curved spacetime from it. Notice that the vacuum should be invariant, i.e. all observers should see the same vaccum, and the only invariant (0,2)-tensor field in Minkowski spacetime is the metric $\eta_{\mu\nu}$. Therefore the energy-momentum tensor of the vacuum, more precisely its vacuum expectation value (VEV), should be proportional to $\eta_{\mu\nu}$. The generalization of this statement to a curved background is $\langle 0|T^{vac}_{\mu\nu}|0\rangle \equiv \left\langle T^{vac}_{\mu\nu}\right\rangle_0 \propto g_{\mu\nu}$ leading to $\left\langle T^{vac}_{\mu\nu}\right\rangle_0 = f(x)g_{\mu\nu}$, where f is a function of spacetime coordinates x. Since the stress-energy tensor should be conserved for every type of matter, $\nabla_\mu T^\mu{}_\nu = 0$, as it was already assumed in the derivation of the continuity equation (1.23), $f(x) = k$ (constant). Moreover, since for any classical type of energy content it is true that $\rho = u^\mu u^\nu T_{\mu\nu}$, where u^μ is the 4-velocity of the observer, then:

$$\begin{aligned} \langle 0|T^{vac}_{\mu\nu}|0\rangle = k g_{\mu\nu} \therefore \langle 0|T^{vac}_{\mu\nu}|0\rangle\, u^\mu u^\nu = -k = \rho_{vac} \Rightarrow \\ \langle 0|T^{vac}_{\mu\nu}|0\rangle \equiv \left\langle T^{vac}_{\mu\nu}\right\rangle_0 = -\rho_{vac} g_{\mu\nu}\,; \end{aligned} \tag{2.1}$$

where in the first line, after the $\therefore$ symbol, $u^\mu u^\nu$ was multiplied on both sides of the equation and the fact $u^\mu u^\nu g_{\mu\nu} = -1$ was used.

Since the equation of state of this vacuum energy density is precisely the same as the one for the cosmological constant, a natural first guess is to try and explain the observed cosmological constant energy density as due uniquely to zero-point energy of Standard Model fields, i.e. vacuum energy. However, as we will see, this generates a huge disagreement which goes by the name of vacuum catastrophe.

2.1 Vacuum Catastrophe

There are two variations of the of the cosmological constant problem and different ways to analyze them, see [34] for a comprehensive review. What is going to be described right now is the zero-point energy of a simplified flat (Minkowski) universe where there is just a real scalar field, for the sake of simplicity, and the vacuum energy density generated by it. After considering this simplification, the generalization to a more realistic case will be discussed. This zero-point energy is usually suppressed by the normal ordering procedure, with the valid excuse, as already discussed, that one can only measure energy differences. We start from the scalar field expansion into creation and annihilation operators, see [35]:

$$\phi(t,\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2\omega_p}}[a(t,\vec{p}) + a^\dagger(t,-\vec{p})]e^{-i\vec{p}\cdot\vec{x}}\,; \tag{2.2}$$

$$\pi(t,\vec{x}) = \dot{\phi}(t,\vec{x}) = \int \frac{d^3p}{(2\pi)^3}\left(-i\sqrt{\frac{\omega_p}{2}}\right)[a(t,\vec{p}) - a^\dagger(t,-\vec{p})]e^{-i\vec{p}\cdot\vec{x}}\,; \tag{2.3}$$

where $\omega_p^2 = |\vec{p}|^2 + m^2$ and m is the mass of the field. The different minus signs in the exponentials in comparison with book [35] is due to a different convention in the definition of Fourier transform.

The energy momentum tensor for a general scalar field with potential $V(\phi)$ was calculated in equation (1.66). For a free field theory in Minkowski spacetime, with $V(\phi) = m^2\phi^2/2$, which is of our interest since there is no other field to interact with ϕ and other self-interactions have been neglected for simplicity, we get:

$$T_{00} = \frac{\dot{\phi}^2}{2} + \frac{1}{2}\delta^{ij}\partial_i\phi\partial_j\phi + \frac{m^2}{2}\phi^2 \tag{2.4}$$

Note that if we wisely choose an observer such that $u^\mu = (1,0,0,0)$ then $\rho_{vac} = \langle T_{\mu\nu}u^\mu u^\nu\rangle_0 = \langle T_{00}(u^0)^2\rangle_0 = \langle 0|T_{00}|0\rangle$. So everything we need is the VEV of the expression for T_{00} in equation (2.4). Let us define $a_p \equiv a(t,\vec{p})$ and its analog $a_p^\dagger \equiv a^\dagger(t,\vec{p})$ to simplify notation. From equations (2.2) and (2.3) plus the fact that from $[a_p, a_k^\dagger] = (2\pi)^3\delta^3(\vec{p}-\vec{k})I$, where I is the identity operator, and $a_p|0\rangle = 0 \Leftrightarrow \langle 0|a_p^\dagger = 0$ we have $\langle [a_p, a_k^\dagger]\rangle_0 = \langle a_p a_k^\dagger\rangle_0 = (2\pi)^3\delta^3(\vec{p}-\vec{k})$, $\langle T_{00}\rangle_0$ can be calculated term by term as:

(i). Let us start with $\langle\dot{\phi}^2\rangle_0 = \langle\pi^2\rangle_0$:

$$
\begin{aligned}
\langle\dot{\phi}^2\rangle_0 &= -\int\frac{d^3p}{(2\pi)^3}\int\frac{d^3p'}{(2\pi)^3}\frac{\sqrt{\omega_p\omega_{p'}}}{2}\left\langle(a_p a_{p'} + a_{-p}^\dagger a_{-p'}^\dagger - a_p a_{-p'}^\dagger - a_{-p}^\dagger a_{p'})\right\rangle_0 e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}} \\
&= \int\frac{d^3p}{(2\pi)^3}\int\frac{d^3p'}{(2\pi)^3}\frac{\sqrt{\omega_p\omega_{p'}}}{2}\left\langle a_p a_{-p'}^\dagger\right\rangle_0 e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}} \\
&= \int\frac{d^3p}{(2\pi)^3}\int d^3p'\frac{\sqrt{\omega_p\omega_{p'}}}{2}\delta^3(\vec{p}+\vec{p}')e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}} \\
&= \int\frac{d^3p}{(2\pi)^3}\frac{\omega_p}{2}
\end{aligned}
\tag{2.5}
$$

(ii). Now, $\langle\delta^{ij}\partial_i\phi\partial_j\phi\rangle_0$:

$$
\begin{aligned}
\left\langle\delta^{ij}\partial_i\phi\partial_j\phi\right\rangle_0 &= \iint\frac{d^3p}{(2\pi)^3}\frac{d^3p'}{(2\pi)^3}\frac{-\vec{p}\cdot\vec{p}'}{2\sqrt{\omega_p\omega_{p'}}}e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}}\left\langle a_p a_p' + a_p a_{-p'}^\dagger + a_{-p}^\dagger a_{p'} + a_{-p}^\dagger a_{-p'}^\dagger\right\rangle_0 \\
&= \iint\frac{d^3p}{(2\pi)^3}d^3p'\frac{-\vec{p}\cdot\vec{p}'}{2\sqrt{\omega_p\omega_{p'}}}e^{-i(\vec{p}+\vec{p}')\cdot\vec{x}}\delta^3(\vec{p}+\vec{p}') \\
&= \int\frac{d^3p}{(2\pi)^3}\frac{|\vec{p}|^2}{2\omega_p}
\end{aligned}
\tag{2.6}
$$

(iii). Finally $\langle\phi^2\rangle_0$:

$$
\begin{aligned}
\left\langle\phi^2\right\rangle_0 &= \iint\frac{d^3p}{(2\pi)^3}\frac{d^3p'}{(2\pi)^3}\frac{1}{2\sqrt{\omega_p\omega_{p'}}}e^{-i(\vec{p}+\vec{p'})\cdot\vec{x}}\left\langle a_p a_p' + a_p a_{-p'}^\dagger + a_{-p}^\dagger a_{p'} + a_{-p}^\dagger a_{-p'}^\dagger\right\rangle_0 \\
&= \iint\frac{d^3p}{(2\pi)^3}d^3p'\frac{1}{2\sqrt{\omega_p\omega_{p'}}}e^{-i(\vec{p}+\vec{p'})\cdot\vec{x}}\delta^3(\vec{p}+\vec{p}') \\
&= \int\frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_p}
\end{aligned}
\tag{2.7}
$$

Thus, from equations (2.5), (2.6) and (2.7):

$$\rho_{vac} = \langle 0|T_{00}|0\rangle = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{\omega_p^2 + |\vec{p}|^2 + m^2}{2\omega_p} = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\omega_p(|\vec{p}|)$$
$$= \frac{1}{4\pi^2}\int_0^\infty dp p^2\sqrt{p^2+m^2}\,; \tag{2.8}$$

where in the last equality it was used the notation $|\vec{p}| \equiv p$.

It is clear, from equation (2.8), that the value of ρ_{vac} diverges. Since a divergent type of energy is not measured (how would we even measure something infinite?), this fact seems to be a problem a priori. However, we could regularize the integral. If the integral in momentum is cut off at an energy (or equivalently mass) scale M it is still nicely interpreted as an effective field theory valid up to this scale only. This scale can be the Planck one, where the Standard Model is certainly expected to be replaced by a theory which includes gravity like string theory, or the scale of other phase transitions that may have happened before Planck scale is reached going into the past, e.g. supersymmetry breaking scale or string scale - which can be lower than M_{Pl}. Then, for a general cut-off M:

$$\begin{aligned}
\rho_{vac} &= \frac{1}{4\pi^2}\int_0^M dp p^2\sqrt{p^2+m^2} = \frac{1}{4\pi^2}\left[\frac{M}{4}(M^2+m^2)^{3/2} - \frac{m^2}{4}\int_0^M dp\sqrt{p^2+m^2}\right]\\
&= \frac{1}{4\pi^2}\left\{\frac{M}{4}(M^2+m^2)^{3/2} - \frac{m^2}{4}\left[\frac{M}{2}(M^2+m^2)^{1/2} + \frac{m^2}{2}\int_0^M \frac{dp}{\sqrt{p^2+m^2}}\right]\right\}\\
&= \frac{1}{4\pi^2}\left[\frac{M^4}{4}\left(1+\frac{m^2}{2M^2}\right)\left(1+\frac{m^2}{M^2}\right)^{1/2} - \frac{m^4}{8}\int_0^{M/m}\frac{dx}{\sqrt{x^2+1}}\right]\\
&= \frac{1}{4\pi^2}\left[\frac{M^4}{4}\left(1+\frac{m^2}{2M^2}\right)\left(1+\frac{m^2}{M^2}\right)^{1/2} - \frac{m^4}{8}\log\left|x+\sqrt{x^2+1}\right|\Big|_0^{M/m}\right]\\
&= \frac{1}{4\pi^2}\left[\frac{M^4}{4}\left(1+\frac{m^2}{2M^2}\right)\left(1+\frac{m^2}{M^2}\right)^{1/2} - \frac{m^4}{8}\log\left(\frac{M}{m}+\frac{M}{m}\sqrt{1+\frac{m^2}{M^2}}\right)\right]\\
&= \frac{M^4}{16\pi^2}\left[1+\mathcal{O}\left(\frac{m^2}{M^2}\right)\right]\,;
\end{aligned} \tag{2.9}$$

where it was considered $M \gg m$ in the last equality since this breaking scale should be much higher than any particle mass in the Standard Model. Note that in the second and third equalities above the following partial integrations were used:

(i). In the second equality:

$$\begin{aligned}\int_0^M dpp^2\sqrt{p^2+m^2} &= \left[\frac{p}{3}(p^2+m^2)^{3/2}\right]\Big|_0^M - \frac{1}{3}\int_0^M dp(p^2+m^2)^{3/2} \\ &= \frac{M}{3}(M^2+m^2)^{3/2} - \frac{1}{3}\left(\int_0^M dpp^2\sqrt{p^2+m^2} + m^2\int_0^M dp\sqrt{p^2+m^2}\right)\end{aligned}$$

$$\Rightarrow \int_0^M dpp^2\sqrt{p^2+m^2} = \frac{M}{4}(M^2+m^2)^{3/2} - \frac{m^2}{4}\int_0^M dp\sqrt{p^2+m^2} \tag{2.10}$$

(ii). And in the third:

$$\begin{aligned}\int_0^M dp\sqrt{p^2+m^2} &= [p\sqrt{p^2+m^2}]\big|_0^M - \int_0^M dp\frac{p^2}{\sqrt{p^2+m^2}} \\ &= M(M^2+m^2)^{1/2} - \int_0^M \sqrt{p^2+m^2} + m^2\int_0^M \frac{dp}{\sqrt{p^2+m^2}}\end{aligned}$$

$$\Rightarrow \int_0^M dp\sqrt{p^2+m^2} = \frac{M}{2}(M^2+m^2)^{1/2} + \frac{m^2}{2}\int_0^M \frac{dp}{\sqrt{p^2+m^2}} \tag{2.11}$$

Back to equation (2.9), note that if it is considered a cut-off about M_{Pl} we have:

$$\rho_{vac} \sim M_{Pl}^4 \simeq (2\times 10^{27})^4\,\mathrm{eV}^4 \sim 10^{109}\,\mathrm{eV}^4 \tag{2.12}$$

This value must be compared against the observed one. From CMBR observations of equation (1.97) we have $H_0 \simeq 67.4\,(\mathrm{km/s})/\mathrm{Mpc} \simeq 70\times3\times10^{-20}\,\mathrm{s}^{-1} \simeq 70\times3\times7\times10^{-36}\,\mathrm{eV} \sim 10^{-33}\,\mathrm{eV}$ and $\Omega_\Lambda^{(0)} \equiv \rho_\Lambda^{(0)}/\rho_c^{(0)} = \rho_\Lambda^{(0)}/(3H_0^2M_{Pl}^2) \simeq 0.69$, thus:

$$\rho_\Lambda^{(0)} \sim 3\times10^{-66}\times(2\times10^{27})^2\,\mathrm{eV}^4 \sim 10^{-11}\,\mathrm{eV}^4 \tag{2.13}$$

From the results of equations (2.12) and (2.13), if the observed cosmological constant value was generated just by the effect of this predicted vacuum energy, then we would reach a theory-observation disagreement of:

$$\frac{\rho_{vac}}{\rho_\Lambda^{(0)}} \sim 10^{120}\,; \tag{2.14}$$

which is usually referred to as the largest disagreement between theory and experiment ever.

Of course, this calculation was simplified in several ways. First, we chose to analyze the simplest type of field without interactions. A more precise calculation would consider every Standard Model field and then sum up their vacuum energies. The scalar field calculation gives us a clear statement of the problem that is also present in more realistic models though. The second simplification, or rather guess, was the cut-off scale, which could be fairly lower than $\sim M_{Pl}$. Nevertheless, any expected phase transition at higher energies would already happen at too high of an energy scale to yield the observed value of $\rho_\Lambda^{(0)}$. One last problem arises with our choice of regularization procedure. Lorentz symmetry is broken when a cut-off in momentum is imposed and this leads to a few problems. For instance, even massless fields contribute to this vacuum energy from equation (2.9). On the other hand, if we have performed a dimensional regularization, we would see that massless fields, e.g. the electromagnetic field, would not contribute to the vacuum energy density, see [34]. This choice of regularization scheme diminishes the discrepancy between theoretical prediction and observed value but it is not able to solve the problem completely.

If one assumes that vacuum energy gravitates, i.e. that it indeed contributes to the right-hand side of the Einstein equations (1.6), and that it is given by the above calculation – or by its counterpart using a dimensional regularization – then, since the observed cosmological constant is much lower than its theoretically expected value, the simplest way to solve the problem is to consider the observed cosmological constant is formed by, at least, two contributions: vacuum energy and something else needed to compensate for the mismatch in the vacuum energy density and the observed cosmological constant energy density. The simplest and least theoretically satisfying possibility for this “something else” to be is to consider a constant of nature, i.e. one that we cannot predict, only figure out from experiments. The problem is, to compensate for the huge vacuum energy density, the value of such constant would have to be negative and extremely large in absolute value to yield the correct observed net value of $\rho_{\Lambda}^{(0)}$. That would be an enormous fine-tuning, indicating something is wrong (or at least incomplete) with our theoretical model. Perhaps, we are pushing the limits of conventional QFT too far outside its validity region. A quantum gravity theory, like string theory, insight may be needed to solve the problem without resorting to such a fine-tuning.

Equation (2.14) represents the problem which is also called vacuum catastrophe. In some sense, it is the same problem as the cosmological constant one. However, the cosmological constant problem can be generalized to the dark energy problem which is broader than the vacuum catastrophe. One might try to make a point that vacuum energy does not gravitate like any other type of energy in order to get rid of the huge disagreement of equation (2.14) or even find some mechanism to make the overall vacuum energy vanish, as supersymmetry would do if it was not, at least, broken. However, the cosmological constant (or dark energy, as we will see) energy density cannot be set to zero exactly. That is why we may look for a model which still gives us an accelerated universe today, as required by observations, but can be dynamical with $w \to -1$ not necessarily $w = -1$, as in the cosmological constant case. In practice, we would be ignoring the vacuum catastrophe or, more precisely, considering this problem may be solved in the future within a stronger theoretical framework such that $\rho_{vac} \to 0$ and we look for a model to create a dark energy density that matches the observations.

2.2 Dark Energy

As it was stated in the last paragraph, what researchers are more confident of is that the universe is now not only expanding but doing so at an accelerated pace. There is no need for a cosmological constant to explain this acceleration a priori. From the second Friedmann equation (1.22), in order to get $\ddot{a} > 0$ we only need the dominant energy content in the universe today to have equation of state $\rho + 3P < 0 \therefore w < -1/3$ (not necessarily $w = -1$). It is important to keep in mind though that Planck satellite results [18] heavily favour a cosmological constant $w = -1$. Anyway, it feasible to construct models which, under certain conditions, give us $w \simeq -1$. In these models, dark energy density is not constant as in the case of the cosmological constant, thus these are dynamic models.

In summary, an energy component whose equation of state yields at least $w \simeq -1$, which leads to an almost exponential increase in the scale factor, is needed to explain the current configuration of the universe. We have already seen a model which has this characteristic: the scalar field model for slow-roll inflation. Recall that the first slow-roll condition $\varepsilon_H \ll 1$ leads to $\dot{\phi}^2 \ll V(\phi)$ which gives $w_\phi \simeq -1$. Therefore, dark energy is not only a relevant type of energy nowadays but it was also important in the very early universe.

Inspired by scalar field models for inflation, one could come up with an analogous model for dark energy today. This is indeed a usual practice. Although quintessence was originally defined as any type of dynamic matter – even with a possible time dependence in the equation of state parameter, $w = w(t)$ – which would originate the accelerated expansion of the universe [36], replacing the cosmological constant, the term is widely associated today with scalar field models for dark energy. Such a model would have to obey a condition similar to the first slow-roll one which enforces the potential term dominates over the kinetic one to obtain $w \simeq -1$, matching observations. The similarity between scalar field models for dark energy and inflation, both represent periods of nearly exponential expansion of the universe, even motivates the search for models in which a unique scalar field does both jobs, e.g. see [37]. It must be said that scalar field models for dark energy are not a requirement. In principle, one could come up with a vector field model, for example. The scalar field is just the simplest option and it seems promising. The same goes for inflationary models, as discussed in chapter 1.

One last comment is in order. The existence of a light scalar field may violate some GR experiments. The scalar field would represent a fifth force for which we have found no experimental evidence yet. Therefore, it is not that simple to postulate the existence of a scalar field for dark energy. We will come back to this issue especially when chameleon cosmology is introduced, see chapter 6.

Chapter 3

Relevant Aspects of String Theory

This chapter is actually a needed digression from our main cosmological discussion. Here, some relevant concepts from string theory will be discussed in order to give us the right tools to properly introduce the models in the last three chapters of this thesis. Only the most relevant information needed to understand the last chapters will be highlighted. Therefore, this is far from a complete and detailed review of string theory, which by itself is a much broader subject than one would hope to describe in a master thesis. More details can be found on book [12] and on the lecture notes [38].

Before introducing the dynamics of strings and how they generate an entire spectrum of particles, two recurrent subjects in string theory must be issued: higher dimensions and supersymmetry (SUSY), more precisely its local version called supergravity (SUGRA).

3.1 Higher Dimensions

We start by describing higher dimensional theories. The idea of considering a world with more than four dimensions – three spatial plus one time dimension, often called $3+1$ dimensions – dates at least back to the 1920s through the works of Theodor Kaluza [39] and Oskar Klein [40]. Their papers and subsequent work by others were of interest of physicists in different epochs within the 20th century, see [41] for a historical overview (and technical review). Theories which consider very small and compact extra dimensions, so that one can only move through $3+1$ dimensions and the rest is unobservable, are called Kaluza-Klein theories. Other scenarios in which the extra dimensions don't need to be too small, i.e. comparable to the Planck length, was also proposed by physicists in the 1990s, see for example [42, 43, 44, 45]. However, the relevant scenario for this thesis will be the Kaluza-Klein one.

3.1.1 Kaluza-Klein Theories

As anticipated, Kaluza-Klein theories are higher dimensional theories which contain small, compactified extra dimensions. The reason why these theories are so appealing is because they represent a simple way to unify fundamental forces of Nature. The original theory by Kaluza and Klein, to be described shortly, considers one extra dimension and is able to unify gravity and classical electromagnetism – by the time of the original papers the only known fundamental interactions – into a unique framework in five dimensions. These fundamental interactions only look different for us, four-dimensional beings, because after the dimensional reduction, to be explained later, the single five dimensional fundamental field splits into $g_{\mu\nu}$ and A_μ as we are used to. This original setup has its problems though, like the presence of an unwanted scalar field besides the metric and the electromagnetic field. But let us start from the beginning, introducing the higher dimensional space.

In general, Kaluza-Klein (KK) theories consider the background space as the product of a 4-dimensional observable space with a n-dimensional compact space, $K_D = M_4 \times K_n$, where D is the total dimension of spacetime. We will denote the 4D usual coordinates as $x^\mu = (t, \vec{x})$ and the coordinates of the compact subspace as $\vec{y}$. All in all, fundamental fields in D dimensions are functions of both x^μ and $\vec{y}$, for instance a scalar field is $\hat{\phi} = \hat{\phi}(t, \vec{x}, \vec{y}) \equiv \hat{\phi}(x^M)$. Throughout this section, hatted quantities represent higher dimensional ones, while their 4D versions appear as usual, i.e. without hats.

Note that, since the background space is a described by the product $M_4 \times K_n$, the background metric should be written in the diagonal form:

$$\hat{g}^{(0)}_{MN}(x^P) = \begin{pmatrix} g^{(0)}_{\mu\nu}(x^\rho) & 0 \\ 0 & g^{(0)}_{mn}(\vec{y}) \end{pmatrix} ; \tag{3.1}$$

where the superscript (0) stands for background.

Of course, this diagonal form plus the fact that $g^{(0)}_{\mu\nu}$ is only a function of x^ρ while $g^{(0)}_{mn}$ is only a function of $\vec{y}$ is a feature of the background metric only. The full metric will have the form:

$$\hat{g}_{MN}(x^P) = \begin{pmatrix} g_{\mu\nu}(x^P) & g_{\mu n}(x^P) \\ g_{m\nu}(x^P) & g_{mn}(x^P) \end{pmatrix} \tag{3.2}$$

But the question is: How are we going to conciliate this 5D versions of fields with the 4D fields which are used with great success in our conventional 4-dimensional theories? The answer is the KK dimensional reduction procedure. First, one needs to find the 4-dimensional fields within this 5D setup. Consider, for simplicity, a first example where $K_n = S^1$ (a circle). In this case, one can Fourier expand the fields in the y coordinate. The simplest case is the scalar field:

$$\hat{\phi}(x^M) = \sum_{n=-\infty}^{\infty} e^{\frac{iny}{R}} \phi_n(x^\mu) ; \tag{3.3}$$

where R is the radius of the extra dimension.

In equation (3.3), $\{\phi_n\}$ is an infinite set of 4D fields. These are the fields one would deal with in 4D theories. If one considers now the second most usual example $K_n = S^2$, the 6D fields could be expanded in spherical harmonics $Y_{lm}(\vec{y}) \equiv Y_{lm}(\theta, \varphi)$. In both cases, one finds the functions in which one should expand the higher dimensional fields by evaluating the eigenfunctions of the Laplacean operator (Δ_n):

$$\begin{aligned} \Delta_1 e^{\frac{iny}{R}} \equiv \partial_y^2 e^{\frac{iny}{R}} &= -\frac{n^2}{R^2} e^{\frac{iny}{R}} ; \\ \Delta_2 Y_{lm}(\theta, \varphi) &= -\frac{l(l+1)}{R^2} Y_{lm}(\theta, \varphi) \end{aligned} \tag{3.4}$$

Notice that the eigenvalues are $\propto R^{-2}$, this point will be relevant later on. For a general K_n subspace, one is always able to expand in generalized spherical harmonics $Y_q^{I_q}$ where q is analogous to l and n in equation (3.4) defining the eigenvalues of the Laplacean operator while I_q is analogous to m, just an index in the representation of the symmetry group of the compact space. Let us call the eigenvalue of the Laplacean $-m_q^2$ (q is to remind us of the q-dependence of the eigenvalues), for a general K_n, such that:

$$\Delta_n Y_q^{I_q}(\vec{y}) = -m_q^2 Y_q^{I_q}(\vec{y}) \tag{3.5}$$

A D-dimensional scalar field is then expanded as:

$$\hat{\phi}(x^M) = \sum_{q,I_q} \phi_q^{I_q}(x^\mu) Y_q^{I_q}(\vec{y}) \tag{3.6}$$

Note now that, since the background space is defined by the product $M_4 \times K_n$, the D'Alembertian operator in D dimensions becomes $\Box_D = \Box_4 + \Delta_n$. Therefore, if one starts with a D-dimensional massless scalar field the infinite set of fields in 4 dimensions may be massive. This can be shown starting from the equation of motion of the D-dimensional scalar:

$$\Box_D \hat{\phi} = 0 \therefore \sum_{q,I_q} (\Box_4 + \Delta_n) \phi_q^{I_q}(x^\mu) Y_q^{I_q}(\vec{y}) = \sum_{q,I_q} [(\Box_4 - m_q^2) \phi_q^{I_q}(x^\mu)] Y_q^{I_q}(\vec{y}) = 0$$

$$\Rightarrow (\Box_4 - m_q^2) \phi_q^{I_q} = 0 \tag{3.7}$$

It is important to highlight that in Kaluza-Klein theories spacetime is fundamentally D-dimensional, therefore the action for the fields, and thus the equations of motion also, should be written in D dimensions first and then we rewrite it in terms of 4-dimensional fields to realize the implications to our 4D world. That is why the D-dimensional Klein-Gordon (KG) equation is satisfied in equation (3.7) and the second line is just a consequence of it.

Equation (3.7) shows us that the mass of each 4D field depends directly on the value of q. The problem now is that there are analogous expansions for non-scalar fields too, in analogy to what was just derived, but one doesn't find an infinite set of metrics or vector fields in the Standard Model. Where do all these fields go? Even more intriguing is the fact that physics seems to only depend on x^μ coordinates. Why is the $\vec{y}$-dependence irrelevant? The solution comes with the reduction ansatz. Since m_q is expected to be $\propto R^{-1}$, as previously pointed out, at least for manifolds in which defining a radius make sense, if we guess that R is very small, then the KK expansion is dominated by the zero mode $\phi_0^{I_0} Y_0^{I_0}$ and the other 4-dimensional fields, $\phi_q^{I_q}$ for $q > 0$, can be put to zero. Thus:

$$\hat{\phi}(x^M) \simeq \phi_0^{I_0}(x^\mu) Y_0^{I_0}(\vec{y}) \tag{3.8}$$

Moreover, if $Y_0^{I_0}$ is $\vec{y}$-independent as it is in the simple cases where $K_n = S^1$ or $K_n = S^2$, we have completely gotten rid of the $\vec{y}$-dependence. It is important to note that the KK reduction is an ansatz, i.e. a guess, and it can perhaps generate inconsistencies in the theory. An inconsistent truncation, i.e. reduction guess, is one which doesn't obey the equations of motion in D dimensions, an explicit example will be seen when we discuss the original KK scenario. On the other hand, the KK expansion is always valid.

In summary, the dimensional reduction procedure consists of two main steps: the KK expansion, which is always consistent, and the KK reduction, which may bring inconsistencies to the theory. In the middle steps it was important to note that the KK background metric assumes a diagonal form allowing us to write the D-dimensional D'Alembertian as a 4-dimensional one plus a $(D-4)$-dimensional Laplacean.

Although we have been focusing on scalar fields, for simplicity, this procedure is generalized to any type of field. For example, a D-dimensional gauge field splits into:

$$\hat{A}_M(x^N) = \big(A_\mu(x^N), A_m(x^N)\big) \tag{3.9}$$

Assuming the D-dimensional theory is invariant under $SO(1,3+n)$, $\hat{A}_M$ transforms as a vector under this group transformations and A_μ transforms as a vector under $SO(1,3) \subset SO(1,3+n)$. However, since the KK reduction procedure breaks $SO(1,3+n)$ into $SO(1,3) \times SO(n)$, A_m transforms as a scalar under $SO(1,3)$. Therefore, we end up with one 4-vector field and n 4-dimensional scalars. The $SO(n)$ group becomes, from the viewpoint of the 4D theory, an internal symmetry whose group elements act on the 4-scalars A_m. That is another appealing feature of KK theories. One may find a way to geometrize internal symmetries of the 4-dimensional Standard Model.

The KK expansion of A_M is:

$$A_\mu(x^M) = \sum_{q,I_q} A_\mu^{qI_q}(x^\mu) Y^{qI_q}(\vec{y})\,;$$
$$A_m(x^M) = \sum_{q,I_q} A^{qI_q}(x^\mu) Y_m^{qI_q}(\vec{y})\,; \tag{3.10}$$

where m, in $Y_m^{qI_q}$, is an index in the local Lorentz group of the space, while I_q is an index in the isometry group of it.

The generalization from this case to the metric one is straight forward. The metric in equation (3.2) contains a $(0,2)$-tensor field in 4D, which can be associated with the 4-dimensional metric, m 4-vector fields $g_{\mu n}$ and m^2 4-scalars g_{mn}. Those are KK expanded in a way which is analogous to equation (3.10).

Before we introduce the KK original theory it will be useful to define Newton's constant and the reduced Planck mass in arbitrary dimensions. Newton's constant is defined such that the Einstein-Hilbert action in any dimensions has the same form as in equation (1.3).

$$\hat{S}_{E-H} = \frac{1}{16\pi G^{(D)}} \int d^D x \sqrt{-\hat{g}}\,\hat{R}\,; \tag{3.11}$$

where hatted quantities are D-dimensional as previously defined, except for $G^{(D)}$ which for now we explicitly write the dimensionality.

Considering $\hat{g}_{MN}$ as the background metric, since this metric is block diagonal, one can write $\sqrt{-\hat{g}} = \sqrt{-g}\sqrt{\tilde{g}}$, where $\tilde{g}$ is the determinant of the compact subspace metric. Furthermore, the KK expansion allows us to split the dependence of $\hat{R}$ into a piece which depends only on x^μ and another which depends on $\vec{y}$. In the case one ends up with only a x^μ-dependence it is clear to see that:

$$\hat{S}_{E-H} = \frac{1}{16\pi G^{(D)}} \left(\int d^n y \sqrt{\tilde{g}} \right) \int d^4 x \sqrt{-g} \hat{R}(x^\mu) \Rightarrow G^{(D)} \equiv G^{(4)} \int d^n y \sqrt{\tilde{g}}\,; \tag{3.12}$$

where the integral over $\vec{y}$ is just the volume of the compact subspace so that $G^{(D)} = G^{(4)} V^{(n)}$.

In order to define the higher dimensional reduced Planck mass, $M_{Pl}^{(D)}$, one must note that the only source of dimensionality of $\hat{R}$ comes from the derivatives, since the metric is dimensionless in any dimensions. Therefore, since in any dimensions $R \sim g^{-1}(\partial\Gamma + \Gamma\Gamma)$ and $\Gamma \sim g^{-1}\partial g$ then $[\sqrt{-\hat{g}}\hat{R}] = 2$. Moreover, since an action should be dimensionless and $[M_{Pl}^{(D)}] = 1$, after all it is a mass, it is easy to see that a natural definition for $M_{Pl}^{(D)}$ which makes the D-dimensional Einstein-Hilbert action resemble its 4D version is:

$$\hat{S}_{E-H} = \frac{[M_{Pl}^{(D)}]^{D-2}}{2} \int d^D x \sqrt{-\hat{g}} \hat{R} \Rightarrow \left[M_{Pl}^{(D)} \right]^{2-D} = 8\pi G^{(D)}\,; \tag{3.13}$$

which gives the right relation, $M_{Pl}^{-2} = 8\pi G$, in 4 dimensions.

Finally, we are able to introduce the original KK theory. The analysis done here is strongly based in the reviews [41, 46] and in the paper [47]. It is done in some level of detail, including the appendix D, since it will be important in chapter 6. Kaluza and Klein proposed, as anticipated, an unification of classical electromagnetism and gravity. This unification is realized in five dimensions. The background space is $M_4 \times S^1$ leading to a background metric of the form $\hat{g}^{(0)}_{MN} = \mathrm{diag}(\eta_{\mu\nu}, 1)$. The KK expansion of such a metric is:

$$\hat{g}_{MN}(x^P) = \begin{pmatrix} \eta_{\mu\nu} + \sum_n h^{(n)}_{\mu\nu}(x^\rho) e^{\frac{iny}{R}} & \sum_n h^{(n)}_{\mu 5}(x^\nu) e^{\frac{iny}{R}} \\ \sum_n h^{(n)}_{5\nu}(x^\nu) e^{\frac{iny}{R}} & 1 + \sum_n \varphi_n(x^\mu) e^{\frac{iny}{R}} \end{pmatrix} \tag{3.14}$$

A linear parametrization of the KK reduction ansatz (linear ansatz) means $g_{\mu\nu}(x^\rho) \equiv \eta_{\mu\nu} + h^{(0)}_{\mu\nu}(x^\rho)$, $A_\mu(x^\nu) \equiv h^{(0)}_{\mu 5}(x^\nu)$ and $\phi(x^\mu) \equiv 1 + \varphi_0(x^\mu)$, leading to a dimensionally reduced metric of the form:

$$\hat{g}_{MN} = \begin{pmatrix} g_{\mu\nu} & A_\mu \\ A_\nu & \phi \end{pmatrix} \; ; \tag{3.15}$$

where $g_{\mu\nu}$, A_μ and ϕ are interpreted as the metric, a gauge field and a scalar field, respectively, in four dimensions. Note that the y-dependence vanished as expected. Even though this linear ansatz is perhaps the simplest possible guess, it is not the most convenient. The following non-linear parametrization yields an action which better resembles the well-known Einstein-Hilbert plus electromagnetic action (when ϕ is set to one) in four dimensions, as shown in appendix D:

$$\hat{g}_{MN} = \phi^{-1/3} \begin{pmatrix} g_{\mu\nu} + \kappa^2 \phi A_\mu A_\nu & \kappa\phi A_\mu \\ \kappa\phi A_\nu & \phi \end{pmatrix} \; ; \tag{3.16}$$

where $\kappa \equiv 4\sqrt{\pi G}$ is a dimensional factor introduced to ensure $[A_\mu] = 1$ as usual in four dimensions and to generate the usual normalization factor in front of the electromagnetic term of the action. The dimensionality of the scalar field is not the usual, but it is redefined later in order to yield the usual kinetic term for scalars, see appendix D. We find out in equation (D.33) that:

$$\hat{S}_{E-H} = \int d^4x \sqrt{-g} \left(\frac{R}{16\pi G} - \frac{1}{4} e^{-\sqrt{6}\frac{\varphi}{M_{Pl}}} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi \right) + \oint_{\partial\Omega} d\Sigma_\mu \sqrt{-g} X^\mu \; ; \tag{3.17}$$

where $\varphi \equiv -\frac{\log\phi}{4\sqrt{3\pi G}} = -\frac{M_{Pl}}{\sqrt{6}} \log\phi$ and X^μ is a vector field which is irrelevant for the equations of motion.

Therefore, when one puts $\phi = 1$, or equivalently $\varphi = 0$, one is left with exactly a gravity plus electromagnetism action in four dimensions! This was done by Kaluza. However, it is inconsistent to do so, meaning that at least one of the equations of motion in 5D is not respected. More precisely, note, from equation (D.32) for instance, that if one calculates the equation of motion for ϕ they can be written as $F_{\mu\nu}F^{\mu\nu}$ = something which is constant if ϕ is set to 1. Therefore, setting $\phi = 1$ means $F_{\mu\nu}F^{\mu\nu}$ = constant, which is not valid in general, since A_μ (and $F_{\mu\nu}$ as a consequence) is not constant. Inconsistencies happen when one sets to a constant a field which appears in a linear term in the action, as it is the case of setting $\phi = 1$. Notice A_μ can be put to zero consistently, since the only term containing this field in the action (the one proportional to $F_{\mu\nu}F^{\mu\nu}$) is quadratic in it. This case was considered in appendix D.

3.2 Basics of Supersymmetry and Supergravity

The other important prerequisite to string theory is supersymmetry (SUSY). This is how one calls the symmetry that relates bosonic to fermionic degrees of freedom. SUSY is an extension of the Poincaré symmetry, whose group is called $ISO(1,3)$ including the $SO(1,3)$ group generators plus spacetime translations. This section is based on chapters 14 and 15 of [12] and on chapters 3 and 4 of [48]. More details can be found there.

We shall illustrate the role of supersymmetry introducing the simple Wess-Zumino model in $d = 2$ dimensions. In this model, one considers a Majorana fermion, which can be thought to have real components. Since a Dirac fermion has $n = 2^{\lfloor d/2 \rfloor}$ complex components, a Majorana fermion has the same amount of real components, which in $d = 2$ means $n = 2$ degrees of freedom without using the equations of motion, i.e. off-shell. When one considers the equation of motion, the Dirac equation for spin-1/2 fermions, the number of degrees of freedom drops to half the off-shell one, since $(\partial\!\!\!/ + m)\psi = 0$ relates half the components to the other half. All in all, a 2D Majorana fermion has $n/2 = 1$ degree of freedom on-shell, i.e. taking the equation of motion into consideration. In order to match the only degree of freedom of an on-shell Majorana fermion one must consider a real scalar field, which off and on-shell has one degree of freedom since the Klein-Gordon equation only restricts the 4-momentum of the field. Thus, the action for a free $d = 2$ Wess-Zumino model is:

$$S = -\frac{1}{2}\int d^2x \left(\partial_\mu\phi\partial^\mu\phi + \bar{\psi}\partial\!\!\!/\psi\right) \; ; \tag{3.18}$$

where $\partial\!\!\!/ \equiv \gamma^\mu\partial_\mu$, ϕ is a real scalar field and ψ is a Majorana fermion.

To show the above action is invariant under the supersymmetric transformation one must first define such transformation. Since we want to relate the bosonic to the fermionic degree of freedom in it, otherwise it would not be a supersymmetry by construction, the variations of the fields under this transformation should be like $\delta\phi \propto \psi$ and $\delta\psi \propto \phi$. In order to respect Lorentz symmetry, i.e. to make sure the left and right hand sides of the variation equations do transform in the same manner under a Lorentz transformation, it is clear that the transformation parameter ϵ must be a spinor itself. One then could define ϵ such that:

$$\delta\phi = \bar{\epsilon}\psi = \epsilon^C\psi \equiv \epsilon^T C\psi \, ; \tag{3.19}$$

where ϵ^C is the Majorana conjugate of ϵ and C is the charge conjugation matrix. Recall that a Majorana spinor satisfies the reality condition $\bar{\chi} = \chi^C$.

The charge conjugation matrix is defined in $d = 2$ and $d = 4$ dimensions as:

$$C^T = -C \quad ; \text{and} \quad C\gamma^\mu C^{-1} = -(\gamma^\mu)^T \tag{3.20}$$

If one guesses the variation of ψ as $\phi\epsilon$, in analogy to equation (3.19), the dimensions would not match. Note, from equation (3.18), that since $[S] = 0$, $[d^2x] = -2$ and $[\partial_\mu] = 1$ the dimension of the fields are $[\phi] = 0$ and $[\psi] = 1/2$. Furthermore, from equation (3.19), $[\epsilon] = -1/2$, therefore one needs something with dimension of mass, i.e. dimension = 1, along with $\phi\epsilon$ in order to match the dimension of $\delta\psi$. Since the theory is massless, one cannot multiply $\phi\epsilon$ by any mass. An invariant variation of ∂_μ would do the job though. Fortunately, one can get such an invariant quantity by contracting ∂_μ with γ^μ. Thus, our guess is:

$$\delta\psi = \partial\!\!\!/ \phi\epsilon \tag{3.21}$$

It turns out that our guess is correct, as will be shown now. From equation (3.18):

$$\delta S = -\int d^2x \left[-\delta\phi\Box\phi + \frac{1}{2}\delta\bar{\psi}\partial\!\!\!/\psi + \frac{1}{2}\bar{\psi}\partial\!\!\!/(\delta\psi)\right] \tag{3.22}$$

In order to rewrite the second and third terms as only one in equation (3.22), note that:

(i). First, $(C\gamma^\mu)$ is symmetric. From equation (3.20):

$$C\gamma^\mu C^{-1} = -(\gamma^\mu)^T \therefore C\gamma^\mu = -(\gamma^\mu)^T C = (\gamma^\mu)^T C^T = (C\gamma^\mu)^T$$
$$\Rightarrow (C\gamma^\mu)_{\alpha\beta} = (C\gamma^\mu)_{\beta\alpha} \tag{3.23}$$

(ii). Then, using the fact derived above:

$$\bar{\psi}\partial\!\!\!/(\delta\psi) = \psi^\alpha (C\gamma^\mu)_{\alpha\beta}[\partial_\mu(\delta\psi)]^\beta = -[\partial_\mu(\delta\psi)]^\beta (C\gamma^\mu)_{\beta\alpha}\psi^\alpha$$
$$= -\partial_\mu[\delta(\psi^\beta C_{\beta\gamma})](\gamma^\mu)^\gamma{}_\alpha \psi^\alpha = -\partial_\mu(\delta\bar{\psi})\gamma^\mu\psi\,; \tag{3.24}$$

where the minus sign in the second equality comes from the commutation of two Grassmann variables ψ^α and $[\partial_\mu(\delta\psi)]^\beta$.

(iii). Finally, using the previous result:

$$\int d^2x\bar{\psi}\partial\!\!\!/(\delta\psi) = -\int d^2x\partial_\mu(\delta\bar{\psi})\gamma^\mu\psi = \int d^2x\delta\bar{\psi}\partial\!\!\!/\psi\,; \tag{3.25}$$

where we just partially integrated it in the last equality.

Therefore, back to equation (3.22), one can rewrite the second term as being the same as the third one such that:

$$\delta S = -\int d^2x \left[-\delta\phi\Box\phi + \bar{\psi}\partial\!\!\!/(\delta\psi)\right] = -\int d^2x \left[-\bar{\epsilon}\psi\Box\phi + \bar{\psi}\partial\!\!\!/\partial\!\!\!/(\epsilon\phi)\right] \tag{3.26}$$

Two more steps are needed to arrive at the conclusion $\delta S = 0$.

(i). Note that:

$$\bar{\epsilon}\psi \equiv \epsilon^\alpha C_{\alpha\beta}\psi^\beta = -\psi^\beta C_{\alpha\beta}\epsilon^\alpha = \psi^\beta C_{\beta\alpha}\epsilon^\alpha \equiv \bar{\psi}\epsilon\,; \tag{3.27}$$

where in the second equality the minus sign is due to the commutation of two Grassmann variables ϵ^α and ψ^β and there is another minus sign in the third equality from the fact $C^T = -C$.

(ii). Moreover:

$$\partial\!\!\!/\partial\!\!\!/(\epsilon\phi) = \gamma^\mu\gamma^\nu\epsilon\partial_\mu\partial_\nu\phi = \frac{1}{2}\{\gamma^\mu,\gamma^\nu\}\epsilon\partial_\mu\partial_\nu\phi = g^{\mu\nu}I\epsilon\partial_\mu\partial_\nu\phi = \epsilon\Box\phi\,; \tag{3.28}$$

where in the second equality it was used the fact that μ and ν indices are symmetrized, since one can permute the partial derivatives freely, and in the third equality the Clifford algebra definition, $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}I$, was used, where I is the identity in the spinorial space.

Back to equation (3.26), we end up with:

$$\delta S = -\int d^2x \left(-\bar{\psi}\epsilon\Box\phi + \bar{\psi}\epsilon\Box\phi\right) = 0\,; \tag{3.29}$$

as promised.

It may seem weird the fact that we didn't use the equations of motion anywhere to show that $\delta S = 0$, despite having only an on-shell supersymmetry. Nevertheless, the invariance of the action is only one requirement of a symmetry. One also needs to check if the algebra of that symmetry close on the fields, meaning that if we apply the left hand side of the algebra relation on the fields we must arrive at right hand side of it at the end. This is what happens in the $d = 2$ Wess-Zumino model, but one has to use the fermionic equation of motion, $\partial\!\!\!/\psi = 0$, in order to make it work. Therefore, we indeed have an on-shell SUSY. We will not show it here, but it is done in full detail in book [48] chapter 3.

Up until now we have been considering a global supersymmetry, i.e. the parameter ϵ does not depend on spacetime coordinates. Inspired by gauge symmetries in usual non-supersymmetric quantum field theories (QFTs), one can try and gauge the symmetry, i.e. make it local $\epsilon = \epsilon(x^\mu)$. In general, this procedure requires the introduction of a gauge field. One of the simplest examples is found when we gauge the global $U(1)$ symmetry of quantum electrodynamics (QED). Before gauging the symmetry we have only fermions in the theory, but after it, the electromagnetic (photon) field A_μ arises naturally to ensure the symmetry is respected. The gauge field associated to local supersymmetry will be a spin-3/2 field called gravitino, the superpartner of the graviton – analogous to what ψ was to ϕ in the $d = 2$ free Wess-Zumino model. Gauging a rigid, i.e. global, SUSY gives us supergravity.

As explained in the last paragraph, supergravity can be seen as a local supersymmetric theory. However, one can also view it as a supersymmetric theory of gravity. In order to do that, one must use the Vielbein-Spin Connection formulation of GR.

First, we must check whether a supersymmetry can be defined with just the graviton and the gravitino by counting their degrees of freedom. In d dimensions the metric is a symmetric tensor field, represented by a symmetric matrix, which has $d(d+1)/2$ independent components. However, due to general coordinate transformation symmetry, which can be used to fix d components, we end up with $d(d+1)/2 - d = d(d-1)/2$ off-shell degrees of freedom for the graviton. On-shell, we analyze the graviton field $h_{\mu\nu}$ defined from $g_{\mu\nu} = \eta_{\nu\mu} + 2\kappa h_{\mu\nu}$, where $\kappa \equiv 4\sqrt{\pi G} = \sqrt{2}M_{Pl}^{1-\frac{d}{2}}$ in d arbitrary dimensions. A small field expansion in κ of the Einstein-Hilbert action leads us to the Fierz-Pauli action in terms of the graviton field, see [49] chapter 40. In analogy to the electromagnetic case, one also has a gauge freedom for $h_{\mu\nu}$. One can choose, for instance, the de Donder gauge, in analogy to the Lorenz gauge of electromagnetism:

$$\partial^\nu \bar{h}_{\mu\nu} \equiv \partial^\nu \left(h_{\mu\nu} - \frac{h^\rho{}_\rho}{2} \eta_{\mu\nu} \right) = 0 \tag{3.30}$$

The above condition removes d graviton degrees of freedom. In summary, while the graviton has $d(d-1)/2$ off-shell degrees of freedom, on-shell it has $d(d-1)/2 - d = (d^2 - 3d)/2 = [(d-1)(d-2)/2] - 1$ degrees of freedom. The on-shell degrees of freedom were rewritten this way to point out it corresponds to a $(d-2) \times (d-2)$ symmetric traceless matrix.

For the gravitino, ψ_μ^α one has the spinorial index α running from 0 to $n = 2^{\lfloor d/2 \rfloor}$ as in the case of a spin-1/2 particle, but now one also has to consider the spacetime index μ running from 0 to d. This leads to nd components. However, the supergravity (local supersymmetry) transformation $\delta\psi_\mu = D_\mu \epsilon$ can fix n components of it, since ϵ is a spinor as we will see. Thus, one ends up with $nd - n = n(d-1) = 2^{\lfloor d/2 \rfloor}(d-1)$ degrees of freedom off-shell. Note that, in $n(d-1)$ degrees of freedom, n comes from the spinorial behavior of the gravitino while $(d-1)$ seems to be a characteristic of a spin-1 gauge field – the electromagnetic field also has $d-1$ degrees of freedom off-shell, reducing to $d-2$ on-shell. Therefore, naively one would expect the degrees of freedom on-shell to be $(n/2)(d-2)$, since the equations of motion for a spin-1/2 fermion reduces the degrees of freedom to half the number of components, as it was already discussed, and a spin-1 gauge field also loses a degree of freedom due to the gauge condition applied. It turns out this guess is almost right, but one has to realize that combining a spinorial and a spacetime index, more precisely via a product of the spin-1/2 and the spin-1 representations of the Lorentz group, decomposes into two irreducible representations, one spin-1/2 and the other spin-3/2:

$$\frac{1}{2} \otimes 1 = \frac{1}{2} \oplus \frac{3}{2} \tag{3.31}$$

All in all, since we are only interested in the on-shell degrees of freedom of a spin-3/2 field, we must subtract the $n/2$ degrees of freedom of the spin-1/2 irreducible representation from $(n/2)(d-2)$, leaving us with $(n/2)(d-3) = 2^{\lfloor d/2 \rfloor}(d-3)/2$ on-shell degrees of freedom for the gravitino.

In four dimensions both the graviton, $[3(2)/2] - 1 = 2$, and the gravitino, $2^2(1/2) = 2$, have the same number of on-shell degrees of freedom. Therefore, we are able to construct a $\mathcal{N} = 1$, i.e. only one supersymmetry (or equivalently one symmetry parameter), supergravity theory in four dimensions with just the graviton and the gravitino.

The Vielbein-Spin Connection formulation of general relativity is obtained by, first, rewriting the metric in terms of the vielbein e^a_μ:

$$g_{\mu\nu}(x) = e^a_\mu(x) e^b_\nu(x) \eta_{ab} \, ; \tag{3.32}$$

where x stands for spacetime coordinates, usually called x^μ, and a and b are indices in the flat space tangent to the generically curved manifold at the point $p \in \mathcal{M}$ whose coordinates are x.

In this description of GR, the vielbein makes the local Lorentz symmetry explicit:

$$e^a_\mu \to \Lambda^a{}_b e^b_\mu \tag{3.33}$$

Therefore, although one may think the vielbein has more independent components than the metric ($d^2 \geq d(d+1)/2 \, ; \forall \, d \in \mathbb{N}$), which would make the descriptions necessarily inequivalent, equation (3.33) fixes $d(d-1)/2$ components, the number of components of the antisymmetric matrix $\Lambda^a{}_b$, leaving only $d^2 - d(d-1)/2 = d(d+1)/2$ independent components for the vielbein, matching perfectly the number of components of the metric.

We are yet to define the action of covariant derivatives on fermions. Since the supergravity theory must include the gravitino, it is essential to define it. The spin connection ω^{ab}_μ arises then as a connection coefficient function, or a gauge field in analogy to A^a_μ of a non-abelian gauge theory, in the definition of such derivative:

$$D_\mu \psi = \partial_\mu + \frac{1}{4}\omega_\mu^{ab}\Gamma_{ab}\psi\,; \tag{3.34}$$

where $(1/4)\Gamma_{ab}$ is the generator of Lorentz transformations in the spinor representation.

It is important to point out that the spin connection cannot be independent of the vielbein. We have already counted the number of independent components of the vielbein and realized it matches the number of components of the metric. This number of components are related to the number of degrees of freedom and two theories with different degrees of freedom cannot be equivalent. Therefore, we must find a way to write the spin connection as a function of the vielbein. This is achieved by imposing the no-torsion constraint, see [48] chapter 4.

In order to rewrite the Einstein-Hilbert action in terms of e_μ^a and $\omega_\mu^{ab} = \omega_\mu^{ab}(e_\nu^c)$, one must realize that the field strength tensor of the spin connection $R_{\mu\nu}^{ab} = \partial_\mu\omega_\nu^{ab} - \partial_\nu\omega_\mu^{ab} + \omega_\mu^{ac}\omega_{\nu c}{}^b - \omega_\nu^{ac}\omega_{\mu c}{}^b$ is related to the Riemann tensor through:

$$R_{\rho\sigma}^{ab}(\omega(e)) = e_\mu^a(e^{-1})^{\nu b}R^\mu{}_{\nu\rho\sigma}(\Gamma(g(e))) \tag{3.35}$$

It won't be done here but one can check the veracity of equation (3.35) by rewriting the Christoffel symbols, which give us the expression for the Riemann tensor, in terms of the vielbein using equation (3.32) and then contracting the Riemann tensor with e and e^{-1} as it is done above arriving at the expression for the spin connection field strength given in the last paragraph.

From equation (3.35) one finds out the expression for the Ricci scalar in terms of e and $\omega = \omega(e)$. First, note that:

$$\begin{gathered} g_{\mu\nu} = e_\mu^a e_{\nu a}\ ;\ \text{and}\ g^{\mu\nu} = (e^{-1})_a^\mu (e^{-1})^{\nu a} \\ \Rightarrow \delta^\rho{}_\nu = g^{\rho\mu}g_{\mu\nu} = (e^{-1})_a^\rho(e^{-1})^{\mu a}e_{\mu b}e_\nu^b = (e^{-1})_a^\rho\delta^a{}_b e_\nu^b = (e^{-1})_a^\rho e_\nu^a = (e^{-1})^{\rho a}e_{\nu a} \end{gathered} \tag{3.36}$$

Similarly, one gets $\delta^a{}_b = (e^{-1})^{\mu a}e_{\mu b} = e_\mu^a(e^{-1})_b^\mu$. Then, from equation (3.35):

$$
\begin{aligned}
(e^{-1})^{\lambda}_{a} e_{\xi b} R^{ab}_{\rho\sigma} = (e^{-1})^{\lambda}_{a} e^{a}_{\mu} (e^{-1})^{\nu b} e_{\xi b} R^{\mu}{}_{\nu\rho\sigma} = R^{\lambda}{}_{\xi\rho\sigma} \\
\Rightarrow R_{\mu\nu} \equiv R^{\rho}{}_{\mu\rho\nu} = (e^{-1})^{\rho}_{a} e_{\mu b} R^{ab}_{\rho\nu} \\
\Rightarrow R \equiv g^{\mu\nu} R_{\mu\nu} = (e^{-1})^{\mu c} (e^{-1})^{\nu}_{c} (e^{-1})^{\rho}_{a} e_{\mu b} R^{ab}_{\rho\nu} = (e^{-1})^{\nu}_{b} (e^{-1})^{\rho}_{a} R^{ab}_{\rho\nu}
\end{aligned}
\tag{3.37}
$$

With the result of equation (3.37) on hands, we only need to rewrite $\sqrt{-g}$ in terms of e to fully find the Einstein-Hilbert action in terms of the vielbein. From equation (3.32):

$$
\det(g_{\mu\nu}) = (\det e)^2 \det(\eta_{ab}) = -(\det e)^2 \therefore \sqrt{-g} = \det e \tag{3.38}
$$

Finally, in $d = 4$ dimensions, the Einstein-Hilbert action becomes:

$$
S_{E-H} = \frac{1}{16\pi G} \int d^4x (\det e) (e^{-1})^{\nu}_{b} (e^{-1})^{\rho}_{a} R^{ab}_{\rho\nu} \tag{3.39}
$$

To complete the $\mathcal{N} = 1$ supergravity action one must add to the above action the contribution coming from a the spin-3/2 gravitino field. In flat space, Rarita and Schwinger wrote the action for a spin-3/2 field that goes by their names. In $d = 4$ dimensions:

$$
S_{R-S} = -\frac{1}{2} \int d^4x \bar{\psi}_{\mu} \gamma^{\mu\nu\rho} \partial_{\nu} \psi_{\rho} \Rightarrow S_{R-S} = -\frac{1}{2} \int d^4x (\det e) \bar{\psi}_{\mu} \gamma^{\mu\nu\rho} D_{\nu} \psi_{\rho}\,; \tag{3.40}
$$

where after the $\Rightarrow$ we generalize it to curved space as usual, by introducing a factor of $\sqrt{-g} = \det e$ and replacing $\partial_{\nu} \to D_{\nu}$. Note that $\gamma^{\mu\nu\rho} \equiv \gamma^{[\mu}\gamma^{\nu}\gamma^{\rho]}$.

Finally, we are able to define the full $\mathcal{N} = 1$ supergravity action in 4D as:

$$
S_{\mathcal{N}=1\,SUGRA} = S_{E-H} + S_{R-S}\,; \tag{3.41}
$$

where S_{E-H} and S_{R-S} were described in equations (3.39) and (3.40), respectively.

The above action is invariant under the supersymmetric transformation, see [48], defined by:

$$\delta e_\mu^a = \sqrt{2\pi G}\bar{\epsilon}\gamma^a\psi_\mu \;\; ; \text{ and } \;\; \delta\psi_\mu = \frac{1}{\sqrt{8\pi G}} D_\mu \epsilon \, ; \tag{3.42}$$

which can be found by educated guesses, using dimensional analysis and matching the different types of indices. The constant factors multiplying the transformations are conventional. One can include $\sqrt{2\pi G}$ in the definition of ϵ in order to get a unity multiplicative factor in the vielbein variation as long as the factor in the other transformation law changes accordingly.

Up until now we have introduced one example of rigid supersymmetry and another one of local supersymmetry, i.e. supergravity. In both cases, it seems like one has to guess the supersymmetry transformation laws. Thus, a natural question that may arise is: Can we find a formalism in which supersymmetry is manifest? The answer is yes, we can! This is called the superspace formalism, see [48] chapter 3. In this formalism one considers an expansion of our notion of space. Fields are not only functions of the usual spacetime (bosonic) coordinates x anymore. They are rather functions of x and a fermionic coordinate θ.

We want to couple matter multiplets to supergravity. One possibility is adding Wess-Zumino, or chiral, multiplets. These multiplets are composed, in $d = 4$ dimensions, by a Majorana fermion and a complex scalar, since the number of degrees of freedom of fermions doubles from $d = 2$ (previously studied case) to $d = 4$ dimensions and the bosonic degrees of freedom has to follow this increase. The Wess-Zumino multiplet is said to have spin content $(1/2, 0)$ for that reason. In the superspace formalism, the multiplet is characterized by the chiral superfield $\Phi = \Phi(x, \theta)$. No detail will be given here, but the relevant point for later chapters is that the most general action for a chiral superfield is written in terms of the so-called Kähler potential, $K = K(\Phi, \Phi^\dagger)$, and the superpotential, $W = W(\Phi)$. The Kähler potential gives us kinetic terms while W gives us the interactions.

Another possibility is to couple a gauge multiplet, whose spin content is $(1, 1/2)$, to $\mathcal{N} = 1$ supergravity in four dimensions. This multiplet is composed of a gauge field A_μ^a, in the adjoint of a given group G, plus a fermion, hence the spin content $(1, 1/2)$. It is characterized by the function $\mathcal{F}(\{\Phi\})$, where Φs are chiral scalar multiplets. An important quantity for later chapters is the supersymmetric potential for scalars ϕ^i. In general, this potential for the scalars assumes the form:

$$V = e^{\frac{K}{M_{Pl}^2}} \left[\sum_{i\bar{j}} (g^{-1})^{i\bar{j}} D_i W (D_j W)^* - \frac{3}{M_{Pl}^2} |W|^2 \right] + \frac{1}{2} (\mathcal{F}^{-1})^{ab} \left(\frac{\partial K}{\partial \phi_i} (T_a)_{ij} \phi^j \right) \left(\frac{\partial K}{\partial \phi_k} (T_b)_{kl} \phi^l \right) ; \tag{3.43}$$

where the covariant derivative D_i is defined by $D_i = (\partial/\partial\phi^i) + (1/M_{Pl}^2)(\partial K/\partial\phi^i)$.

The first line of equation (3.43) is a contribution from the Wess-Zumino multiplets called F-term while the second line comes from the gauge multiplet and is known as D-term. T_a are the generators of the group G of the gauge multiplet.

3.2.1 $\mathcal{N} = 4$ Super Yang-Mills (SYM)

Now that we have had a taste of what supersymmetry is, there is only one more supersymmetric theory we must introduce now that will be essential later on when the AdS/CFT correspondence is introduced. That is the $\mathcal{N} = 4$ Super Yang-Mills (SYM) theory.

First, it should be pointed out that four is the maximum amount of supersymmetry possible for theories which contain fields with, at most, spins equal to one, i.e. spins ≤ 1, in $d = 4$ dimensions. The reason for that is $-1 \leq$ vacuum helicity $\leq +1$ and each supersymmetry changes the helicity by 1/2, see again chapter 3 of [48]. Therefore, in four steps ($\mathcal{N} = 4$ supersymmetries, i.e. four ϵ^I parameters) one goes from the maximum to the minimum helicity. One more supersymmetry would require the existence of a spin-3/2 field, which is not the case by construction.

We must describe now the field content of $\mathcal{N} = 4$ SYM. In order to earn the name "Yang-Mills", it has to be a gauge theory. It can be viewed, in 4D, as a dimensionally reduced theory coming from a 10-dimensional gauge theory. In the Kaluza-Klein section we have learned that a 10-dimensional vector field is decomposed into a 4-vector field and 6 scalars after the dimensional reduction. Therefore, one has $4-2=2$ on-shell degrees of freedom coming from the gauge fields A^a_μ plus 6 more coming from the scalar fields ϕ^i (i runs from 1 to 6), giving us a total of 8 bosonic on-shell degrees of freedom. In order to match that, one has to consider four spinors ψ^I (I runs from 1 to 4), since the number of on-shell degrees of freedom of each of them is $2^{\lfloor 4/2 \rfloor}/2 = 2$, leaving us with $4 \times 2 = 8$ on-shell fermionic degrees of freedom. Note that the scalar fields transform under $SO(6)$, since the KK reduction breaks $SO(1,9) \to SO(1,3) \times SO(6)$, and the spinors transform under $SU(4)$. Since those groups are isomorphic, one can rewrite the $SO(6)$ fundamental index i in the antisymmetric representation of $SU(4)$ as $[IJ]$, thus $\phi^i \to \phi^{[IJ]}$.

The SUSY transformation laws nor the action of $\mathcal{N}=4$ SYM will be fully written here but one can see that we must vary the gauge field into the fermions such that $\delta A \propto \bar{\epsilon}_I \psi^I$ (the gauge field indices were dropped since we are not caring about them here, but matching the indices would give the correct guess $\delta A^a_\mu = \bar{\epsilon}_I \gamma_\mu \psi^{Ia}$), and since $I = 1, \dots, 4$ it is clear we indeed have four supersymmetries, i.e. four fermionic symmetry parameters. The importance of $\mathcal{N}=4$ SYM lies in its uniqueness. The gauge theory with four supersymmetries and with canonical kinetic terms (no more than two derivatives) is $\mathcal{N}=4$ SYM.

3.3 Quantization of the Bosonic String

Finally, we are able to introduce string theory in its first version: the bosonic string. The theory starts off being classically described in full analogy with a point particle. The simplest free string action one can write is the generalization of the action for a free point-particle. The free relativistic point particle action is written as $S = -m \int d\tau$, where m is the mass of the particle, a physical quantity that describes it, and τ is the particle's proper time. Therefore, the stationary action principle, i.e. applying $\delta S = 0$, tells us the particle classically moves in a way to minimize (or maximize) its proper time, essentially the length of its worldline through spacetime. Analogously, a string sweeps out a worldsheet, i.e. 2-

dimensional version of the worldline, through spacetime. Therefore, one can guess that the action of the string will tend to minimize (or maximize) its worldsheet area. It turns out that this is the right guess. We may write $S = -T \int dA$ as the free string action, where T is the dimensional constant which ensures $[S] = 0$. Therefore, $[T] = 2$ and is the analogous of the particle's mass, describing the string. It is called the string tension, which has units of energy per length or energy squared in our units. For historical reasons, it is usually rewritten as $T = 1/(2\pi\alpha')$.

Since the worldsheet is a two-dimensional object, it has to be characterized by two parameters, τ and σ, instead of just one as in the point particle case. Instead of the proper time, we could have rewritten the particle's action such that it could assume any parametrization of the worldline, i.e. $X'^{\mu}(\tau') = X^{\mu}(\tau)$ where X^{μ} are worldline coordinates on spacetime and $\tau' = \tau'(\tau)$ is a monotonic function of τ. Similarly, τ and σ for the string can be any parameters that span the worldsheet and we must make sure the action is reparametrization invariant $X'^{\mu}(\tau'(\tau,\sigma), \sigma'(\tau,\sigma)) = X^{\mu}(\tau,\sigma)$ considering monotonic functions again. We saw that the action involves an integral over the area of the worldsheet, therefore, we must know how to express this area element $dA = d\sigma d\tau\sqrt{-\det(\gamma_{ab})}$, where γ_{ab} is the worldsheet metric and $a, b = 0, 1$, in terms of spacetime coordinates X^{μ}, where $\mu = 0, \ldots, D-1$ for a generic D-dimensional spacetime. More precisely, we must be able to write γ_{ab} as an induced metric, h_{ab}, from the 4-dimensional spacetime metric $g_{\mu\nu}$. This is done by starting with the spacetime line element, and considering the $\sigma^0 = \tau$ and $\sigma^1 = \sigma$, such that $X^{\mu}(\sigma^a)$. Thus:

$$ds^2 = g_{\mu\nu} dX^{\mu} dX^{\nu} \text{ ; and } dX^{\mu} = \frac{\partial X^{\mu}}{\partial \sigma^a} d\sigma^a$$

$$\Rightarrow ds^2_{induced} = g_{\mu\nu}(X) \frac{\partial X^{\mu}}{\partial \sigma^a} \frac{\partial X^{\nu}}{\partial \sigma^b} d\sigma^a d\sigma^b \equiv h_{ab} d\sigma^a d\sigma^b \therefore h_{ab} = \partial_a X^{\mu} \partial_b X^{\nu} g_{\mu\nu} \tag{3.44}$$

Therefore, the simplest string action, the so-called Nambu-Goto action is:

$$S_{N-G} = -\frac{1}{2\pi\alpha'} \int d^2\sigma \sqrt{-\det(h_{ab}(\tau,\sigma))}\,; \tag{3.45}$$

where $d^2\sigma \equiv d\sigma^0 d\sigma^1 \equiv d\tau d\sigma$.

Because of the square root factor, the action above is not very suitable for quantization. Were we able to rewrite the Nambu-Goto action without the square root, leaving just the quadratic in derivatives term, its path integral would be more easily evaluated. It turns out we are indeed able to do that by introducing the worldsheet metric as an auxiliary field. Note in equation (3.45) that from the viewpoint of the worldsheet, X^μ a D scalar fields of a field theory. Moreover, $S_{N-G} = S_{N-G}[X^\mu]$ is a functional of X^μ. The new Polyakov action is a functional now of both X^μ and γ_{ab}, $S_P = S_P[X^\mu, \gamma_{ab}]$. All in all one can write, since we are interested in describing a theory in flat spacetime $g_{\mu\nu} = \eta_{\mu\nu}$:

$$S_P = -\frac{1}{4\pi\alpha'} \int d^2\sigma \sqrt{-\gamma}\gamma^{ab}\partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu} \tag{3.46}$$

One can check that by varying the action with respect to γ^{ab} and setting $\delta S_P/\delta\gamma^{ab} = 0$ we obtain $\gamma_{ab} = \partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu}$, see [12] chapter 17. When plugged back into the Polyakov action it gives us $S_P = -T/2 \int d^2\sigma \sqrt{-\gamma}\gamma^{ab}\gamma_{ab} = -T \int d^2\sigma \sqrt{-\gamma} = S_{N-G}$ and hence the two formulations are equivalent, at least classically. However, the Polyakov action has one more symmetry than the Nambu-Goto one. Spacetime Poincaré invariance and worldsheet reparametrization were already a symmetry of the Nambu-Goto action, which were nicely carried to the Polyakov formulation too. But now, the later is Weyl invariant despite the fact that the first one is not. A Weyl transformation is defined as:

$$\begin{aligned} X^\mu &\to X'^\mu(\sigma^a) = X^\mu(\sigma^a) \\ \gamma_{ab} &\to \gamma'_{ab}(\sigma^c) = e^{2\omega}\gamma_{ab}(\sigma^c)\,; \quad \forall\, \omega = \omega(\sigma^a) \end{aligned} \tag{3.47}$$

and the Polyakov action is invariant under it since $\sqrt{-\gamma} \to e^{2\omega}\sqrt{-\gamma}$ while $\gamma^{ab} \to e^{-2\omega}\gamma^{ab}$.

Now, we must analyze the equations of motion. First, by varying the action with respect to γ^{ab} one gets:

$$T_{ab} \equiv -\frac{4\pi}{\sqrt{-\gamma}}\frac{\delta S_P}{\delta\gamma^{ab}} = \frac{\eta_{\mu\nu}}{\alpha'}\left(\partial_a X^\mu \partial_b X^\nu - \frac{1}{2}\gamma_{ab}\gamma^{cd}\partial_c X^\mu \partial_d X^\nu\right) = 0\,; \tag{3.48}$$

Note the extra factor of 2π conventionally used in string theory to define the energy-momentum tensor. For a peculiarity of two dimensional manifolds, we are able to do this redefinition without changing factors in the equations of motion even if we add an Einstein-Hilbert contribution to the Polyakov action, see [12]. Moreover, due to reparametrization and Weyl invariance of the Polyakov action, respectively, one gets the conservation of T_{ab} as well as $T^a{}_a = 0$ (off-shell).

The most interesting results come when varying the action with respect to X^μ. Consider τ as a time coordinate describing the motion of the string through space and σ as its spatial coordinate. Then, considering the boundaries of the worldsheet to be $-\infty < \tau < \infty$ and $0 \leq \sigma \leq l$, the variation of the action becomes:

$$\begin{aligned}\delta S_P &= -\frac{1}{2\pi\alpha'}\int d^2\sigma\sqrt{-\gamma}\gamma^{ab}\partial_a(\delta X^\mu)\partial_b X_\mu \\ &= \frac{1}{2\pi\alpha'}\left[\int d\tau\int_0^l d\sigma\sqrt{-\gamma}\delta X^\mu \Box X_\mu - \int d\tau\left(\sqrt{-\gamma}\delta X^\mu\partial_\sigma X_\mu\right)\Big|_{\sigma=0}^{\sigma=l}\right]\,;\end{aligned} \tag{3.49}$$

where $\partial_a(\sqrt{-\gamma}\gamma^{ab}\partial_b X_\mu) = \sqrt{-\gamma}\nabla_a(\gamma^{ab}\partial_b X_\mu) = \sqrt{-\gamma}\gamma^{ab}\nabla_a\partial_b X_\mu \equiv \sqrt{-\gamma}\Box X_\mu$.

When we set $\delta S/\delta X^\mu = 0$ the bulk variation of the metric leads to the equation of motion $\Box X_\mu = 0$, the wave equation in $1+1$ dimensions. On the other hand, the boundary term gives us:

$$(\delta X^\mu \partial_\sigma X_\mu)\big|_{\sigma=0}^{\sigma=l} = 0\,; \tag{3.50}$$

which is satisfied in three different, physically relevant ways:

1. Closed strings: The first way to satisfy the above equation is by considering $\sigma = 0$ and $\sigma = l$ as the same point, thus forming a loop, i.e. a closed string:

$$X^\mu(\tau, 0) = X^\mu(\tau, l) \ ; \ \gamma_{ab}(\tau, 0) = \gamma_{ab}(\tau, l) \tag{3.51}$$

2. Open strings: One is able to obtain open strings in two different fashions according to the following boundary conditions:

   - Neumann boundary conditions:

   $$\partial_\sigma X_\mu(\tau, 0) = \partial_\sigma X_\mu(\tau, l) = 0 \tag{3.52}$$

     With this condition the endpoints of the string move at the speed of light, see [12] chapter 17.

   - Dirichlet boundary conditions:

   $$\delta X^\mu(\tau, 0) = \delta X^\mu(\tau, l) = 0 \tag{3.53}$$

     This way the endpoints are fixed and the strings end on D-branes, to be discussed later.

   The possibility of having one end of the open string respecting a Neumann boundary condition while the other obeys a Dirichlet boundary condition will not be discussed here.

It is important to notice that since we have three worldsheet invariances of the Polyakov action (two reparametrizations plus Weyl invariance), we can use them to set the three independent components of the worldsheet metric to whatever we find suitable by fixing a gauge. The simple conformal gauge, or more precisely unit gauge, is defined by setting $\gamma_{ab} = \eta_{ab}$. In this gauge, the equations of motions for the string becomes:

$$\Box X^\mu = 0 \therefore \partial^a \partial_a X^\mu = (-\partial_\tau^2 + \partial_\sigma^2) X^\mu = 0 \tag{3.54}$$

However, one must not be fooled by the simplicity of equation (3.54) and has to recall the equations of motion for the worldsheet metric (3.48) still have to be respected, which gives us two constraints:

$$T_{ab} = \frac{1}{\alpha'} \left(\partial_a X^\mu \partial_b X_\mu - \frac{1}{2} \eta_{ab} \partial_c X^\mu \partial^c X_\mu \right) = 0$$
$$\Rightarrow \begin{cases} \alpha' T_{00} & = \alpha' T_{11} = \frac{1}{2}(\dot{X}_\mu \dot{X}^\mu + X'_\mu X'^\mu) = 0 \therefore \dot{X}_\mu \dot{X}^\mu + X'_\mu X'^\mu = 0 \\ \alpha' T_{01} & = \alpha' T_{10} = \dot{X}^\mu X'_\mu = 0 \end{cases} ; \tag{3.55}$$

where $\dot{X}^\mu \equiv \partial_\tau X^\mu$ and $X'^\mu \equiv \partial_\sigma X^\mu$.

It is useful to define light-cone coordinates on the worldsheet:

$$\sigma^\pm \equiv \tau \pm \sigma \Rightarrow ds^2_{ws} = -d\tau^2 + d\sigma^2 = -d\sigma^+ d\sigma^- ; \tag{3.56}$$

where the ws subscript stands for worldsheet.

In such coordinates, the equations of motion become:

$$\Box X^\mu = -4 \frac{\partial}{\partial \sigma^+} \frac{\partial}{\partial \sigma^-} X^\mu = 0 \therefore \partial_+ \partial_- X^\mu = 0 ; \tag{3.57}$$

where $\partial_\pm \equiv (\partial / \partial \sigma^\pm)$.

The most general solution to equation (3.57) is:

$$X^\mu = X^\mu_R(\sigma^-) + X^\mu_L(\sigma^+) ; \tag{3.58}$$

meaning that we decompose the string spacetime coordinates in right-moving modes, X^μ_R, and left-moving modes, X^μ_L.

From now on, we will proceed to the quantization of closed strings. There is an analogous procedure to the quantization of open strings with Neumann boundary conditions, see for instance [38] chapter 3. For a closed string, one can always rescale σ such that the periodicity, in $X^\mu(\tau,0) = X^\mu(\tau,l)$, becomes $l = 2\pi$, making σ correspond to an angle. Therefore the general expansions of right and left-moving modes are:

$$X^\mu_R(\sigma^-) = \frac{1}{2}x^\mu + \frac{\alpha'}{2}p^\mu\sigma^- + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\alpha^\mu_n e^{-in\sigma^-}\,; \tag{3.59}$$

$$X^\mu_L(\sigma^+) = \frac{1}{2}x^\mu + \frac{\alpha'}{2}p^\mu\sigma^+ + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\tilde{\alpha}^\mu_n e^{-in\sigma^+}\,; \tag{3.60}$$

where x^μ, p^μ, α^μ_n and $\tilde{\alpha}^\mu_n$ were defined such that the coefficients were the ones described above in each term for later convenience. Moreover, in order to have real left and right-moving modes, we must have x^μ, $p^\mu \in \mathbb{R}$ and $\alpha^\mu_{-n} = (\alpha^\mu_n)^*$ plus the same for $\tilde{\alpha}$s.

Rewriting the constraints of equation (3.55) in terms of light-cone coordinates one gets $\partial_+ X^\mu \partial_+ X_\mu = 0 = \partial_- X^\mu \partial_- X_\mu$. This leads us to:

$$\begin{aligned}\partial_- X^\mu = \partial_- X^\mu_R = \frac{\alpha'}{2}p^\mu + \sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\alpha^\mu_n e^{-in\sigma^-} \equiv \sqrt{\frac{\alpha'}{2}}\sum_{n}\alpha^\mu_n e^{-in\sigma^-} \Rightarrow \\ (\partial_- X^\mu)^2 = \frac{\alpha'}{2}\sum_{n,p}\alpha^\mu_n\alpha_{p\mu}e^{-i(n+p)\sigma^-} = \frac{\alpha'}{2}\sum_{n,m}\alpha^\mu_{m-n}\alpha_{n\mu}e^{-im\sigma^-} \equiv \alpha'\sum_m L_m e^{-im\sigma^-} = 0\,;\end{aligned} \tag{3.61}$$

where in the first line we've defined $\alpha^\mu_0 \equiv \sqrt{\frac{\alpha'}{2}}p^\mu$ in order to include the p^μ term into the sum. In the second line L_m were defined as:

$$L_m \equiv \frac{1}{2}\sum_n \alpha^\mu_{m-n}\alpha_{n\mu}\,; \tag{3.62}$$

and they vanish to satisfy the constraints.

If we repeat the same steps for the other constraint $(\partial_+ X^\mu)^2 = 0$ we would get analogous results replacing αs $\rightarrow \tilde{\alpha}$s. That means $\tilde{\alpha}^\mu_0 = \alpha^\mu_0 \equiv \sqrt{\frac{\alpha'}{2}}p^\mu$ and:

$$\tilde{L}_m \equiv \frac{1}{2}\sum_n \tilde{\alpha}^\mu_{m-n}\tilde{\alpha}_{n\mu} = 0 \tag{3.63}$$

One last comment is in order before we proceed to quantization. One could have proven that p^μ is the momentum of the center of mass of the string by showing it is the conserved charge due to spacetime translation symmetry, see [38] chapter 2. Therefore, it is reasonable to define the mass of the string as its 4-momentum squared (with a minus sign due to the metric signature) in analogy to point particles:

$$M^2 \equiv -p^\mu p_\mu = \begin{cases} -\frac{2}{\alpha'}\alpha^\mu_0\alpha_{0\mu} = \frac{2}{\alpha'}\sum_{n\neq 0}\alpha^\mu_{-n}\alpha_{n\mu} = \frac{4}{\alpha'}\sum_{n>0}\alpha^\mu_{-n}\alpha_{n\mu} \\ -\frac{2}{\alpha'}\tilde{\alpha}^\mu_0\tilde{\alpha}_{0\mu} = \frac{2}{\alpha'}\sum_{n\neq 0}\tilde{\alpha}^\mu_{-n}\tilde{\alpha}_{n\mu} = \frac{4}{\alpha'}\sum_{n>0}\tilde{\alpha}^\mu_{-n}\tilde{\alpha}_{n\mu} \end{cases} ;\ \text{or}$$
$$= -\frac{1}{\alpha'}(\alpha^\mu_0\alpha_{0\mu} + \tilde{\alpha}^\mu_0\tilde{\alpha}_{0\mu}) = \frac{2}{\alpha'}\sum_{n>0}(\alpha^\mu_{-n}\alpha_{n\mu} + \tilde{\alpha}^\mu_{-n}\tilde{\alpha}_{n\mu})\,; \tag{3.64}$$

where we have used the facts that $L_0 = 0 \Rightarrow \alpha^\mu_0\alpha_{0\mu} = \sum_{n\neq 0}\alpha^\mu_{-n}\alpha_{n\mu}$ and its analogous in terms of $\tilde{L}_0$, besides the definition of α^μ_0 and $\tilde{\alpha}^\mu_0$ in terms of p^μ.

One can see, from $L_0 = \tilde{L}_0 = 0$ and $\alpha^\mu_0 = \tilde{\alpha}^\mu_0$ or, equivalently, from the above equation that the following condition must be satisfied:

$$\sum_{n>0}\alpha^\mu_{-n}\alpha_{n\mu} = \sum_{n>0}\tilde{\alpha}^\mu_{-n}\tilde{\alpha}_{n\mu}\,; \tag{3.65}$$

this is the so-called level-matching condition (to be justified after quantization).

Now, it is finally time to quantize the theory. We will do so by introducing light-cone coordinates on spacetime too, not only on the worldsheet anymore:

$$X^\pm = \frac{X^0 \pm X^{D-1}}{\sqrt{2}} \Rightarrow ds^2_{st} = -2dX^+dX^- + dX^i dX_i\,;\ \text{where:}\, i = 1,\ldots,D-2\,; \tag{3.66}$$

the subscript st stands for spacetime.

Then, our coordinates become $X = \{X^+, X^-, X^i\}$. The important point to notice is that setting $\gamma_{ab} = \eta_{ab}$, the unit gauge, does not completely fix the gauge, one is still left with a residual gauge symmetry. This is analogous to the electromagnetic case. To completely fix the residual freedom we have, the following condition can be imposed:

$$X^+ = x^+ + \alpha' p^+ \tau = x^+ + \alpha' p^+ \left(\frac{\sigma^+ + \sigma^-}{2} \right) ; \tag{3.67}$$

this is the so-called light-cone gauge condition and it implies $\alpha_n^+ = \tilde{\alpha}_n^+ = 0$.

Note now that, adopting spacetime light-cone coordinates, the constraints coming from the equations of motion for γ_{ab}, namely $\partial_+ X^\mu \partial_+ X_\mu = 0 = \partial_- X^\mu \partial_- X_\mu$, become, from equation (3.66):

$$-2\partial_+ X^+ \partial_+ X^- + \partial_+ X^i \partial_+ X_i = 0 \Rightarrow \partial_+ X_L^- = \left(\frac{1}{\alpha' p^+} \right) \partial_+ X^i \partial_+ X_i \, ; \tag{3.68}$$

$$-2\partial_- X^+ \partial_- X^- + \partial_- X^i \partial_- X_i = 0 \Rightarrow \partial_- X_R^- = \left(\frac{1}{\alpha' p^+} \right) \partial_- X^i \partial_- X_i \tag{3.69}$$

Therefore all coefficients, except for x^- which is an integration constant, in the expansion of $X_{R/L}^-$ following equations (3.59) and (3.60) are determined by coefficients of other coordinates modes via integrations of the above equations. Doing such integrations gives us:

$$\frac{\alpha'}{2} p^- = \frac{\alpha'}{4p^+} p^i p_i + \frac{1}{2p^+} \sum_{n \neq 0} \alpha_{-n}^i \alpha_i^n = \frac{\alpha'}{4p^+} p^i p_i + \frac{1}{2p^+} \sum_{n \neq 0} \tilde{\alpha}_{-n}^i \tilde{\alpha}_i^n \, ; \tag{3.70}$$

$$\sqrt{\frac{\alpha'}{2}} \alpha_n^- = \frac{1}{2p^+} \sum_{n \neq m \neq 0} \alpha_{n-m}^i \alpha_i^m \, ; \tag{3.71}$$

$$\sqrt{\frac{\alpha'}{2}} \tilde{\alpha}_n^- = \frac{1}{2p^+} \sum_{n \neq m \neq 0} \tilde{\alpha}_{n-m}^i \tilde{\alpha}_i^m \tag{3.72}$$

Note that, using equation (3.70) one can rewrite M^2 just in terms of transverse modes, i.e. modes an i index which are transverse to the $+$ and $-$ directions.

$$M^2 \equiv -p^\mu p_\mu = 2p^+p^- - p^ip_i = p^ip_i + \frac{2}{\alpha'}\sum_{n\neq 0}\alpha^i_{-n}\alpha^n_i - p^ip_i = \frac{2}{\alpha'}\sum_{n\neq 0}\alpha^i_{-n}\alpha^n_i$$
$$= \frac{4}{\alpha'}\sum_{n>0}\alpha^i_{-n}\alpha^n_i = \frac{4}{\alpha'}\sum_{n>0}\tilde{\alpha}^i_{-n}\tilde{\alpha}^n_i = \frac{2}{\alpha'}\sum_{n>0}(\alpha^i_{-n}\alpha^n_i + \tilde{\alpha}^i_{-n}\tilde{\alpha}^n_i)\,; \tag{3.73}$$

where, in the last equality, we've rewritten the expression in order to make it straight forward to compare against the general result of equation (3.64).

Finally, we have presented enough information to discuss the quantization process. As usual, one has to define a momentum associated to each worldsheet field X^μ, write the canonical Poisson brackets and promote them to commutators via $\{\}_{PB} \to [,]/(i\hbar) = -i[,]$, in our units ($\hbar = 1$). One starts with the Polyakov action in the unit gauge, equation (3.46) with $\gamma^{ab} = \eta^{ab}$. That means the action and the momenta (density) can be written as:

$$S_P = \int d\tau L = \frac{1}{4\pi\alpha'}\iint d\tau d\sigma(\dot{X}^2 - X'^2) \Rightarrow P_\mu \equiv \frac{\delta L}{\delta \dot{X}^\mu} = \frac{\dot{X}_\mu}{2\pi\alpha'} \tag{3.74}$$

The equal-time canonical Poisson brackets and their commutators, i.e. their quantum versions, are:

$$\{X^\mu(\tau,\sigma), P_\nu(\tau,\sigma')\}_{PB} \equiv \int d\sigma''\left[\frac{\delta X^\mu(\sigma)}{\delta X^\rho(\sigma'')}\frac{\delta P_\nu(\sigma')}{\delta P_\rho(\sigma'')} - \frac{\delta P_\nu(\sigma')}{\delta X_\rho(\sigma'')}\frac{\delta X^\mu(\sigma)}{\delta P^\rho(\sigma'')}\right] = \delta(\sigma-\sigma')\delta^\mu{}_\nu$$
$$\Rightarrow \{X^\mu(\tau,\sigma), P^\nu(\tau,\sigma')\}_{PB} = \delta(\sigma-\sigma')\eta^{\mu\nu} \to [X^\mu(\tau,\sigma), P^\nu(\tau,\sigma')] = i\delta(\sigma-\sigma')\eta^{\mu\nu}\,; \tag{3.75}$$
$$\{X^\mu(\tau,\sigma), X^\nu(\tau,\sigma')\}_{PB} = \{P^\mu(\tau,\sigma), P^\nu(\tau,\sigma')\}_{PB} = 0$$
$$\to [X^\mu(\tau,\sigma), X^\nu(\tau,\sigma')] = [P^\mu(\tau,\sigma), P^\nu(\tau,\sigma')] = 0 \tag{3.76}$$

From equation (3.75) plus the clever rewriting of x^μ and p^μ, which can be seen from summing left and right-moving modes in equations (3.59) and (3.60) and then taking a τ-derivative, as:

$$x^\mu = \frac{1}{2\pi}\int_0^{2\pi} d\sigma X^\mu \ ; \ p^\mu = \frac{1}{2\pi\alpha'}\int_0^{2\pi} d\sigma \dot{X}^\mu = \int_0^{2\pi} d\sigma P^\mu \,; \tag{3.77}$$

one gets:

$$[x^\mu, p^\nu] = \frac{1}{2\pi}\iint d\sigma d\sigma' [X^\mu(\tau,\sigma), P^\nu(\tau,\sigma')] = \frac{i\eta^{\mu\nu}}{2\pi}\int_0^{2\pi} d\sigma \int_0^{2\pi} d\sigma' \delta(\sigma-\sigma') = i\eta^{\mu\nu} \tag{3.78}$$

Similarly, but with the implementation of a smart trick to get rid of mixing αs and $\tilde{\alpha}$s when writing their commutators in terms of the ones in equations (3.75) and (3.76), see [50] chapter 13, one finds the following commutation relations:

$$[\alpha^\mu_m, \alpha^\nu_n] = [\tilde{\alpha}^\mu_m, \tilde{\alpha}^\nu_n] = m\eta^{\mu\nu}\delta_{m+n,0}\,; \tag{3.79}$$

$$[\alpha^\mu_m, \tilde{\alpha}^\nu_n] = 0 \tag{3.80}$$

The above commutation relations looks like harmonic oscillators relations. However, the m in the right-hand side of equation (3.79) is unusual. That is not a problem though, since we can redefine operators $a^\mu_m \equiv \alpha^\mu_m/\sqrt{m}$ and $a^{\dagger\mu}_m \equiv \alpha^\mu_{-m}/\sqrt{m}$ (for $m > 0$) in order to generate the usual commutation relations.

Note that $[\alpha^0_m, \alpha^0_n]$, likewise for $\tilde{\alpha}s$, has a negative sign, leading to states with negative norm. This situation is analogous to that we encounter when trying to covariantly quantize electromagnetism. In both situations, it means we have more modes than needed, i.e. some of them must be unphysical. Equations (3.70), (3.71), (3.72) and (3.73) summarize the usefulness of light-cone coordinates and the light-cone gauge condition to quantization. It removes unphysical degrees of freedom by setting $\alpha^+_n = \tilde{\alpha}^+_n = 0$ and making the minus-direction modes dependent on the i-direction ones. All in all, one ends up with only α^i_n and

$\tilde{\alpha}_n^i$ as independent modes and their commutation relations are not problematic, see equation (3.79).

It turns out that equation (3.73) is only valid classically, i.e. when $\hbar \to 0$ and all modes commute. The correct expression, due to the fact α_n^i and α_{-n}^i do not commute, at the quantum level is, see [38]:

$$\begin{aligned} M^2 &= \frac{4}{\alpha'}\left(\sum_{n>0}\alpha_{-n}^i\alpha_i^n - a\right) = \frac{4}{\alpha'}\left(\sum_{n>0}\tilde{\alpha}_{-n}^i\tilde{\alpha}_i^n - a\right) \\ &= \frac{2}{\alpha'}\left[\sum_{n>0}(\alpha_{-n}^i\alpha_i^n + \tilde{\alpha}_{-n}^i\tilde{\alpha}_i^n) - 2a\right] \end{aligned} \tag{3.81}$$

Note that we can define:

$$N \equiv \sum_{n>0}\alpha_{-n}^i\alpha_i^n = \sum_{n>0} n a_n^{\dagger i} a_i^n\,; \tag{3.82}$$

$$\tilde{N} \equiv \sum_{n>0}\tilde{\alpha}_{-n}^i\tilde{\alpha}_i^n = \sum_{n>0} n \tilde{a}_n^{\dagger i} \tilde{a}_i^n\,; \tag{3.83}$$

where $N = \tilde{N}$. This is the previous mentioned level matching condition. Notice N and $\tilde{N}$ are kind of number operator, but there is an extra factor of n inside the sums which is called level, justifying the name of the condition.

The constant a in equation (3.81) is due to normal ordering. It can be shown it is related to the number of spacetime dimensions through $a = (D-2)/24$, see [12]. It turns out that, for consistency arguments which are easier to realize in the open string case and therefore we are not giving details here, a has to be equal to 1. That means, the bosonic string theory lives in $D = 26$ spacetime dimensions! This conclusion is also valid to closed strings.

String theory contains an infinite number of particles. Each kind of particle is characterized by its level n and one constructs a general state by applying the creation operators $\alpha_{-n}^i \propto a_n^{\dagger i}$, and their left-moving modes counterparts, to the vacuum state $|0, \vec{p}\rangle$, where $\vec{p}$

stands for the momentum eigenvalue of the vacuum. Thus, the vacuum is defined as the state which is annihilated by any annihilation operator $a_n^i \left|0, \vec{p}\right\rangle = \tilde{a}_n^i \left|0, \vec{p}\right\rangle = 0$, as usual. It can be seen then, from equation (3.81), that the vacuum has $N = \tilde{N} = 0$ and therefore $M^2 = -4/\alpha'$, i.e. it is a tachyon. That means it is unstable, a problem which is solved with supersymmetry when one considers superstrings instead of bosonic ones.

Let us analyze the first excited closed string states. Since we must match levels, $N = \tilde{N}$, this state is $a_1^{\dagger i} a_1^{\dagger j} \left|0, \vec{p}\right\rangle \equiv \left|ij\right\rangle$ and it a massless state, $M^2 = (2/\alpha')(N + \tilde{N} - 2) = 0$, which has two transverse (to the plus and minus directions) indices. Breaking it into irreducible representations of the Lorentz group we end up with an antisymmetric state plus a trace part and a symmetric traceless state. The theory contains a massless symmetric traceless state which is identified with the graviton, see [51] for an argument of why a particle with such characteristics must satisfy the Einstein equations. Therefore, string theory is a quantum theory of gravity! The antisymmetric state is called the B-tensor and the scalar is the dilaton. Together with the graviton, they condense and form a background for the string, modifying the Polyakov action, see [12] chapter 17.

3.3.1 D-Branes

Most of the last section we have only considered closed strings. However, as we saw, string theory also contains open strings which may appear in two, a priori different, fashions. Neumann boundary conditions are consistent from the beginning but Dirichlet ones may seem weird since fixing the ends of a string to a spacetime point would break translational symmetry. This strange feature was overcome when it was realized the ends of strings are attached to objects called Dirichlet-branes, or more commonly D-branes, which themselves are dynamical and thus do not break translational invariance [52].

The name D-branes come from a generalization of the word membrane, which is a two-dimensional object in space, i.e. it has a three dimensional worldvolume (analogous to the worldsheet for a string), to any number of space dimensions they span. More precisely, they should be called Dp-branes, where p is their spatial (not spacetime) dimensionality. For instance, a membrane is a D2-brane while the string is a D1-brane and a particle can be called a D0-brane.

Consider an open string which obeys Neumann boundary conditions in $p+1$ spacetime directions and Dirichlet conditions in the other $D-p-1$ dimensions. Then, the ends of the string would be attached to a $p+1$ dimensional wall, a Dp-brane. Following this ideia, we can think of an open string which obeys Neumann boundary conditions in all 26 directions as ending up in a D25-brane, rather than having free endpoints. This way, all types of open strings are explained in the same fashion and the difference, coming from their boundary conditions, will only affect the dimensionality of the wall on which they end. Note that we have been considering open strings with Neumann boundary conditions in the time dimension. The weird case of Dirichlet boundary conditions in the time direction, the endpoints of a string would be fixed in a moment of time, is relevant though, but it will not be analyzed in this thesis.

Since a generic worldvolume is now $p+1$ dimensional, instead of the two-dimensional worldsheet of strings, we will use spacetime light-cone coordinates $X = (X^+, X^-, X^a, X^i)$ such that $\{X^+, X^-, X^a\}$ are $p+1$ coordinates which are parallel to the worldvolume and $\{X^i\}$ are $D-p-1$ transverse coordinates, i.e. coordinates which obey Dirichlet boundary conditions. In analogy to the point-particle and the string action, one could guess the form of the D-brane action is such that its worldvolume is minimized (or maximized). Namely:

$$S = -T_p \int d^{p+1}\xi \sqrt{-\det(h_{ab})}\,; \tag{3.84}$$

where T_p is a generalization of the string tension, now energy per unit spatial volume of the D-brane, ξ^a are worldvolume coordinates and a, b runs through all of them.

The induced metric h_{ab} is defined in the same way as the one in the Nambu-Goto action, i.e. it is the induced metric from the spacetime one $h_{ab} = \partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu}$ (for flat spacetime). But since it was pointed out before, the massless closed string states condensate and form a background which is obtained by $\eta_{\mu\nu} \rightarrow g_{\mu\nu}(X) + \alpha' B_{\mu\nu}(X)$ in the expression for the induced metric, where $g_{\mu\nu}$ is the full curved metric (including the graviton) and $B_{\mu\nu}$ is the antisymmetric B-tensor. The coupling of the D-brane with the dilaton, ϕ, comes from a longer argument which is not going to be done here, see [12] chapter 19. All in all, the changes in equation (3.84) lead us to:

$$S = -T_p \int d^{p+1}\xi e^{-\phi} \sqrt{-\det \left[\partial_a X^\mu \partial_b X^\nu (g_{\mu\nu} + \alpha' B_{\mu\nu})\right]} \tag{3.85}$$

We are working in the static gauge. This means:

$$X^+ = \xi^+ \ ; \ X^- = \xi^- \ ; \ X^a = \xi^a \tag{3.86}$$

plus the identification of coordinates transverse to the D-brane's worldvolume with scalar fields through the following rescaling:

$$X^i(\xi^a) = \frac{\phi^i(\xi^a)}{\sqrt{T_p}} \ ; \tag{3.87}$$

where the rescaling constant was chosen so that we get the usual kinetic term for a scalar when expanding the D-brane action in a particular way [12].

Were we to include the effect of massless open string modes on the D-brane action, equation (3.85), we would need to consider the massless vector, A_a, one finds when quantizing the open string with Neumann boundary conditions. This was not shown here since we focused on the quantization of closed strings only. Anyway, the way to covariantly include such field is through its strength tensor, F_{ab}. More precisely, to include it inside the determinant one follows the Born-Infeld action for electromagnetism. All in all, including the massless closed string states plus A_a we arrive at the Dirac-Born-Infeld action, see [12]:

$$S_{DBI} = -T_p \int d^{p+1}\xi e^{-\phi} \sqrt{-\det \left[\partial_a X^\mu \partial_b X^\nu (g_{\mu\nu} + \alpha' B_{\mu\nu}) + \alpha' F_{ab}\right]} \tag{3.88}$$

The existence of D-branes changes the open string states in a relevant way. First, notice that open strings might have both ends attached to the same D-brane or to two different D-branes. Therefore, considering N possible D-branes for generality, in order to identify at which D-brane each endpoint of the string ends one must introduce the so-called Chan-Paton factors, ij, to open string states. Meaning that the string has one of its endpoints on the ith D-brane and the other on the jth one. D-branes also break the $SO(1, D-1)$ group into $SO(1,p) \times SO(D-p-1)$, meaning that, from the point of view of the D-brane worldvolume, $\alpha^a_{-1} |ground, ij\rangle$ are vectors while transverse excitations $\alpha^i_{-1} |ground, ij\rangle$ are scalars. $|ground, ij\rangle$ is the open string ground state whose dependence on Chan-Paton factors was made explicit but also depends on other quantities which are not relevant to our point and were thereby hidden, avoiding a notational overkill. The point is the worldvolume vectors are gauge fields and the Chan-Paton factors, more precisely whether the string has both endpoints on the same D-brane or not, can be shown to determine the group of the gauge theory. If both ends of the string are attached to the same D-brane, $i = j$, the group is $U(N)$, recall that N is the total number of D-branes available to the string's endpoint to end on, while if $i \neq j$ the group becomes $U(1)^N$, see [12] chapter 19.

The fact gauge theories can be said to live on the D-branes, since the gauge fields are parallel to the D-brane excitations, may give us an interesting idea of how to get $3+1$ dimensional particle physics theories out of string theory. Since the Standard Model of particle physics is a mostly composed of a $SU(3) \times SU(2) \times U(1)$ gauge theory, one might suppose Standard Model interactions are constrained to a D3-brane, see for example [44].

In summary, the picture we end up with is that string theory contains both closed and open strings, see figure 3.1. Open strings end on dynamical objects called Dp-branes, whose spatial dimensionality p is determined by the number of dimensions in which the open string obeys Neumann boundary conditions. String excitations which are transverse to the D-brane are scalars from the D-brane's worldvolume viewpoint while parallel excitations are nonabelian gauge fields. Thus, one can say that the gauge theory lives on the D-brane.

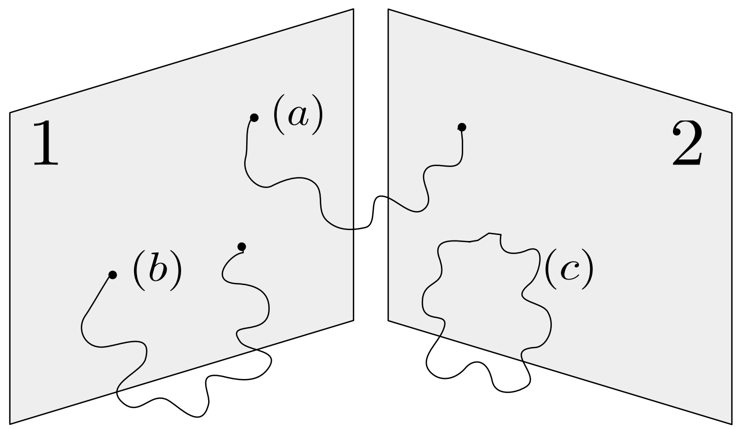

Figure 3.1: Representation of different configurations fundamental strings can assume. There are open strings with ends attached to different D-branes, string (a) in the figure, open strings with both ends on the same D-brane, string (b), and closed strings detached from any D-brane, as in (c). Although D-branes 1 and 2 are both D2-branes, this is not necessarily the case, their dimensionalities were chosen for illustrational purposes.

3.4 Dualities

So far, we have seen many deep consequences of considering the most elementary objects of nature as being one-dimensional strings rather than point-particles. As we have seen, symmetries of the string action play an important role in that. However, string theory has other unusual and interesting features that may seem initially like ordinary symmetries but are actually duality symmetries. Duality is a somewhat deeper symmetry that allows us to describe a system in two a priori unrelated yet completely equivalent ways. Gauge-gravity and T-duality will be the relevant ones for this thesis. But since we still need more information to understand gauge-gravity duality as a generalization of the AdS/CFT correspondence, we will leave that to the end of the chapter and start by describing T-duality. It is important to point out these two are not the only known string theory dualities. There is at least the extra S-duality, but it isn't relevant to this work and thus it will not be described here.

3.4.1 T-Duality

T-duality is a feature of both open and closed strings. Nevertheless, only T-duality of closed strings will be discussed here, see [12] chapter 21 for an introduction to T-duality of open strings. The relevant concepts will already be introduced in the closed string case.

Let us start by considering one compactified dimension, the 25th direction for concreteness, out of the 26 spacetime dimensions in the bosonic string theory. For simplicity, let us assume the background spacetime is $M_{25} \times S^1$, where M_{25} is a 25-dimensional Minkowski space and S^1 is the circle, as usual. We are considering the same type of spacetime as in the Kaluza-Klein original theory which was previously discussed. There are just 21 extra large dimensions that won't change our discussion. The fact that the 25th direction is a circle means we must identify by an equivalence relation $X^{25} \sim X^{25} + m(2\pi R)$, where R is the radius of the circle and $m \in \mathbb{Z}$ is the so-called winding number. That means our previous closed string boundary condition $X^\mu(\tau, \sigma + 2\pi) = X^\mu(\tau, \sigma)$ gets changed, only in the 25th direction, to:

$$X^{25}(\tau, \sigma + 2\pi) = X^{25}(\tau, \sigma) + 2\pi m R \sim X^{25}(\tau, \sigma) \tag{3.89}$$

The equation above means that the string winds around the 25th dimension m times. It will be useful to define the winding as:

$$w \equiv \frac{mR}{\alpha'} \tag{3.90}$$

As we have seen, using worldsheet light-cone coordinates leads us to equations of motion whose solutions are a sum of left and right-moving modes $X^\mu_{L/R}$. Recall that $\sigma^\pm \equiv \tau \pm \sigma$ leading to $\sigma \to \sigma + 2\pi \Rightarrow \sigma^\pm \to \sigma^\pm \pm 2\pi$. This allows us to rewrite equation (3.89) as:

$$\begin{aligned} X^{25}_R(\sigma^- - 2\pi) + X^{25}_L(\sigma^+ + 2\pi) &= X^{25}_R(\sigma^-) + X^{25}_L(\sigma^+) + 2\pi m R \Rightarrow \\ X^{25}_R(\sigma^- - 2\pi) - X^{25}_R(\sigma^-) &= X^{25}_L(\sigma^+) - X^{25}_L(\sigma^+ + 2\pi) + 2\pi\alpha' w\,; \end{aligned} \tag{3.91}$$

where in the last line the definition of equation (3.90) was applied.

Note that we have periodicity of both left and right moving modes since the left hand side of the above equation only depends on σ^- while the right hand side of it only depends on σ^+. We will have expressions similar to those of equations (3.59) and (3.60).

$$X_R^{25}(\sigma^-) = x_R^{25} + \sqrt{\frac{\alpha'}{2}}\alpha_0^{25}\sigma^- + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\alpha_n^{25}e^{-in\sigma^-} \; ; \tag{3.92}$$

$$X_L^{25}(\sigma^+) = x_L^{25} + \sqrt{\frac{\alpha'}{2}}\tilde{\alpha}_0^{25}\sigma^+ + i\sqrt{\frac{\alpha'}{2}}\sum_{n\neq 0}\frac{1}{n}\tilde{\alpha}_n^{25}e^{-in\sigma^+} \tag{3.93}$$

Now, what is changed due to the winding of the string is that instead of having $\alpha_0^{25} = \tilde{\alpha}_0^{25}$, like we have for other coordinates, we have $\tilde{\alpha}_0^{25} - \alpha_0^{25} = \sqrt{2\alpha'}w$. This can be checked by substituting equations (3.92) and (3.93) into (3.91).

Recall that, from equation (3.74), the momenta are $P^\mu = \dot{X}^\mu/(2\pi\alpha')$. These are actually momenta per unit σ, i.e. length of the string. Therefore, the integrated momentum in the 25th direction is:

$$\begin{aligned} p \equiv p^{25} &= \int_0^{2\pi} d\sigma P^{25} = \frac{1}{2\pi\alpha'}\int_0^{2\pi} d\sigma(\dot{X}_R^{25} + \dot{X}_L^{25}) = \frac{1}{2\pi\alpha'}\int_0^{2\pi} d\sigma(\partial_- X_R^{25} + \partial_+ X_L^{25}) \\ &= \frac{1}{2\pi\alpha'}\sqrt{\frac{\alpha'}{2}}\left[2\pi(\alpha_0^{25} + \tilde{\alpha}_0^{25})\right] = \frac{\alpha_0^{25} + \tilde{\alpha}_0^{25}}{\sqrt{2\alpha'}} \end{aligned} \tag{3.94}$$

So we can solve α_0^{25} and $\tilde{\alpha}_0^{25}$ for p and w:

$$\tilde{\alpha}_0^{25} - \alpha_0^{25} = \sqrt{2\alpha'}w \; ; \text{and } p = \frac{\alpha_0^{25} + \tilde{\alpha}_0^{25}}{\sqrt{2\alpha'}} \Rightarrow \alpha_0^{25} = \sqrt{\frac{\alpha'}{2}}(p - w) \; ; \; \tilde{\alpha}_0^{25} = \sqrt{\frac{\alpha'}{2}}(p + w) \tag{3.95}$$

Moreover, in order for e^{ipx} to be single-valued in the 25th direction, due to the identification (3.89), we must have:

$$e^{i(2\pi R)p} = e^0 = 1 \Rightarrow p = \frac{n}{R} \; ; \; n \in \mathbb{Z} \tag{3.96}$$

Our goal is to compute the spectrum of the closed string, i.e. M^2, in our case to compare against equation (3.81) and see what changes when we consider one compactified dimension. From the point of view of the large dimensions, using spacetime light-cone coordinates as before, the new mass spectrum is $M^2 = 2p^+p^- - p^I p_I$, where I does not include the compactified dimension. In our notation $X = (X^+, X^-, X^i) = (X^+, X^-, X^I, X^{25})$. Thus, considering the light-cone gauge condition $X^+ = x^+ + \alpha' p^+ \tau$, we proceed to find out an expression for p^- in terms of other directions' modes in the exact same way it was done in equation (3.70). The obtained expression is:

$$\begin{aligned} p^- &= \frac{1}{2p^+} p^I p_I + \frac{1}{\alpha' p^+} (\alpha_0^{25})^2 + \frac{1}{\alpha' p^+} \sum_{n \neq 0} \alpha^i_{-n} \alpha^n_i \\ &= \frac{1}{2p^+} p^I p_I + \frac{1}{\alpha' p^+} (\alpha_0^{25})^2 + \frac{2}{\alpha' p^+} \left(\sum_{n>0} \alpha^i_{-n} \alpha^n_i - 1 \right) \\ &= \frac{1}{2p^+} p^I p_I + \frac{1}{\alpha' p^+} (\alpha_0^{25})^2 + \frac{2}{\alpha' p^+} (N - 1) \, ; \end{aligned} \tag{3.97}$$

where in the second line we have broken the sum $\sum_{n \neq 0} \to \sum_{n>0} + \sum_{n<0}$ then redefined $n \to -n$ in the second sum and regrouped them both together obtaining a factor of 2 in front of it. The constant $a = 1$ subtracting the final sum is due to normal ordering, i.e. we must commute α^n_i past α^i_{-n} in the second sum after the redefinition $n \to -n$ in order to be able to recombine the two sums. In the last line the definition (3.82) was used.

Equation (3.97) was obtained by integrating equation (3.69). The analogous expression one finds by integrating equation (3.68) instead is:

$$p^- = \frac{1}{2p^+} p^I p_I + \frac{1}{\alpha' p^+} (\tilde{\alpha}_0^{25})^2 + \frac{2}{\alpha' p^+} (\tilde{N} - 1) \, ; \tag{3.98}$$

where the definition (3.83) was used to make the expression shorter.

Subtracting equation (3.98) from (3.97) one obtains a modification of the level matching condition:

$$N - \tilde{N} = -\frac{1}{2}\left[(\alpha_0^{25})^2 - (\tilde{\alpha}_0^{25})^2\right] = \alpha' p w \tag{3.99}$$

Finally, we are able to find an expression for M^2 which is easy to compare against equation (3.81):

$$\begin{aligned} M^2 &= 2p^+p^- - p^I p_I = p^I p_I + \frac{2}{\alpha'}(\alpha_0^{25})^2 + \frac{4}{\alpha'}(N-1) - p^I p_I \\ &= \frac{2}{\alpha'}(\alpha_0^{25})^2 + \frac{2}{\alpha'}(N + \tilde{N} + \alpha' p w - 2) = \frac{2}{\alpha'}\left[\frac{\alpha'}{2}(p^2 + w^2) + N + \tilde{N} - 2\right] \\ &= p^2 + w^2 + \frac{2}{\alpha'}(N + \tilde{N} - 2) = \left(\frac{n}{R}\right)^2 + \left(\frac{mR}{\alpha'}\right)^2 + \frac{2}{\alpha'}(N + \tilde{N} - 2)\,; \end{aligned} \tag{3.100}$$

where in the second equality and third equalities equations (3.97) and (3.99) were used, respectively. In the last equality we have plugged in the definitions (3.90) and (3.96).

The last term in the spectrum expression was the same we have already obtained in the case of no compact dimensions, see equation (3.81). On the other hand, the two first terms are due to the identification (3.89). Notice the spectrum remains the same if we do the following transformation:

$$R \to \tilde{R} \equiv \frac{\alpha'}{R}\ ; \text{ and } m \leftrightarrow n\,; \tag{3.101}$$

meaning that $M^2(R, n, m) = M^2(\tilde{R}, m, n)$.

At first sight, this seems just like an unexpected symmetry of the closed string spectrum when one considers a compactified dimension. However, it can be shown that it is actually a hint of an equivalence between two string theory descriptions, see [12] chapters 20 and 21, [50] chapter 17 or [53] chapter 8 for a more comprehensive discussion. One might think of the string length as the radius at which the two descriptions are the same, i.e. a self-dual description, this happens when $R = \tilde{R} = \sqrt{\alpha'} \equiv l_s$. A string theory whose compact dimension radius is lower than this threshold can always be interchanged with another description whose radius is $> l_s$. Point-particle theories don't have this symmetry between two apparently unrelated descriptions, therefore it is hard to make sense of lengths less than Planck length, for instance. This is not much of a problem in string theory since every time a complication of this kind arises in understanding physical theories one can work in the dual description where the distance will be greater than Planck length so the theory should be well-behaved. It should be pointed out that two dual descriptions are indistinguishable in the sense that their predictions are all the same. So there should not be any way, in principle, to choose a preferred description.

Dualities should have a duality map or dictionary, translating quantities in one description to their counterparts in the dual description. T-duality is no exception to the rule and one could show it is basically implemented through $X^{25} = X_L^{25} + X_R^{25} \to \tilde{X}^{25} = X_L^{25} - X_R^{25}$, see [50].

3.5 Superstring Theory

Up until now all the excitations of strings we have described represent bosonic particles. Nevertheless it is known this is only one of the two kinds of particle composing our universe. Fermionic modes are obtained with the implementation of supersymmetry within string theory, i.e. superstring theory. Furthermore, the ground state of bosonic strings is tachyonic (at least for closed strings as it was previously shown but maybe also for open strings, see [12] chapter 19), thus perturbatively unstable. Superstrings don't have any tachyon on their spectrum.

It turns out there are several ways of implementing supersymmetry in string theory, see [12] chapter 20. One can start with worldsheet supersymmetry only in the Neveu-Schwarz-Ramond (NSR) formulation or with spacetime SUSY in the Green-Schwarz formulation. There is also the pure spinor or Berkovits formulation which also has spacetime supersymmetry. In all formulations the conclusion is the same and, analogously to what have been found for the bosonic string, the theory itself demands a certain number of spacetime dimensions for consistency. Superstrings live on $D = 10$ dimensions rather than 26 of the bosonic case.

Historically, five different superstring theories were found. They differ by how one can implement fermions to the worldsheet. If we add fermions both in the right and left moving sectors of the string we end up with a type II superstring theory. Instead, if we add fermions just to one sector we end up with a heterotic theory. In the type II case one has $\mathcal{N} = 2$ supersymmetries leading to two fermions. Those two can have opposite (type IIA) or the same chirality (type IIB). The heterotic string theory is also split into two: the heterotic $SO(32)$ and the heterotic $E_8 \times E_8$ theories. They are named after the Yang-Mills groups they give rise in 10 dimensions. Furthermore, there is yet another superstring theory called type I which contains both open and closed strings. It is important to point out that the massless bosonic closed string states (the graviton associated with $g_{\mu\nu}$, the antisymmetry B-field $B_{\mu\nu}$ and the dilaton ϕ) which would condensate and form a background for the string are also present in all superstring theories.

The fact that we have several kinds of superstring theories instead of just one as in the bosonic case may seem like a setback to the original excitement of trying to construct a rigid, i.e. non-adjustable, theory of Nature. This setback was overcome by the realizations of Witten in 1995 [54], see [12] chapter 21 to a brief, yet more detailed than here, discussion. He was able to relate the different types of superstring theories using S-duality and T-duality, so that one theory in the perturbative limit is equivalent to another in a non-perturbative regime. All in all, the different types of string theories are now seen as different corners of the same underlying theory, called M-theory for unknown reasons.

D-branes are also present in superstring theories, but now they are supersymmetric themselves. A very important observation that will be used when introducing the AdS/CFT correspondence is that, since D-branes have gauge theories living on them, when one considers a D3-brane the gauge theory manifested on it is the previously highlighted $\mathcal{N} = 4$ Super Yang-Mills in 4 dimensions. This is so since $\mathcal{N} = 4$ is the maximum supersymmetry one can have in 4 dimensions for spins ≤ 1 and, as we know, the gauge theory with $\mathcal{N} = 4$

supersymmetries and canonical kinetic terms is unique, $\mathcal{N} = 4$ SYM theory.

Two more points must be added to this brief discussion on supersymmetric strings in order to introduce later needed ideas: SUGRA as superstring theories low energy limit and the problems with dimensional reduction of higher dimensional string theories.

3.5.1 Low Energy Limit: Supergravity

One of the things which make string theory so promising is the absence of adjustable dimensional parameters. The unique dimensional parameter we have encountered, which just sets the string scale as we saw in the T-duality discussion, is the so-called Regge slope α', or simply alpha prime. Therefore, it would be interesting to look for low energy limits, $E \ll (\alpha')^{-1/2}$, of superstring theories since one wishes to make contact with experiments. It turns out it is very complicated to reproduce the Standard Model as we know it, but one finds out that the lower energy theories are supersymmetric and involve the previously mentioned massless states, including $g_{\mu\nu}$ which is the relevant one to our point, thus they are a supergravity theories. We will return to this observation shortly when the original AdS/CFT setup is introduced. In this context the relevant supergravity theory will be the type IIB one, i.e. the low energy limit of type IIB superstring theory, which contains even form potentials as massless Ramond-Ramond (R-R) modes: a pseudo-scalar with field strength $F_{(1)}$ coupled to a D-instanton or D(-1)-brane (Dirichlet boundary conditions even in the time direction), a 2-form $A_{(2)}$ with field strength $F_{(3)}$ coupled to D1-branes and a 4-form $A_{(4)}$ with field strength $F_{(5)}$, see [12] chapter 21 for a proper introduction to types IIA and IIB supergravities.

3.5.2 The Problem of Moduli Stabilization

String theory is a candidate for theory of the real world even at enormously high energies. But since one wishes to make contact with experiment to prove the theory, it is natural that one considers the compactification of 6 out of the 10 dimensions of superstring theories, leaving us with only the usual 4 observable ones. The process of compactification by itself generates scalar fields in the dimensionally reduced 4-dimensional theories. We have already seen in detail an example of that when the original 5-dimensional Kaluza-Klein theory was reduced giving rise to a free massless scalar field, see appendix D. Such scalar fields without potential, at least perturbatively, are called moduli and they are common in compactified string theories.

Moduli are a problem to the effective 4-dimensional theories since they would represent a fifth force which is constrained by experiments, at least on Earth. In chapter 6 we will see the possibility of having such scalars on cosmological scales without violating current experimental bounds through the chameleon mechanism. Nevertheless, in string theory one usually has to stabilize the moduli, i.e. find a way to give them a non-perturbative potential. Since string theory is currently only defined perturbatively, it is needless to say this is a hard task. It will be seen that T-duality is crucial to stabilize moduli in string gas cosmology, chapter 4, while chameleon scalars are a good alternative to having to stabilize all moduli of a theory.

3.6 AdS/CFT and Gauge/Gravity Duality

To end this chapter's discussion another duality will be introduced, namely gauge-gravity duality. The issue started as the so-called holographic principle inspired by the black hole information paradox, see [55] for a review on the subject. Eventually, people started applying this principle to many areas of physics, including cosmology, and nowadays even books were written with further developments of it, see for instance [48]. The first more concrete realization of the principle came through the work of Juan Maldacena [56]. The main features of Maldacena's paper will be reviewed to later generalize it to gauge-gravity duality. First, we need to briefly introduce conformal field theories, which compose one side of AdS/CFT correspondence.

3.6.1 Conformal Field Theories (CFT)

First, we must introduce what a conformal transformation is. A conformal transformation is a generalization of scale transformations $ds'^2 = \lambda^2 ds^2$, where λ is a constant, in flat space. The generalization is done by replacing $\lambda \to \Omega(x)$, a spacetime function.

$$ds^2 = dx^\mu dx_\mu \to ds'^2 = dx'^\mu dx'_\mu = \Omega(x)^2 ds^2\,; \tag{3.102}$$

where Ω^2 is called a conformal factor.

Scale-invariant theories at the quantum level are usually also conformal invariant. Note that scale invariance is a global symmetry, thus there are some scale-invariant theories at the classical level which are not invariant at the quantum level. That is true since when we renormalize the theory a regularization procedure has to be considered and the regularization introduces a scale to it, e.g. the dimensional transmutation μ in dimensional regularization or the cutoff Λ in cut-off regularization, see [48] chapter 8. Moreover, since scale invariance is not a local symmetry, there is no problem in breaking it at the quantum level. On the other hand, conformal symmetry should not be broken at the quantum level since it is local, just like gauge symmetries, and breaking it would give rise to an inconsistency in the theory. All in all, one must check whether the scale invariance is valid quantum mechanically or not before trying to associate it with conformal symmetry.

Our interest here will be to analyze $d > 2$ dimensional theories. In such case the conformal group is finite dimensional. An infinitesimal conformal transformation is given by $x'^\mu = x^\mu + v^\mu(x)$ and $\Omega(x) = 1 + \sigma_v(x)$. Thus:

$$\begin{gathered} x'^\mu = x^\mu + v^\mu(x) \therefore dx'^\mu = dx^\mu + \partial_\nu v^\mu dx^\nu \;\; ; \text{and} \;\; \Omega(x) = 1 + \sigma_v(x) \therefore \Omega^2 = 1 + 2\sigma_v \Rightarrow \\ ds'^2 = dx'^\mu dx'_\mu = dx^\mu dx_\mu + 2\partial_\nu v_\mu dx^\mu dx^\nu = (1 + 2\sigma_v) dx^\mu dx^\nu \Rightarrow \\ 2\partial_{(\nu} v_{\mu)} = \partial_\nu v_\mu + \partial_\mu v_\nu = 2\sigma_v \eta_{\mu\nu} \therefore \sigma_v(x) = \frac{\partial_\mu v^\mu}{d} \end{gathered} \tag{3.103}$$

For $d > 2$ dimensions the equation above has as a most general solution:

$$v_\mu(x) = a_\mu + \omega_{\mu\nu}x^\nu + \lambda x_\mu + b_\mu x^2 - 2x_\mu b_\nu x^\nu \, ; \tag{3.104}$$

where $\omega_{\mu\nu} = -\omega_{\nu\mu}$ and σ_v is found to be:

$$\begin{aligned}\sigma_v &= \frac{1}{d}\left(\omega_{\mu\nu}\partial^\mu x^\nu + \lambda\partial^\mu x_\mu + 2b_\mu x_\nu \partial^\mu x^\nu - 2\partial^\mu x_\mu b_\nu x^\nu - 2x_\mu b_\nu \partial^\mu x^\nu\right) \\ &= \lambda - 2b_\mu x^\mu + \frac{1}{d}\left(\omega_{\mu\nu} + 2b_\mu x_\nu - 2b_\nu x_\mu\right)\eta^{\mu\nu} = \lambda - 2b_\mu x^\mu\end{aligned} \tag{3.105}$$

So, the conformal transformation contains a scale transformation λ, a translation a_μ, a rotation $\omega_{\mu\nu}$ and a different type of transformation called special conformal transformation b_μ. It can be shown that their generators can be put into the form of $d+2$-dimensional anti-symmetric matrices which are generators of an $SO(2,d)$ group, see [48]. It will be important to keep in mind that this is the group of conformal transformations in d dimensions ($d > 2$).

The other important fact to AdS/CFT is that the previously introduced $\mathcal{N} = 4$ Super Yang-Mills theory in $d = 4$ dimensions is conformal invariant. Thus, being a 4-dimensional conformal field theory, its conformal group is $SO(2,4)$.

3.6.2 The Original Realization by Maldacena

Finally, we are able to introduce the famous anti-de Sitter/conformal field theory (AdS/CFT) correspondence found by Maldacena [56]. The first indications of a duality are found analyzing the isometry group of a d-dimensional Anti-de Sitter space. As anticipated in chapter 1, an anti-de Sitter space is a maximally symmetric Lorentzian space with constant and negative curvature, i.e. Ricci scalar. It is the analogous of a Lobachevsky space in Euclidean signature. Similarly to a Lobachevsky space, see [12] chapter 23, a d-dimensional AdS space, AdS_d for short, can be defined by embedding in a $(d+1)$-dimensional space with two time-like coordinates:

$$ds^2 = -(dX^0)^2 + \sum_{i=1}^{d-1}(dX^i)^2 - (dX^d)^2\,; \tag{3.106}$$

by imposing the constraint:

$$-(X^0)^2 + \sum_{i=1}^{d-1}(X^i)^2 - (X^d)^2 = -R^2 \tag{3.107}$$

which implies a $SO(2,d-1)$ symmetry.

A choice of coordinates which manifestly satisfies the constraint (3.107), i.e. by simply substituting the below definitions into the left hand side of the above equation one gets $-R^2$ after a few cancellations, is the following:

$$X^0 = \frac{1}{2u}\left[1 + u^2\left(R^2 + x^i x_i - t^2\right)\right]\,; \tag{3.108}$$

$$X^i = Rux^i\,; \tag{3.109}$$

$$X^{d-1} = \frac{1}{2u}\left[1 - u^2\left(R^2 - x^i x_i + t^2\right)\right]\,; \tag{3.110}$$

$$X^d = Rut\,; \tag{3.111}$$

where $i = 1, \ldots, d-2$.

In terms of the coordinates defined in equations (3.108), (3.109), (3.110) and (3.111), the line element (3.106) becomes, see [48] chapter 2:

$$ds^2 = R^2\left[u^2\left(-dt^2 + \sum_{i=1}^{d-2} dx^i dx_i\right) + \frac{du^2}{u^2}\right]\,; \tag{3.112}$$

where $u > 0$.

Coordinates (t, x^i, u) are called Poincaré coordinates. They only cover part of the AdS space though, the so-called Poincaré patch. Another way of writing the metric (3.112) that will be useful later is by replacing $u \to y \equiv u^{-1}$.

$$ds^2 = \frac{R^2}{y^2}\left(-dt^2 + \sum_{i=1}^{d-2} dx^i dx_i + dy^2\right) ; \tag{3.113}$$

where $0 < u < \infty \Rightarrow 0 < y < \infty$.

The first indication of a possible duality between a theory in AdS space and a conformal field theory is the compatibility of their symmetry groups. The isometry group of AdS_{d+1} space is $SO(2,d)$ as shown above. This matches exactly the conformal group of a conformal field theory in d dimensions. This is certainly not enough to say there is a duality between two theories of these kinds. Nonetheless we learn that, if there is such a duality, it must be realized between a $(d+1)$-dimensional theory in AdS space and a d-dimensional conformal field theory.

Now, let us rewrite equation (3.113) in terms of $\tilde{y} \equiv -\log(y/R) \Rightarrow e^{-\tilde{y}} = y/R$.

$$ds^2 = e^{2\tilde{y}}\left(-dt^2 + \sum_{i=1}^{d-2} dx^i dx_i\right) + R^2 d\tilde{y}^2 \tag{3.114}$$

It is clear to see from the above line element that it takes light a finite amount of time to reach the boundary $\tilde{y} \to \infty$. Supposing light travel in the $\tilde{y}$-direction only, i.e. $dx^i dx_i = 0$:

$$e^{2\tilde{y}} dt^2 = R^2 d\tilde{y}^2 \Rightarrow \int_{t_0}^{t_f} dt = R\int_{y_0}^{\infty} d\tilde{y} e^{-\tilde{y}} \therefore \Delta t \equiv t_f - t_0 = Re^{-y_0} < \infty \tag{3.115}$$

The above observation leads us to the conclusion that light can go to the boundary of the Poincaré patch (see [12] for a generalization valid for the whole AdS space), bounce off of it, and come back in a finite amount of time. Therefore, it raises the possibility that light could leave an imprint of what happens in the bulk of space at its boundary. That is an indication of holography, justifying the name of the principle. Note that the idea of reducing one dimension from the bulk to the boundary endorses the matching of AdS_{d+1} space isometry group with a conformal group of a field theory with one less dimension.

As of yet, we have described only the compatibility between a theory in an AdS background and a conformal field theory that would live in its boundary. But a specific realization of such a duality still needs to be given. By the time Maldacena had the idea of the first realization, it had been found in previous works [57, 58, 59] that a near-extremal, i.e. mass just slightly larger than charge, black hole could be described by a system with a large number N of D-branes. He found an equivalence between two different theories by analyzing this low energy limit, or decoupling limit as we will see, in two different ways: one gravitational and the other analyzing the physics on D-branes.

First, the production of Hawking radiation can be seen, from a string theory perspective, as two open strings ending on a D-brane which can interact to form a closed string, moving away from the D-brane. This closed string would vibrate, as usual, creating a mode which is identified as the particle radiated by the black hole, see figure 3.2. Then, the system is composed of a D-brane, open strings attached to it, closed strings free to move in the bulk of space, and the interaction term between open and closed strings creating the Hawking radiation. It turns out that the coupling of this interaction is proportional to the gravitational constant $G^{(10)}$ in $d = 10$ dimensions which is proportional to α'^4. Therefore, the low energy limit $\alpha' \to 0$ (more precisely $\alpha' E^2 \to 0 \Leftrightarrow E \ll (\alpha')^{-1/2}$ since α' isn't dimensionless, where E is the energy scale considered in this analysis) also decouples D-brane processes from the bulk physics, hence the name decoupling limit. In summary, we end up with free supergravity, the low energy limit of a superstring theory, in the 10-dimensional bulk and the action on the D-brane, which becomes just the $\mathcal{N} = 4$ SYM action for a D3-brane as pointed out in the superstrings section. The large number N of D-branes becomes the rank of the gauge group $SU(N)$ in the SYM theory.

The other way of analyzing the same system is by adopting a gravitational viewpoint, meaning that instead of considering D-branes one considers only the space curved by them. It turns out that the decoupling limit $\alpha' E^2 \to 0$ also implies a decoupling between bulk and near-boundary physics in this case. It could be shown that this limit leads to a low energy limit near the boundary, see [48] chapter 10. In the end, there is again a free gravity theory

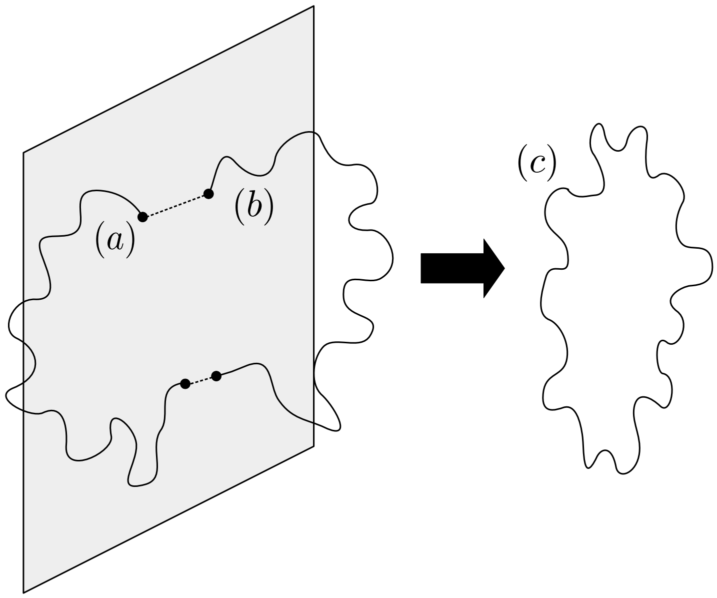

Figure 3.2: Production of Hawking radiation from black holes. From a string theory perspective, the process is understood as the interaction of two open strings, (a) and (b), to form a closed string (c) which is understood as radiation moving away from the black hole.

in the bulk plus low energy gravity excitations near the boundary.

Matching the two viewpoints, it is clear that in the bulk (far from the boundary) they are already described by the same theory, but at the boundary (or near it in the second case) the theories we got were a priori unrelated. However, since they come from the same system seen from two different perspectives, the two (near) boundary theories should be equivalent. The equivalence is between $\mathcal{N}=4$ SYM, a conformal field theory in $d=4$ dimensions, and a low energy supergravity theory in AdS space, hence the name AdS/CFT correspondence! It should be pointed out that this was an heuristic derivation of the correspondence by Maldacena. Many attempts to a full derivation can be found in the literature, see for instance [60, 61, 62].

About the gravitational dual in the above AdS/CFT setup, one could show that, in the decoupling limit, the line element of type IIB supergravity becomes, see [48]:

$$ds^2 \simeq \frac{r^2}{R^2}\left(-dt^2+\sum_{i=1}^{3} dx^i dx_i\right)+\frac{R^2}{r^2}dr^2+R^2 d\Omega_5^2\,; \tag{3.116}$$

where $d\Omega_5$ is an infinitesimal 5-dimensional solid angle.

The coordinate r is called the holographic coordinate since this is the one we impose a limit on, in this case $r \to 0$, in order to reach the boundary of the AdS space, see [48]. It can be shown it is related to the energy scale E of the SYM theory through $E \propto r/\alpha'$.

By a redefinition of the r coordinate $r \to y \equiv R^2/r$ equation (3.116) becomes:

$$ds^2 \simeq \frac{R^2}{y^2}\left(-dt^2 + \sum_{i=1}^{3} dx^i dx_i + dy^2\right) + R^2 d\Omega_5^2\,; \tag{3.117}$$

which, comparing to equation (3.113), can be seen to be the line element of an $AdS_5 \times S^5$ space of radius R.

Apart from the compact space which is not that relevant at the boundary of AdS_5, the duality is indeed holographic, with the conformal field theory having one dimension less than the bulk theory. It should be pointed out that although the heuristic derivation by Maldacena was done in a very specific low energy limit, nowadays the correspondence is believed by many to hold at the full string theory level, not only at $\alpha' E^2 \to 0$ and $g_s \to 0$ (the string coupling).

One of the most important features of AdS/CFT, crucial in order to earn the name duality, is that it is a correspondence between a strongly coupled field theory and a weakly coupled gravity dual, or vice-versa. This can be seen from the fact that the radius R in equation (3.117) has the specific form $R^4 = 4\pi g_s N \alpha'^2$ coming from the type IIB supergravity solution implicitly used in the second point of view taken above to find the equivalence between the dual theories, see [48] chapter 10. Furthermore, the limit $\alpha' \ll 1$ does not have a meaning by itself unless we specify its units, as previously discussed. A more precise limit would be $\alpha'/R^2 \ll 1$. From that it follows:

$$\frac{R^2}{\alpha'} = \sqrt{4\pi g_s N} = \sqrt{g_{YM}^2 N} \equiv \sqrt{\lambda} \gg 1\,; \tag{3.118}$$

where g_{YM} is the SYM coupling which relates to g_s through $4\pi g_s = g_{YM}^2$ basically because a closed string interaction can be seen as two open string interactions (with coupling g_o) leading to $g_s \propto g_o^2$ and since Yang-Mills fields live on the D-brane $g_o \propto g_{YM}$, see [48]. The 't Hooft coupling $\lambda \equiv g_{YM}^2 N$ is the effective coupling to expand on for $SU(N)$ gauge theories with large N [63]. Therefore, when we have a perturbative, i.e. weakly coupled gravity theory $\alpha'/R^2 \ll 1$, the field theory dual is nonperturbative $\lambda \gg 1$. Considering a duality valid also for $\alpha'/R^2 \gg 1$, the opposite will be valid too.

The final issues we need to address are the relation between physical quantities on both sides of the duality, i.e. how to construct a map between them, and how one relates observables on both sides. The prescription to relate physical quantities goes by the name of duality map or dictionary. After the dimensional reduction, fields in AdS_5 should be related to gauge-invariant operators in the field theory dual. Here we will only list some field-operators pairs that will be useful later, see [48] for more details. Considering a global symmetry group G in the field theory dual, its Noether current J_μ corresponds to the gauge field of the local counterpart of the same symmetry group in the gravity dual. Furthermore, the energy momentum tensor $T_{\mu\nu}$ in the field theory is related to the metric $g_{\mu\nu}$ in the gravity dual, see [12] chapter 23, which is a special case of the last statement since $T_{\mu\nu}$ is the current for spacetime translations.

The relation between observable quantities on both sides of the duality was found through the works of Gubser, Klebanov, Polyakov and Witten [64, 65] and therefore is called GKPW construction. It can be shown that AdS massless fields, e.g. scalars $\phi(x^i, y)$ where the coordinates follow the notation of equation (3.117), are y-independent at the boundary $y \to 0$, $\phi_0(x^i)$. The y-independent field at the boundary will work as a source for the field theory operator $\mathcal{O}$ which is related to it by the duality map. Therefore, one can write the generating functional of the field theory as:

$$Z_{\text{CFT}}[\phi_0] = \int \mathcal{D}[\Phi]\, e^{-S_{SYM}[\Phi] + \int d^d x \mathcal{O}(x)\phi_0(x)}\,; \tag{3.119}$$

where we are working in Euclidean space, obtained via Wick rotation from Minkowski space as usual, and $\mathcal{D}[\Phi]$ represents a path integration over all $\mathcal{N} = 4$ SYM fields for short.

Field theory observables are related to n-point functions which are obtained by functional differentiation of Z_{CFT}.

The GKPW construction simply equates the partition functions of the fiel theory and the one of its gravity dual.

$$Z_{\mathrm{CFT}}[\phi_0] = Z_{AdS}[\phi[\phi_0]] \tag{3.120}$$

This is believed to be valid in general if the correspondence holds indeed at the full string theory level, but in particular in the decoupling limit, which leads to supergravity in the bulk, we have: $Z_{AdS}[\phi[\phi_0]] \simeq e^{-S_{SUGRA}[\phi[\phi_0]]}$.

3.6.3 Gauge/Gravity Duality: The General Case

The original setup of $AdS_5 \times S^5$ vs. $\mathcal{N} = 4$, $d = 4$ Super Yang-Mills was described above. However, this is believed to be just the most solid realization we have nowadays of a broader holographic principle. In general, we would have an analogous duality between a field theory, which might not be conformal invariant, and a gravitational theory. One could consider a case with less supersymmetry also. All in all, the generalized duality is called gauge/gravity duality. Many examples of such dualities are studied in book [48] part III.

In general, gauge/gravity dualities have some common properties such as:

1. Symmetries of both sides of the duality have to match;
2. They relate a strongly coupled theory on one side to a weakly coupled theory on the other;
3. The gauge theory has one dimension less than its gravity dual, living at the boundary of it;
4. A relation $E \propto r/\alpha'$ should hold between the energy scale E of the field theory and the holographic coordinate r, the one on which we impose a limit to reach the boundary of the gravity dual space.

There are other general properties, see [12] chapter 23, but those are the most relevant ones for our later developments. In chapter 5 these features will play an important role when defining a field theory dual to our FLRW background of cosmology.

Chapter 4

String Gas Cosmology

The first out of three types of cosmologies that are going to be described in this thesis in which traces of dark energy arise quite naturally is the so-called string gas cosmology (SGC). The goal of SGC is to use string theory insights and unique features, namely T-duality and the extra winding and oscillatory degrees of freedom, which are absent in quantum field theories, to explain the early universe dynamics. The theory was started by Brandenberger and Vafa [1] and it is mainly motivated by the following observation. It is known that if inflationary cosmology is realized in the context of classical GR coupled to a scalar field, then the theory will contain an initial cosmological singularity [66]. Singularities in physics usually indicate the extrapolation of the predictive power of a theory beyond its actual scope. It is natural to look for a theory of quantum gravity, like string theory, to describe physics at a time today's observable universe was shrunk to a very small volume leading to an enormously high mass density. See [67] for a discussion of how string theory would solve the singularity problem. In other words, if inflation lasts for more than an upper limit value, nowadays' cosmological scales $\sim H_0^{-1}$ would be less than the Planck size at the beginning of inflation. If this is the case, the effective field theory approach of inflation cannot be trusted and one must rely on a quantum theory of gravity.

Let us estimate the upper limit on the number of e-folds N_e for inflation to be trustworthy.

What we want to write down is an expression for the physical scale $\lambda_p = \lambda_p(a)$ in the beginning of inflation, starting with $\lambda_p(a_0) \sim H_0^{-1}$ nowadays. The comoving scale λ_c can be written as:

$$\lambda_p(a) = a\lambda_c \therefore \lambda_c = \frac{\lambda_p(a_0)}{a_0} \sim (a_0 H_0)^{-1} \tag{4.1}$$

From the beginning of the radiation-dominated era until now thermodynamic calculations tell us that the temperature of the radiation fluid in the universe has a temperature which is $\propto a^{-1}$. Therefore:

$$T \propto a^{-1} \therefore \frac{T_r}{T_0} = \frac{a_0}{a_r} \therefore T_r = \frac{a_0 T_0}{a_r} \geq 10^9\text{GeV} \sim 10^{22}\text{K} \therefore a_r \lesssim 10^{-22} a_0 \, ; \tag{4.2}$$

where the r subscript stands for the end of reheating, i.e $a_r \equiv a(t_{\text{end of reheating}})$ and $T_r \equiv T(a_r)$, and we have used the facts that $T_r \geq 10^9$GeV and $T_0 \simeq 2.7\text{K} \sim 1\text{K}$, see [12] chapters 5 and 10.

The scale of inflation is yet unknown since no tensor perturbations have been measured, see chapter 1. The only thing we know is that the end of inflation must have happened before the end of reheating, i.e. $a_e \leq a_r$. If inflation ended way before the end of reheating, current scales would be even more shrunk at the beginning of inflation. Therefore, the less stringent scenario would be to consider $a_e \sim a_r$. This is going to be considered here, but one should notice the strong, perhaps unrealistic, approximation we are considering. From the definition of the number of e-folds:

$$N_e \equiv \log\left(\frac{a_e}{a_b}\right) \therefore a_b = a_e e^{-N_e} \sim a_r e^{-N_e} \, ; \tag{4.3}$$

where a_b is the scale factor at the beginning of inflation.

Finally, using the constraint (4.2) and equation (4.3), we can write the physical scale at the beginning of inflation $\lambda_p(a_b)$ as:

$$\lambda_p(a_b) = a_b\lambda_c \sim a_r e^{-N_e}(a_0 H_0)^{-1} \lesssim e^{-N_e} 10^{-22} a_0 (a_0 H_0)^{-1} = e^{-N_e} 10^{-22} H_0^{-1} \tag{4.4}$$

Thus, the maximum number of e-folds one can have in order for $\lambda_p(a_b) \geq l_{Pl} = m_{Pl}^{-1} = (8\pi)^{-1/2} M_{Pl}^{-1} \sim 10^{-1} M_{Pl}^{-1}$ (m_{Pl} is the Planck mass while M_{Pl} is the reduced one) is:

$$l_{Pl} \sim 10^{-1} M_{Pl}^{-1} \lesssim e^{-N_e} 10^{-22} H_0^{-1} \therefore e^{-N_e} \gtrsim 10^{21} \frac{H_0}{M_{Pl}} \sim 10^{21} \times 10^{-60} = 10^{-39}$$
$$N_e \lesssim 39 \log(10) \simeq 90\,; \tag{4.5}$$

since $H_0 \simeq 70(\text{km/s})/\text{Mpc} \sim 10^{-42}\text{GeV}$ and $M_{Pl} \sim 10^{18}\text{GeV}$.

Recall that in this calculation the likely unrealistic approximation $a_e \sim a_r$ was considered. If $a_e \ll a_r$, the constraint would be more stringent since one has to go further back in time to reach the beginning of inflation ($a_b \to 0$) leading to a more severe shrinking of λ_p which can be seen from inequality (4.4) to yield a tighter constraint on N_e. Moreover, the effective field theory approach would break up at the string scale which could be lower than Planck mass, i.e. $m_s \leq m_{Pl}$, leading to $l_s \geq l_{Pl}$ and thus a tighter constraint on N_e also. In the literature, one usually assumes inflation is reliable if $N_e \lesssim 70$, see [12] chapter 32 and [68].

The minimum amount of inflation needed to solve pre-inflationary problems is $N_e \sim 60$, as discussed in chapter 1. Therefore, if $60 \lesssim N_e \lesssim 70$, one can perhaps trust the effective field theory approach of inflation, even though field theory has not been tested to such high energies. Nevertheless, there is no good reason for restricting $N_e \lesssim 70$, so it is important to figure out what string theory has to tell us about the early universe.

4.1 Basics of String Gas Cosmology

Unfortunately, string theory is only defined perturbatively as of yet and one would need a non-perturbative formulation of it below the Planck length. The idea of Brandenberger and Vafa was then to focus on features of string theory that should be valid even non-perturbatively and study a simplified model of a gas of non-interacting strings in a certain background. The important feature taken into consideration for the model was T-duality. Since the classical GR background isn't T-duality symmetric one has to consider another background to preserve the symmetry, see [68, 69]. It is natural to consider the background obtained in string theory by the condensation of the massless mode fields $g_{\mu\nu}$, $B_{\mu\nu}$ and the dilaton ϕ. It was said in chapter 3 that such condensation would change the Polyakov action (3.46) preserving string theories symmetries. The modified action is, see [70]:

$$\begin{aligned} S_P = -\frac{1}{4\pi\alpha'}\int d^2\sigma\big[&\sqrt{-\gamma}\gamma^{ab}g_{\mu\nu}(X)\partial_a X^\mu\partial_b X^\nu + \epsilon^{ab}B_{\mu\nu}(X)\partial_a X^\mu\partial_b X^\nu \\ &+ \alpha'\sqrt{-\gamma}\phi(X)R^{(2)}\big]\,; \end{aligned} \tag{4.6}$$

where ϵ^{ab} is the totally antisymmetric tensor on the worldsheet and $R^{(2)}$ is the Ricci scalar of the worldsheet manifold. Note that although we have previously written the Polyakov action in the bosonic string context, the above fields are also present in superstring theories. More specifically they appear as NS-NS (NS stands for Neveu-Schwarz) modes of superstrings.

Since one wants to write equation (4.6) as the background action and superstrings live on a 10-dimensional spacetime, one may try and rewrite the above worldsheet action (integrated over worldsheet coordinates) as a spacetime action instead (integrated over all $D = 10$ dimensions). It turns out that this is possible, using the adiabatic approximation to be defined later. To achieve that, one considers the equations of motion coming from S_P. We recall that the worldsheet fields are both X^μ and γ^{ab}. Moreover, Weyl invariance still leads to a traceless worldsheet energy-momentum tensor $T^a_{\ a} = 0$. Taking all that into consideration plus the requirement that Weyl symmetry isn't broken at the quantum level, see [70], one finds out that the following spacetime action reproduces the equations of motion up to tree level in α':

$$S_{ST} = \frac{1}{2\kappa_D^2} \int d^D x \sqrt{-g} e^{-2\phi} \left(R + c + 4\nabla_\mu \phi \nabla^\mu \phi - \frac{1}{12} H^2 \right) ; \tag{4.7}$$

where $2\kappa_D^2 = \left(2\pi\sqrt{\alpha'}\right)^{D-2} g_s^2 (2\pi)^{-1}$, g_s is the string coupling we have encountered in chapter 3, $c = 0$ in the critical dimension ($D = 10$ for superstrings) and $H^2 \equiv H_{\mu\nu\rho} H^{\mu\nu\rho}$ is the square of the field strength of $B_{\mu\nu}$ (in differential form notation $H \equiv dB$).

Having defined the background action, one adds free strings to it to form the non-interacting string gas of Brandenberger-Vafa. Those free strings are described by S_P. In the end, the full action is the sum of a contribution coming from the background (4.7) plus a sum of actions of the type (4.6) for each string composing the gas.

Several assumptions are made to further simplify this string gas scenario:

1. The fields are homogeneous in space, meaning that they are only time-dependent;
2. One considers the adiabatic approximation, meaning that t-dependency is small, which ends up leading to α' corrections being negligible. We have already made use of this approximation to write S_{ST} which reproduces S_P equations of motion only up to tree level on α';
3. Weak string coupling, i.e. $g_s \ll 1$;
4. All spatial dimensions are compactified.

This last assumption seems unrealistic since we need three large spatial dimensions to reproduce the physics at low energies, which we are used to in everyday life. Nonetheless, we will later see that SGC has an elegant way of naturally explaining the fact that we live in an effectively three-dimensional world.

4.2 Special Features of the Model

With the basic ingredients of SGC at hand, we must explore its consequences and highlight its different features in comparison to the inflationary scenario. First, coming back to the issue discussed at the beginning of the chapter, we will describe what the model has to say about the initial singularity, more specifically temperature singularity, one finds in inflationary scenarios. The spacetime singularity problem is more straightforwardly solved by T-duality since sub-string scale and large-scale physics are dual to each other. More precisely, [70] shows that T-duality applied to cosmology becomes a scale factor duality, with dual $a \leftrightarrow 1/a$ descriptions. This fact maps the problem of considering $a \to 0$ to a well-behaved model. Therefore, we will focus on the more subtle temperature singularity issue. This will lead to the assumption of having compactified dimensions in all directions. Then it will be discussed how one ends up with three large spatial dimensions in this scenario. After that, it will be seen how SGC solves the horizon problem in pre-inflationary cosmology and compare this solution against the one discussed in chapter 1 within inflationary models. In the end, the problem of moduli stabilization in the model will be issued.

4.2.1 Avoiding the Temperature Singularity

The usual picture one has in mind in the cosmological evolution of the universe is that of a current large and cold universe, with radiation fluid's temperature $\sim$ 1K, which was in the past much smaller and hot. One can make an analogy to any gas in a box. The more one shrinks the box, the higher the gas's temperature gets. Extrapolating this reasoning one may conclude that when $a \to 0$, in other words when the cosmological singularity is reached, the temperature increases indefinitely. The string gas scenario avoids this "temperature singularity" of standard cosmology with the help of T-duality.

Before highlighting the role of T-duality in it, let us observe that a gas of strings in thermal equilibrium has a maximum possible temperature in a fully compactified space. The density of string states increases exponentially with energy $\rho(E) \propto e^{\beta_H E}$, see [71], where for now β_H is just a proportionality constant (which is found to be $\propto \sqrt{\alpha'}$ as expected by dimensional analysis, see [1] appendix A) needed to make $\beta_H E$ dimensionless and whose importance will be seen later. Therefore, the partition function is written, as usual, as $Z = \sum_i \rho(E_i) e^{-\beta E_i}$, where $\beta \equiv 1/T$ is the inverse temperature of the gas of strings. Taking the continuum limit it becomes:

$$Z = \int_0^\infty dE e^{\beta_H E} e^{-\beta E} ; \tag{4.8}$$

which is only finite if $\beta_H - \beta < 0 \therefore \beta > \beta_H \therefore T < T_H$. The Hagedorn temperature was defined as $T_H \equiv 1/\beta_H$ and it sets the upper limit for a temperature of a gas of strings in equilibrium, solving the temperature singularity, as promised.

For sufficiently high energies, the specific heat of the gas is negative if one considers any non-compact dimension [1, 71]. This fact is an indication that, for consistency, one has to assume the string gas lives in a fully compactified space. The reason for this sudden change in the behavior of the gas whether it is put in a box, i.e. in a compact space, or not is due to T-duality. If we consider a non-compact dimension, there will be no winding modes for this direction. The dimensions could be very large, though, which is needed to explain current physics as will be discussed in the next subsection.

To emphasize the role of T-duality in solving the temperature singularity, we will anticipate the fact that the cosmological evolution in SGC at late times agrees with the standard cosmology picture. So, going back in time, the radii R of the large dimensions shrink, increasing the energy through an increase in the momentum contribution to equation (3.100) which, in turns, leads to an increase in temperature. However, when R becomes too small, coming close to the string scale $l_s \equiv \sqrt{\alpha'}$, the winding modes' contribution to the energy, which is $\propto R$, start to compensate the increase in energy due to momentum modes. All in all, the temperature never surpasses T_H as expected for consistency. Even when the radii R go below l_s one can use the T-dual description, i.e. $R \to \alpha'/R$, to see that the temperature never diverges since it have to follow the same pattern as for $R > l_s$, see [68] figure 3.

4.2.2 Three Large Dimensions

In the last subsection, the need for considering compact dimensions in all directions of space was described. However, one still needs to explain then how the three large dimensions we are used to from everyday life experience emerge within SGC. Since there isn't any good reason for us to assume three of the nine, considering superstring theory, compact spatial dimensions are intrinsically different from the others and start large from the beginning, it is natural to start with all dimensions at about the Planck length in the early universe, i.e. when $T \to T_H$. SGC has to explain why only three of these dimensions become large as the universe cools down. It does so from the simple fact that strings are one-dimensional objects.

First, notice that two dynamic point-particles, 0-dimensional objects, generically cross each other's path in, at most, one spatial dimension. Of course, there is the possibility of the two of them having the same velocity in the same direction, but since the distribution of possible velocities is continuous, the probability of them having exactly the same velocity is zero. So, in general, their paths will intersect for $d \leq 1$, where d is the number of spatial dimensions only. If $d > 1$, then, in general, their paths don't intersect. In the case of strings, their paths will, in general, cross if $d \leq 3$ dimensions only. One can extend the argument for two p-dimensional objects, which will cross each other's path only in a $d \leq 2p+1$ dimensional space (the maximum spacetime dimensionality has to match twice – since we are considering two objects – the number of dimensions of the p-dimensional object's worldvolume). Anyway, the point of SGC is that in $d \leq 3$ dimensions, in general, two strings with opposite winding numbers will intersect and then interact to form an unwound string, see figure 4.1. As all dimensions grow large, wound strings would become classical objects which would contain a further expansion of the dimensions around which the strings are wound. Since in three of the nine dimensions wound strings become unwound ones, they can grow large without being pressed back by wound strings forming the three large dimensions of our cold, late time universe. The other six dimensions are held back to small sizes by wound strings.

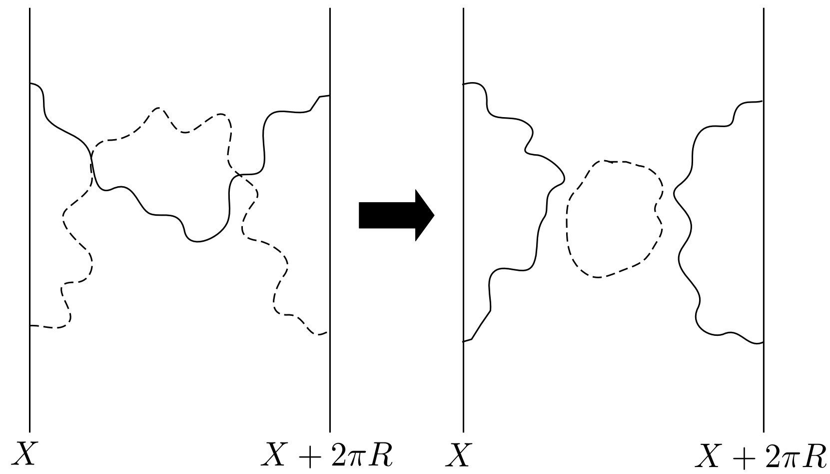

Figure 4.1: Wound strings, on the left, interacting to form unwound strings, on the right, allowing for the dimension X to grow large. The vertical lines symbolize regions of space that are identified, $X \sim X + 2\pi R$, forming a compact dimension.

4.2.3 Solution to the Horizon Problem

We must now solve the equations of motion for the gas of strings in the background introduced in equation (4.7) to figure out what SGC has to tell us about the evolution of the universe. The full action will be, as anticipated, the sum of S_{ST} and the modified Polyakov action, equation (4.6), for each string composing the gas. Strings are described by their stress-energy tensor $T_{\mu\nu}$ since they are the matter composing the gas. Let us start by considering a simplified cosmological ansatz for the form of the background fields:

$$\begin{aligned} ds^2 &= -dt^2 + \sum_{i=1}^{d} a_i^2(t) dx_i^2 = -dt^2 + e^{2\lambda(t)} \sum_{i=1}^{3} dx_i^2 + e^{2\nu(t)} \sum_{i=4}^{d} dx_i^2 \,; \\ H_3 &= h\, dx^1 \wedge dx^2 \wedge dx^3 \,; \\ \phi &= \phi(t) \,; \end{aligned} \tag{4.9}$$

where all the fields can only depend on time since this is an assumption of the model, as it was seen.

Note that one could have guessed a more generic line element $ds^2 = -dt^2 + \sum_{i=1}^{d} a_i^2(t)dx_i^2$ where all the scale factors are, a priori, different. However, anticipating the previously discussed feature of three dimensions growing large while the rest are kept small, we have chosen to define $a_1 = a_2 = a_3 \equiv e^{\lambda}$. For simplicity, we have also assumed all small dimensions behave in the same manner, namely $a_4 = \ldots = a_d \equiv e^{\nu}$. Furthermore, a measure of the volume of the total space, i.e. considering large and small dimensions, is given by:

$$V \equiv \prod_{i=1}^{d} a_i = e^{3\lambda+6\nu}\,; \tag{4.10}$$

while it is shown in [70] that the energy density of the gas of strings and the pressure in the ith direction are:

$$\rho = \sum_s n_s E_s\,; \tag{4.11}$$

$$P_i = -\frac{1}{V}\frac{\partial(\rho V)}{\partial \log a_i} = -\frac{1}{V}\frac{\partial E}{\partial \log a_i}\,; \tag{4.12}$$

where $\sum_s$ is a sum over all species of gases and n_s is the number density of each species. In the second line, we have defined the total energy $E \equiv \rho V$ for short. Notice that considering the ansatz (4.9) we will only have two different pressures in different directions:

$$P_\lambda = -\frac{1}{V}\frac{\partial E}{\partial \lambda}\,; \text{ and} \tag{4.13}$$

$$P_\nu = -\frac{1}{V}\frac{\partial E}{\partial \nu}\,; \tag{4.14}$$

where P_λ is the pressure in any of the large directions while P_ν is the pressure in the other directions.

The equations of motion in the superstring case, where one has $d = 9$ spatial dimensions, are found in [70] to be:

$$R_{\mu\nu} + 2\nabla_\mu \nabla_\nu \phi - \frac{1}{4} H_{\mu\rho\sigma} H_\nu{}^{\rho\sigma} = 2\kappa_{10}^2 e^{2\phi} T_{\mu\nu} = 16\pi G^{(10)} e^{2\phi} T_{\mu\nu}\,; \tag{4.15}$$

$$\nabla_\mu \left(e^{-2\phi} H^{\mu\nu\rho}\right) = 0\,; \tag{4.16}$$

$$4\nabla_\mu \phi \nabla^\mu \phi - 2\nabla_\mu \nabla^\mu \phi - \frac{1}{6} H_{\mu\nu\rho} H^{\mu\nu\rho} = 2\kappa_{10}^2 e^{2\phi} T = 16\pi G^{(10)} e^{2\phi} T\,; \tag{4.17}$$

where $T \equiv T^\mu{}_\mu$, $G^{(10)}$ is the 10-dimensional Newton constant, keeping the notation of chapter 3, and the trace of equation (4.15) was used to rewrite the ϕ equation of motion as in equation (4.17). Note that we are working in the critical dimension, i.e. $c = 0$ in equation (4.7).

Plugging the ansatz (4.9) into the above equations of motion one finds out that equation (4.16) is identically satisfied while the other two equations of motion become:

$$-3\dot{\lambda}^2 - 6\dot{\nu}^2 + \dot{\varphi}^2 - \frac{h^2}{2} e^{-6\lambda} = e^\varphi E\,; \tag{4.18}$$

$$\ddot{\lambda} - \dot{\varphi}\dot{\lambda} - \frac{h^2}{2} e^{-6\lambda} = \frac{1}{2} e^\varphi \bar{P}_\lambda\,; \tag{4.19}$$

$$\ddot{\nu} - \dot{\varphi}\dot{\nu} = \frac{1}{2} e^\varphi \bar{P}_\nu\,; \tag{4.20}$$

$$\ddot{\varphi} - 3\dot{\lambda}^2 - 6\dot{\nu}^2 = \frac{1}{2} e^\varphi E\,; \tag{4.21}$$

where the dot stands for time derivative, as usual, and $\varphi \equiv 2\phi - \log V = 2\phi - \sum_{i=1}^d \lambda_i = 2\phi - 3\lambda - 6\nu$ is the T-duality invariant dilaton field, see [70]. Furthermore, we have defined for short $\bar{P}_i \equiv P_i V = P_i e^{3\lambda + 6\nu}$, where $i = \lambda, \nu$.

In addition to equations (4.18) to (4.21), one has to impose the conservation of the stress-energy tensor, which leads to:

$$\partial^\mu T_{\mu\nu} = 0 \Rightarrow \dot{E} + 3\dot{\lambda}\bar{P}_\lambda + 6\dot{\nu}\bar{P}_\nu = 0 \tag{4.22}$$

Equations (4.18), (4.19), (4.20), (4.21) and (4.22) are the ones from which we analyze how the cosmological evolution is described within SGC. The first thing to notice is that the B-field contribution, through h, is only relevant in the beginning of the universe, i.e. when all $a_i \to 0$, since it is supressed by a factor $a_\lambda^{-6} \equiv e^{-6\lambda}$ and a_λ grows large with time as it was argued in the last subsection, giving rise to the three large spatial dimensions. Computational simulations, see [12] chapter 32, shows that in the early period of evolution, there is a so-called cosmological loitering period in which $\dot{\lambda} = 0$ and $\dot{\varphi} = 0$. Note that $\dot{\lambda} = 0 \therefore \dot{a}_\lambda = 0$ means that the three large spatial dimensions don't evolve. In fact it takes some time until all the winding modes in these directions annihilate among themselves allowing these directions to grow large.

It is usual to call Hagedorn phase [68] the period until winding modes in the future large directions vanish, including the cosmological loitering period. The Hagedorn phase solves the horizon problem within standard cosmology in an alternative fashion, compared to inflation. Recall that the problem is solved by inflation because the largest scales we have access to today were once inside the horizon only to be blown outside it during inflation while H^{-1}, the Hubble radius, remains approximately constant. It is just later, after the end of inflation, that these scales come back inside the horizon. The solution within SGC is quite the opposite, current largest scales are almost constant during the Hagedorn phase while the Hubble radius starts off infinite, $\dot{a_\lambda} \to 0 \therefore H \equiv \dot{a_\lambda}/a_\lambda \to 0 \therefore H^{-1} \to \infty$, with the scales inside it, and H^{-1} starts to shrink as the winding modes are annihilated. At the end of the annihilation we are left with unwound strings, see figure 4.1, which form the matter present in the current universe starting the radiation-dominated era of cosmology. Therefore, in this sense, the annihilation of winding modes would assume the analogous role of reheating in standard cosmology. Reheating connects the inflationary period with the radiation-dominated era, whereas within SGC, winding modes annihilation connects the Hagedorn phase to the radiation-dominated period. After the start of the radiation-dominated era, both standard cosmology plus inflation and SGC pictures coincide, with the scales which went outside the horizon coming back inside eventually. The first scales to go outside the horizon are the last to come back inside it in both pictures.

The main difference between standard cosmology, with inflation, and the string gas scenario is outlined in figure 4.2. A comprehensive analysis of the Hagedorn phase and its comparison with inflation is found in [68, 72].

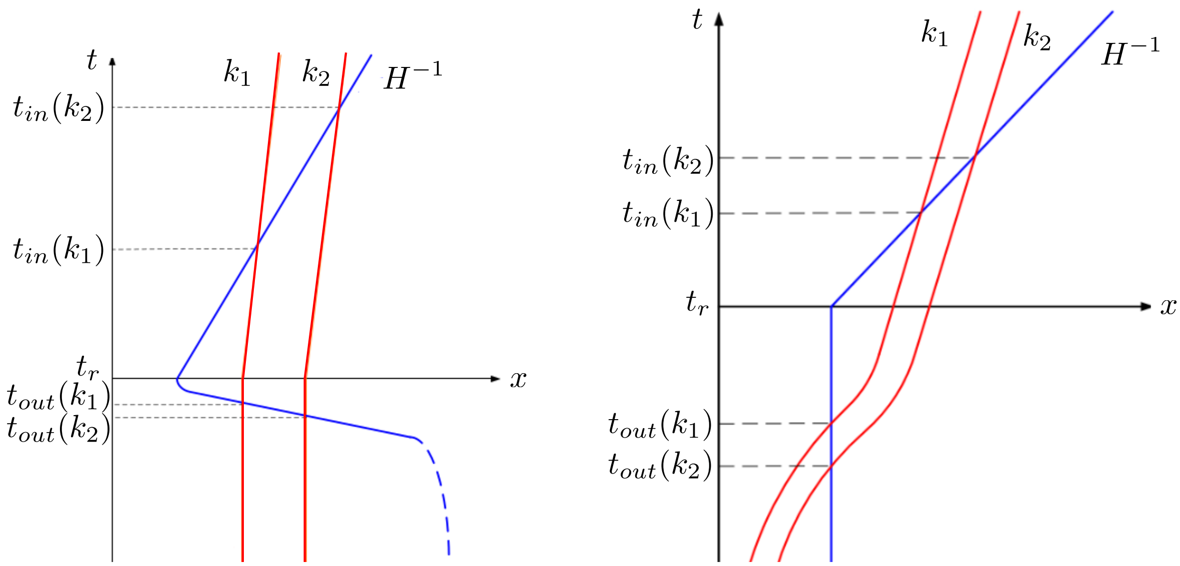

Figure 4.2: Comparison between inflationary and SGC pictures of the early universe. The figure on the left represents the SGC scenario, where the Hubble radius H^{-1} starts off infinite and then shrinks so that the comoving scales k_1 and k_2 get outside it. On the right, inflation says H^{-1} is almost constant in this primordial era while the scales get blown outside it. At the moment $t = t_r$ reheating happens, and its analogous within SGC, thus both behaviors agree for $t > t_r$. In both diagrams, $t_{out}(k_i)$ ($t_{in}(k_i)$) stands for the time at which the scale k_i ($i = 1, 2$) gets outside (comes back inside) the horizon. Note that, in both cases, the earlier a scale gets outside the horizon the later such scale comes back inside it. Both x-axes represent physical distances, rather than comoving ones.

4.2.4 T-Duality and Moduli Stabilization

The issue of moduli stabilization is a rather subtle one. One must be able to stabilize shape, size, and the dilaton moduli to respect experimental observations, as discussed in chapter 3. Here we will only give a qualitative description of the role of T-duality in the stabilization of size moduli. For more details see [68, 73] and the references therein.

In a gas of strings containing both momentum and winding modes, the energy associated with the first ones is $\propto 1/R$, where R is the size of the dimension being analyzed, while the energy associated to winding modes is $\propto R/\alpha'$, see equation (3.100). Therefore, in order to find a minimum energy point, taking these two contributions into account, the size of the dimension in question will tend to the self-dual radius $R = \sqrt{\alpha'} \equiv l_s$, considering $n = m$ and $N + \tilde{N} = 2$ in equation (3.100) for simplicity. Since the small dimensions have a preferred size, as it was argued, one can conclude the size moduli will not be dynamic if they already start with this stable size. Thus, one avoids the possibility of them interacting non-trivially with matter, which is constrained by experiments on Earth. In other words, they have been stabilized.

4.3 Connection to Observations

In the last section, it was shown that the string gas scenario is an alternative to inflation. Although they solve pre-inflationary problems in different ways, both of them seem, at least qualitatively, consistent with the generic picture of scales going outside the horizon during inflation (or the Hagedorn phase in SGC) to eventually come back inside it later. This is not enough, though, to say which one better explains reality. In science, Nature always has the last say. Therefore, one must be able to predict the power spectra of scalar and tensor perturbations within SGC to see whether the predictions agree with experimental data or not, similarly to what is done with inflationary models. In this section, we will give a taste of how one constructs such spectra from the thermodynamics of a string gas.

Note, however, that not all calculations will be shown here. Only the main results and the discussion will be presented. After all, the goal of the thesis is to show how dark energy features arise in different cosmological models. Nonetheless, this is an important discussion to fully introduce the string gas model and shall be presented here. More details are found in [68, 72]. A review of SGC's status after the 2015 Planck data release can be found in [74].

4.3.1 Thermodynamics

Recall from chapter 1 that the quantities one needs to calculate in order to write the power spectra Δ_s^2 and Δ_t^2 are the 2-point functions $\langle \zeta_k \zeta_{k'} \rangle$, of scalar perturbations, and $\langle h_k h_{k'} \rangle$, of tensor perturbations. Working in the Newton gauge with $\Phi = \Psi$ and considering also tensor perturbations, as opposed to equation (1.84), one gets the line element:

$$ds^2 = a^2(\eta) \left\{ -(1+2\Phi)d\eta^2 + \left[(1-2\Phi)\delta_{ij} + h_{ij}\right] dx^i dx^j \right\} \; ; \tag{4.23}$$

where η is the conformal time as usual.

Note that Φ is indeed the Newtonian potential which follows the Newton equation:

$$\nabla^2 \Phi = 4\pi G \delta\rho \, ; \tag{4.24}$$

for a small density fluctuation $\delta\rho$ as a source.

In momentum space, the above equation becomes:

$$-k^2 \Phi(k) = 4\pi G \delta\rho(k) = 4\pi G \delta T^0_{\ 0}(k) \tag{4.25}$$

For tensor perturbations one can show we arrive at an analogous equation:

$$-k^2 h^i_{\ j}(k) = 4\pi G \delta T^i_{\ j}(k) \, ; \, \text{where} \, i \neq j \tag{4.26}$$

Then, the 2-point functions one needs to calculate are:

$$\langle |\Phi(k)|^2 \rangle = \frac{(4\pi G)^2}{k^4} \langle \delta T^0_{\ 0}(k) \delta T^0_{\ 0}(k) \rangle = \frac{(4\pi G)^2}{k^4} \langle (\delta\rho(k))^2 \rangle \, ; \tag{4.27}$$

$$\langle |h(k)|^2 \rangle = \frac{(4\pi G)^2}{k^4} \langle \delta T^i_{\ j}(k) \delta T^i_{\ j}(k) \rangle \, ; \tag{4.28}$$

and therefore the correlators we must calculate from the thermodynamics of the gas of strings are $\langle \delta T^0{}_0 \, \delta T^0{}_0 \rangle$ and $\langle \delta T^i{}_j \, \delta T^i{}_j \rangle$ in order to later relate them to their momentum space counterparts which appear in equations (4.27) and (4.28).

One will note that we should have been looking for 2-point functions of the gauge-invariant scalar $\zeta(k)$ instead of $\Phi(k)$. Nevertheless, it is shown in [72] that these correlators are proportional to each other in the string gas scenario. Therefore, figuring out $\langle |\Phi(k)|^2 \rangle$ is enough to predict the behavior of Δ_s^2.

For a general thermodynamical system, in a spacetime with metric $g_{\mu\nu}$, the partition function Z is given by:

$$Z = \sum_s e^{-\beta\sqrt{-g_{00}}H(s)} = e^{-\beta F} \, ; \tag{4.29}$$

where F is the free energy, $\beta = 1/T$ is the inverse temperature and $\sum_s$ is a sum over every state of the system, each of which with its energy $H(s)$.

It can be shown that, from the definition of the ensemble average of the stress-energy tensor, one can write it in terms of the partition function as:

$$\langle T^\mu{}_\nu \rangle \equiv \frac{1}{Z} \sum_s e^{-\beta\sqrt{-g_{00}}H(s)} T^\mu{}_\nu \Rightarrow \langle T^\mu{}_\nu \rangle = \frac{2g^{\mu\rho}}{\sqrt{-g}} \frac{\delta}{\delta g^{\rho\nu}} \log Z \tag{4.30}$$

The definition of $\langle \delta T^\mu{}_\nu \, \delta T^\rho{}_\sigma \rangle$ comes from the expression of variance as $\langle \delta T^\mu{}_\nu \, \delta T^\rho{}_\sigma \rangle = \langle T^\mu{}_\nu \, T^\rho{}_\sigma \rangle - \langle T^\mu{}_\nu \rangle \langle T^\rho{}_\sigma \rangle$, leading to:

$$\langle \delta T^\mu{}_\nu \, \delta T^\rho{}_\sigma \rangle = \frac{2g^{\mu\lambda}}{\sqrt{-g}} \frac{\delta}{\delta g^{\lambda\nu}} \left(\frac{g^{\rho\xi}}{\sqrt{-g}} \frac{\delta}{\delta g^{\xi\sigma}} \log Z \right) + \frac{2g^{\rho\lambda}}{\sqrt{-g}} \frac{\delta}{\delta g^{\lambda\sigma}} \left(\frac{g^{\mu\xi}}{\sqrt{-g}} \frac{\delta}{\delta g^{\xi\nu}} \log Z \right) \tag{4.31}$$

Setting all free indices to 0 in equation (4.31) and considering the 3-dimensional part of the metric to be $g_{ij} \simeq R^2 \delta_{ij} \therefore \sqrt{-g} = \sqrt{-g_{00}} R^3 = R^3$ one finds, after some algebraic manipulation:

$$\langle \delta T^0{}_0 \, \delta T^0{}_0 \rangle = \langle (\delta\rho)^2 \rangle = \frac{T^2}{R^6} C_V \equiv \frac{T^2}{R^6} \left(\frac{\partial E}{\partial T} \right)_V \tag{4.32}$$

Analogously, setting $\mu = \rho = i$ and $\nu = \sigma = j$ in equation (4.31) one gets, see [68, 72]:

$$\langle \delta T^i{}_j \, \delta T^i{}_j \rangle = \frac{1}{\beta R^2} \left(\frac{\partial P}{\partial R} \right)_T = \frac{T}{R^2} \left(\frac{\partial P}{\partial R} \right)_T ; \tag{4.33}$$

where P is the pressure of the string gas defined as usual in thermodynamics:

$$P \equiv - \left(\frac{\partial E}{\partial V} \right)_S = T \left(\frac{\partial S}{\partial V} \right)_E \tag{4.34}$$

Up until now equations (4.32) and (4.33) have been found in [68, 72] and [12] chapter 33 with general thermodynamics considerations only. We still need to consider the gas of strings as a statistical mechanical system and compute C_V and $(\partial P/\partial R)_T$ in this case. For the Hagedorn phase, in which $T \simeq T_H \simeq l_s^{-1} \equiv (\alpha')^{-1/2}$, these are found in [12] to be:

$$C_V \simeq \frac{R^2}{l_s^3 T \left(1 - \frac{T}{T_H}\right)} \;\; ; \text{and} \;\; \left(\frac{\partial P}{\partial R} \right)_T \simeq \frac{2}{3} \left(\frac{T}{T_H} \right) \frac{\left(1 - \frac{T}{T_H}\right)}{l_s^3 R^2} \log \left[\frac{l_s^2}{R^2 \left(1 - \frac{T}{T_H}\right)} \right] \tag{4.35}$$

Plugging the results of (4.35) into the expressions (4.32) and (4.33) one gets:

$$\langle \delta T^0{}_0 \, \delta T^0{}_0 \rangle = \langle (\delta\rho)^2 \rangle = \frac{T}{R^4 l_s^3} \frac{1}{\left(1 - \frac{T}{T_H}\right)} ; \tag{4.36}$$

$$\langle \delta T^i{}_j \, \delta T^i{}_j \rangle = \frac{2}{3} \left(\frac{T}{T_H} \right) \frac{T \left(1 - \frac{T}{T_H}\right)}{l_s^3 R^4} \log \left[\frac{l_s^2}{R^2 \left(1 - \frac{T}{T_H}\right)} \right] \tag{4.37}$$

4.3.2 Power Spectra

In order to obtain the power spectra, we still need to translate the correlators of equations (4.36) and (4.37) to momentum space and relate them to the desired 2-point functions $\langle|\Phi(k)|^2\rangle$ and $\langle|h(k)|^2\rangle$. First, considering a region of radius $R \simeq k^{-1}$, inspired in equations (4.27) and (4.28) we can write their position space counterparts as:

$$\langle|\Phi(R)|^2\rangle \simeq (4\pi G)^2 R^4 \langle(\delta\rho)^2\rangle = (4\pi G)^2 \frac{T}{l_s^3} \frac{1}{\left(1-\frac{T}{T_H}\right)}\,; \tag{4.38}$$

$$\langle|h(R)|^2\rangle \simeq (4\pi G)^2 R^4 \langle \delta T^i_{\ j}\, \delta T^i_{\ j}\rangle = (4\pi G)^2 \left(\frac{2T^2}{3T_H}\right) \frac{\left(1-\frac{T}{T_H}\right)}{l_s^3} \log\left[\frac{l_s^2}{R^2\left(1-\frac{T}{T_H}\right)}\right] \tag{4.39}$$

Now, from equation (1.88), one can see that, apart from a delta function, the scalar perturbation 2-point functions in momentum and position spaces are related by only one Fourier transform:

$$\langle|\Phi(k)|^2\rangle = (2\pi)^3 \delta^3(\vec{k}+\vec{k}') \int d^3R \langle|\Phi(R)|^2\rangle e^{i\vec{k}\cdot\vec{R}} \equiv (2\pi)^3 \delta^3(\vec{k}+\vec{k}') P_\Phi(k)\,; \tag{4.40}$$

where the last equality is just the definition of the power spectrum $P_\Phi(k)$.

Therefore, still considering a region of radius $R \simeq k^{-1}$, one can argue that $P_\Phi(k) \simeq [V\langle|\Phi(R)|^2\rangle]_{R\simeq k^{-1}} \sim k^{-3}[\langle|\Phi(R)|^2\rangle]_{R\simeq k^{-1}}$, where $V \sim R^3 \simeq k^{-3}$. Furthermore, remember that $\langle|\zeta(k)|^2\rangle \sim \langle|\Phi(k)|^2\rangle \therefore P_\zeta(k) \sim P_\Phi(k)$, see [72]. Thus, finally, one arrives at:

$$\Delta_s^2(k) \equiv \frac{k^3}{2\pi^2} P_\zeta(k) \sim \left(\frac{k^3}{2\pi^2}\right) k^{-3} (4\pi G)^2 \frac{T(k)}{l_s^3} \frac{1}{\left(1-\frac{T(k)}{T_H}\right)} \propto \frac{T(k)}{T_H} \frac{1}{\left(1-\frac{T(k)}{T_H}\right)}\,; \tag{4.41}$$

where, in the last proportionality relation, the approximation $T_H \simeq l_s^{-1}$ was used and the dimensionless proportionality constant was hidden since it would be irrelevant for our discussion anyway. The k-dependence of the temperature stands for the temperature at the moment a certain scale $R \simeq k^{-1}$ goes outside the horizon.

A completely analogous argument can be made to the tensor power spectrum. Thus:

$$P_h(k) \sim k^{-3}[\langle |h(R)|^2 \rangle]_{R \simeq k^{-1}} = k^{-3}(4\pi G)^2 \left[\frac{2\,(T(k))^2}{3T_H}\right] \frac{\left(1-\frac{T(k)}{T_H}\right)}{l_s^3} \log\left[\frac{k^2 l_s^2}{\left(1-\frac{T(k)}{T_H}\right)}\right] \Rightarrow$$

$$\Delta_t^2(k) = 2\Delta_h^2(k) \equiv \frac{k^3}{\pi^2} P_h(k) \propto \left(\frac{T(k)}{T_H}\right)^2 \left(1-\frac{T(k)}{T_H}\right) \log\left[\frac{k^2 l_s^2}{\left(1-\frac{T(k)}{T_H}\right)}\right] \tag{4.42}$$

With both power spectra in hand, one can finally compare against CMBR observations and see whether SGC is as good a fit to Planck data [18] as inflation. The first thing to be noticed is that the scalar power spectrum is almost scale-invariant. The only scale dependency comes from $T(k)$ which is almost a constant during the Hagedorn phase. This scale invariance up to a high degree agrees with experiments. Also Δ_t^2 is almost k-independent. Even though there is a factor of k^2 inside the logarithm, one must notice that $T(k) \sim T_H \therefore \left(1-\frac{T(k)}{T_H}\right) \to 0$ which highly suppresses the k-dependency.

Let us check what can be inferred about the scalar spectral index n_s. From equations (1.91):

$$\log \Delta_s^2(k) = \log\left(\frac{T}{T_H}\right) - \log\left(1-\frac{T}{T_H}\right) + \log(\text{constant}) \Rightarrow$$

$$n_s - 1 \equiv \frac{d\log \Delta_s^2}{d\log k} = k\frac{d\log \Delta_s^2}{dk} = k\left[\frac{1}{T}\frac{dT}{dk} + \left(\frac{1}{1-\frac{T}{T_H}}\right)\frac{1}{T_H}\frac{dT}{dK}\right] = \frac{kT_H}{T(T_H-T)}\frac{dT}{dk} \tag{4.43}$$

Note that, since larger scales (in position space, corresponding to the smaller ks) comes outside the horizon earlier, at higher temperatures. Therefore, for $R < R' \Leftrightarrow k > k'$ then $T(k) < T(k')$, which means $\frac{dT}{dk} < 0$. Therefore, since $T_H > T$, we conclude from equation

(4.43) that $n_s - 1 < 0$ (a so-called red tilt), which agrees with Planck's results as pointed out in the summary section of chapter 1.

On the other hand, since:

$$\log \Delta_t^2(k) = 2\log\left(\frac{T}{T_H}\right) + \log\left(1 - \frac{T}{T_H}\right) + \log\left[\log\left(k^2 l_s^2\right) - \log\left(1 - \frac{T}{T_H}\right)\right] + \log(\text{constant})\,; \tag{4.44}$$

One finds out, from equation (1.95), that the tensor tilt has the opposite behavior:

$$\begin{aligned} n_t \equiv \frac{d\log\Delta_t^2}{d\log k} &= k\frac{d\log\Delta_t^2}{dk} = k\left\{\frac{2}{T}\frac{dT}{dk} - \frac{1}{T_H - T}\frac{dT}{dk} + \log^{-1}\left[\frac{k^2 l_s^2}{\left(1-\frac{T}{T_H}\right)}\right] \times \right. \\ &\left. \times \left(\frac{2}{k} + \frac{1}{T_H - T}\frac{dT}{dk}\right)\right\} \\ &\simeq k\frac{dT}{dk}\left(\frac{2}{T} - \frac{1}{T_H - T}\right) \simeq -\frac{k}{T_H - T}\frac{dT}{dk} > 0\,; \end{aligned} \tag{4.45}$$

where in the fourth equality we have neglected entirely the last term of the previous expression since during the Hagedorn phase $\frac{T}{T_H} \to 1 \therefore \log^{-1}\left[\frac{k^2 l_s^2}{\left(1-\frac{T}{T_H}\right)}\right] \to 0$. In the last equality again it was considered $\frac{T}{T_H} \to 1 \therefore T_H - T \ll T$.

This tensor blue tilt is a distinct characteristic of the string gas scenario. Slow-roll inflation leads to a tensor red tilt, i.e. $n_t < 0$, see [12] chapter 11. Unfortunately, tensor perturbations have not been measures yet. However, if one day it is measured to be $n_t > 0$, that will be a big triumph for SGC.

Planck data also constraints the ratio of tensor to scalar perturbations to be smaller than a certain value, $r \ll 1$. The string gas scenario also predicts this, at least qualitatively. From equations (4.41) and (4.42):

$$r = \frac{\Delta_t^2(k)}{\Delta_s^2(k)} \sim \left(1 - \frac{T}{T_H}\right)^2 \log\left[\frac{k^2 l_s^2}{\left(1 - \frac{T}{T_H}\right)}\right] = -\left(1 - \frac{T}{T_H}\right)^2 \log\left[\frac{\left(1 - \frac{T}{T_H}\right)}{k^2 l_s^2}\right] \ll 1\,; \quad (4.46)$$

where, again, in the Hagedorn phase $\frac{T}{T_H} \to 1$.

In conclusion, SGC was shown to have, at least qualitatively since the precise form of $\frac{dT}{dk}$ is unknown, the desired properties to match experimental data. The tensor blue tilt prediction is an interesting one to compare against possible future observations.

4.4 Brane Gas Cosmology

In chapter 3, it was argued that physicists came to realize string theory is not just a theory of strings [52]. Higher-dimensional objects, D-branes, play a central role in the theory as well. Therefore, it is natural to consider a gas of D-branes as a generalization of the string gas and look for new features this scenario may bring us. We will mainly discuss how a dark energy component arises in brane gas cosmology and comment on how some special features of SGC translate to this new scenario based on [2].

In a gas of D-branes, the spacetime background action can be taken to be the same as before, namely S_{ST} of equation (4.7), but the components of the gas get changed, meaning that the modified Polyakov action (4.6) has to be replaced by a brane action which includes interactions with the background fields $g_{\mu\nu}$, $B_{\mu\nu}$ and ϕ. As it was discussed in chapter 3, the Dp-brane action modified by this NS-NS background fields and the gauge field A_a, with field strength F_{ab}, living on the brane is the Dirac-Born-Infeld (DBI) action (3.88):

$$S_{DBI} = -T_p \int d^{p+1}\xi e^{-\phi} \sqrt{-\det\left(\tilde{g}_{ab} + \tilde{B}_{ab} + \alpha' F_{ab}\right)}\,; \quad (4.47)$$

where a and b are worldsheet indices and T_p is the brane tension. Moreover, $\tilde{g}_{ab}$ and $\tilde{B}_{ab}$ are the induced fields on the D-brane worldvolume:

$$\tilde{g}_{ab} = g_{\mu\nu}\partial_a X^\mu \partial_b X^\nu = g_{ab} + g_{ij}\partial_a\tilde{\phi}^i\partial_b\tilde{\phi}^j + g_{ib}\partial_a\tilde{\phi}^i + g_{ia}\partial_b\tilde{\phi}^i \,; \tag{4.48}$$

$$\tilde{B}_{ab} = \alpha' B_{\mu\nu}\partial_a X^\mu \partial_b X^\nu = \alpha'\left(B_{ab} + B_{ij}\partial_a\tilde{\phi}^i\partial_b\tilde{\phi}^j + 2B_{i[b}\partial_{a]}\tilde{\phi}^i\right) \,; \tag{4.49}$$

where, in the second equalities of both lines, we have applied the static gauge defining conditions: $X^a = \xi^a$ and $X^i = \tilde{\phi}^i$. One should not confuse the dilaton ϕ, nor its T-duality invariant version φ, with the scalars $\tilde{\phi}^i$ coming from brane transverse fluctuations.

For our cosmological purposes, we will consider a background with vanishing $B_{\mu\nu}$ and a metric with the same scale factor in all directions, which is a good ansatz before the annihilation of winding, and their generalization to $p > 1$ dimensional branes, wrapping modes. Recall that in the cosmological loitering period, all dimensions start on equal footing, it is just later that three spatial dimensions would have grown large. The generalization of this picture within brane gas cosmology will be discussed shortly. Anyway, the 10-dimensional line element is:

$$ds^2 = a^2(\eta)\left[-d\eta^2 + \sum_{i=1}^{d}(dx^i)^2\right] \,; \tag{4.50}$$

where $d = 9$ if we are working in the critical dimension of superstring theory.

If transverse fluctuations of the D-brane and the gauge field living on it are small, it is shown in [2] that the DBI action can be expanded as:

$$S_{DBI} \simeq -T_p \int d^{p+1}\xi e^{-\phi} a(\eta)^{p+1}\left[1 + \frac{1}{2}(\partial_a\tilde{\phi}_i)^2 + \frac{\alpha'^2}{4}a^{-4}F_{ab}F^{ab}\right] \tag{4.51}$$

Then, from the above action, one can calculate the energy-momentum tensor of a gas of Dp-branes wrapping modes only, similarly to what was done in the string case in [70]. It turns out that, in this specific case, one finds the equation of state:

$$P = -\frac{p}{d}\rho \Rightarrow w_p = -\frac{p}{d}\,; \tag{4.52}$$

where p is the dimensionality of the D-branes and d is the number of spatial dimensions. If one considers D3-branes in a $d = 3$ ($D = d + 1 = 4$ total spacetime dimensions) background, this gas would play the role of a cosmological constant. Of course, if one is working in the critical dimension $d = 9$, then the effect can only be obtained for a gas of D9-brane wrapping modes, but it would hardly have anything to do with our 4-dimensional cosmological constant.

We must point out that a gas of D-branes can be composed of different types of objects, i.e. D-branes with different ps. More realistic gases would also have momentum and oscillatory modes. The example given above was chosen knowing that it specifically yields dark energy's, or more precisely cosmological constant's, equation of state for the right values of d and p.

Note that this hypothetical gas of D3-brane wrapping modes in a $3+1$ dimensional spacetime, generating a cosmological constant equation of state, would have been relevant in the Hagedorn phase. Therefore, it is not a model for the dark energy we observe dominating the dynamics of the large-scale universe today, as opposed, for example, to the string-motivated model which will be discussed in chapter 6.

To complete our discussion, we will briefly compare some special features of brane gas cosmology with the ones presented for the string gas case, see [2] and the references therein for details. The first feature of SGC was the absence of a cosmological singularity. This remains the case for a gas of D-branes. The spacetime singularity is avoided by T-duality in the same fashion. Furthermore, one can show that the temperature of a gas of branes also has an upper limit, a Hagedorn temperature. The second feature presented was the natural emergence of three large spatial dimensions within SGC. For brane gases that contain D1-branes, i.e. strings, three spatial dimensions will still grow large. The difference is that if we have also higher dimensional Dp-branes in the gas other dimensions will grow large too since their wrapping modes can annihilate in up to $2p+1$ spatial dimensions, as anticipated. However, this is not a problem. For the sake of concreteness, suppose D2-branes wrapping modes annihilate in 5 dimensions, which will grow large. Then, if there are D1-branes winding modes on these large dimensions, they annihilate allowing 3 out of the previous 5 dimensions to grow even larger. Therefore, brane gas cosmology presents a hierarchy of dimensions as in the large extra dimensions scenarios which were pointed out in chapter 3. Finally, the picture one has in mind as a solution to the horizon problem within brane gas cosmology is also similar to that of SGC, since we still have a cosmological loitering period before wrapping modes of D-branes annihilate.

Chapter 5

Holographic Cosmology

Another idea one can use to get insights on the cosmological evolution, even in regimes where the standard approach of inflation plus ΛCDM cosmology would not be valid, is the gauge/gravity duality introduced in chapter 3. A cosmological theory that relies on such duality, usually to study gravity in the strongly coupled regime, goes by the name of holographic cosmology. The most important duality feature which makes it relevant for cosmology is that it is realized between a weakly coupled (perturbative) theory and a strongly coupled (non-perturbative) one. Therefore, even as we approach the strongly coupled gravity regime near the cosmological singularity, one should find a dual field theory that is perturbative. The point is that many of the known theoretical tools used to calculate relevant quantities in a theory rely on perturbative (approximative) methods, so without a duality, it is hard to analyze strongly coupled systems. Of course, the duality should hold also when gravity is weakly coupled, as in inflationary scenarios. In this case, the field theory dual would be strongly coupled.

As discussed in chapter 3, the first known and still best-understood gauge/gravity duality realization is the AdS/CFT correspondence. However, if one intends to apply such duality to cosmology, it is natural to look for a realization of it with a de-Sitter (dS) gravity dual. Recall that, in chapter 1, it was pointed out that cosmological observations tell us we live in a universe whose background is dS. Perhaps even more importantly, the exponential expansion of the universe during inflation also represents a dS phase, or at least almost dS. Based on the works of Strominger [75] and Witten [76], Maldacena started the idea of using holography in cosmology in [3]. He proposed a correspondence between the wavefunction of the universe and the partition function of the field theory dual in the following fashion:

$$\Psi[h_{ij}, \phi] = Z_{\mathrm{QFT}}[h_{ij}, \phi]\,; \tag{5.1}$$

where h_{ij} is the spatial part of the bulk metric and ϕ is the scalar field of the gravity plus scalar field theory considered in the bulk. More on how a wavefunction of an asymptotically dS space is defined can be found in [76].

Generalizing the AdS/CFT case of chapter 3, since the energy-momentum tensor operator T_{ij}, which is the conserved current due to translational symmetry, in the QFT dual is related to the "gauge" field of "local translations" (general coordinate transformations) in the bulk, i.e. the metric, one expects they couple in the expression for Z_{QFT} just like ϕ_0 and $\mathcal{O}$ do in equation (3.119). Therefore, the definition of T_{ij} is:

$$\langle T_{ij}(x)\rangle \equiv \frac{1}{\sqrt{h}} \frac{\delta Z_{\mathrm{QFT}}[h_{kl}, \phi]}{\delta h^{ij}(x)}\bigg|_{h_{kl}=0;\phi=0} = \frac{1}{\sqrt{h}} \frac{\delta \Psi[h_{kl}, \phi]}{\delta h^{ij}(x)}\bigg|_{h_{kl}=0;\phi=0}\,; \tag{5.2}$$

where the factor of 1 over the square root of the spatial metric determinant is conventional. In the last equality, we have just used equation (5.1) to write the stress tensor in terms of the bulk wavefunction.

Analogously to the above construction, the n-point function of the QFT dual operator $\mathcal{O}$, which couples to the bulk scalar field ϕ, is written as:

$$\langle \mathcal{O}(x_1)\ldots\mathcal{O}(x_n)\rangle = \frac{\delta^n \Psi[h_{kl}, \phi]}{\delta\phi(x_1)\ldots\delta\phi(x_n)}\bigg|_{h_{kl}=0;\phi=0}\,; \tag{5.3}$$

in terms of Ψ again, by the equivalence (5.1). Hence, one finds an expression for the wavefunction of the universe by integrating the above equation and considering the analogous of it for T_{ij} correlators:

$$\Psi[h_{ij},\phi] = \exp\left\{ \sum_n \frac{1}{n!} \int dx_1 \dots \int dx_n \Big[\langle \mathcal{O}(x_1)\dots\mathcal{O}(x_n)\rangle \phi(x_1)\dots\phi(x_n) + \langle T_{ij}(x_1)\dots T_{kl}(x_n)\rangle h^{ij}(x_1)\dots h^{kl}(x_n) \Big] \right\} \quad (5.4)$$

The point is, since Ψ is a wavefunction in the bulk, the correlation functions $\langle h_{ij}h_{kl}\rangle$ and $\langle \phi^2\rangle$ are obtained via:

$$\langle h_{ij}(x_1)h_{kl}(x_2)\rangle = \int \mathcal{D}h_{mn}|\Psi|^2 h_{ij}(x_1)h_{kl}(x_2)\,; \quad (5.5)$$

and its analogous for ϕ. Therefore, in the regime where the QFT dual correlators are calculable (perturbative field theory $\Leftrightarrow$ non-perturbative gravity dual), one can calculate $\langle h_{ij}h_{kl}\rangle$ and $\langle \phi^2\rangle$ using the duality, i.e. from field theory dual correlators. Recall that, after defining perturbations around a background and considering some redefinitions to construct gauge-invariant quantities, see chapter 1, both the scalar and tensor power spectra come from the momentum space version of $\langle h_{ij}h_{kl}\rangle$. Thus, roughly speaking, equation (5.5) is what one wants to calculate in cosmology to compare against experiments.

5.1 Holographic Phenomenological Approach

Up until now, we have motivated holographic cosmology, through a sort of heuristic argument, by highlighting how it would be useful to compute relevant cosmological correlators in a strongly coupled gravity regime. However, no information was given about the QFT dual to cosmology other than it is supposed to exist, if the holographic principle is to be believed to its full extent. In [3], Maldacena established the correspondence (5.1) but did not attempt to find the field theory dual. Needless to say, finding this dual theory is a hard task. Ideally, one would hope to realize both sides of the duality as different perspectives of the same phenomenon coming from string theory, as in the original AdS/CFT case. In this thesis, we will instead consider the implementation of holography in cosmology based on the phenomenological approach started by McFadden and Skenderis [5, 4].

In their approach, McFadden and Skenderis consider the analytic continuation, Wick rotation, of FLRW cosmology (neglecting the curvature influence, i.e. $k \to 0$ in the FLRW metric) to its Euclidean version, a domain wall spacetime. The holography is described then between this domain wall theory in the bulk and a phenomenological QFT at the 3-dimensional boundary. To calculate the FLRW correlators leading to the power spectra, which is our goal, one then has to Wick rotate back to Lorentzian signature finding a pseudo-QFT (what is meant by "pseudo" will become clear later) as the field theory dual to FLRW cosmology. In the following subsections, we will briefly review the main aspects of this duality.

5.1.1 Domain Wall/Cosmology Correspondence

First, we need to state the correspondence between a cosmological theory, which we will assume to be composed of a metric $g_{\mu\nu}$ plus a single minimally coupled scalar Φ as usual in inflationary models, and a theory in a domain wall spacetime. This correspondence can be generalized to other kinds of cosmologies though, see [4] and the references therein. The action for both theories can be written as:

$$S = \frac{\tilde{\eta}}{2\kappa^2} \int d^4x \sqrt{|g|} \left[-R + g^{\mu\nu} \partial_\mu \Phi \partial_\nu \Phi + 2\kappa^2 V(\Phi) \right] ; \tag{5.6}$$

where $\tilde{\eta} = -1$ for FLRW cosmology and $\tilde{\eta} = +1$ for the domain wall theory. Note that Φ was rescaled to be dimensionless.

The full ansatz, with background fields plus perturbations around them, for the solutions of the equations of motion are considered to be:

$$\begin{aligned} ds^2 =& \tilde{\eta} \left[1 + 2\varphi(z, \vec{x})\right] dz^2 + 2a(z) \left[\partial_i \nu(z, \vec{x}) + \nu_i(z, \vec{x})\right] dz dx^i \\ &+ a^2(z) \left[\delta_{ij} + h_{ij}(z, \vec{x})\right] dx^i dx^j ; \end{aligned} \tag{5.7}$$

$$\Phi = \phi(z) + \delta\phi(z, \vec{x}) ; \tag{5.8}$$

where z is either the time coordinate in FLRW cosmology ($\tilde{\eta} = -1$) or the radial coordinate in the domain wall case ($\tilde{\eta} = +1$) and ν_i is divergenceless. The 3-metric perturbation h_{ij} can be further decomposed as:

$$h_{ij}(z, \vec{x}) = -2\psi(z, \vec{x})\delta_{ij} + 2\partial_i\partial_j\chi(z, \vec{x}) + 2\partial_{(i}\omega_{j)}(z, \vec{x}) + \gamma_{ij}(z, \vec{x})\,; \tag{5.9}$$

where ω_i is divergenceless and so is γ_{ij}, which is also traceless.

Notice that the metric has the same form as in equation (1.81). Moreover, the background fields are found by setting $\varphi = \nu = \psi = \chi = \delta\phi = 0$, $\nu_i = \omega_i = 0$ and $\gamma_{ij} = 0$, thus we are considering a spatially flat FLRW metric for $\tilde{\eta} = -1$ in (5.7). More precisely, since the curvature k contribution to the energy density is negligible as discussed in chapter 1, one can see that k-dependent terms in the Friedmann equations can be dropped which, in turn, means one can simply consider $k \to 0$ in the background FLRW metric (1.14).

The background equations of motion, coming from the action (5.6), are found in [5, 4] to be:

$$\frac{\dot{a}}{a} = -\frac{1}{2}W(\phi)\,; \qquad \dot{\phi} = W'(\phi)\,; \qquad 2\tilde{\eta}\kappa^2 V(\phi) = [W'(\phi)]^2 - \frac{3}{2}[W(\phi)]^2\,; \tag{5.10}$$

where $W = W(\phi)$ is the so-called fake superpotential and it is defined such that $H(z) \equiv \dot{a}/a = -(1/2)W(\phi(z))$. The dot stands for z-derivative throughout this section.

To find the equations of motion for the linear perturbations defined above, one first has to define gauge-invariant fields, as it was discussed in chapter 1. The gauge-invariant scalar is found to be:

$$\zeta = \psi + \frac{H}{\dot{\phi}}\delta\phi \tag{5.11}$$

It turns out the only independent perturbations are ζ and γ_{ij} while the rest can be written in terms of those. Recalling the definition of the slow-roll parameter $\varepsilon_H \equiv -\dot{H}/H^2 =$

$2[W'(\phi)/W(\phi)]^2$, the equations of motion for perturbations become:

$$\ddot{\zeta} + \left(3H + \frac{\dot{\varepsilon}_H}{\varepsilon_H}\right)\dot{\zeta} - \tilde{\eta}\frac{q^2}{a^2}\zeta = 0\,; \tag{5.12}$$

$$\ddot{\gamma}_{ij} + 3H\dot{\gamma}_{ij} - \tilde{\eta}\frac{q^2}{a^2}\gamma_{ij} = 0\,; \tag{5.13}$$

where $\vec{q}$ is the comoving wavevector of the perturbations.

From equations (5.10), (5.12) and (5.13), one can see that the correspondence between the FLRW cosmology and the domain wall theory is obtained by requiring $\bar{\kappa}^2 = -\kappa^2$ and $\bar{q}^2 = -q^2$, where q and κ are FLRW cosmology variables while $\bar{q}$ and $\bar{\kappa}$ are domain wall quantities. By correspondence we mean that the equations of motion are invariant under the analytic continuation.

Actually, while requiring $\bar{\kappa}^2 = -\kappa^2$ is enough, a technical detail makes us need to choose $\bar{q} = -iq$, instead of $\bar{q} = iq$, coming from $\bar{q}^2 = -q^2$. This is needed since one wants the perturbations in the domain wall theory to be suppressed in the interior of the domain wall to have a holographic duality, see [4]. All in all, the relations:

$$\bar{\kappa}^2 = -\kappa^2 \quad ;\text{ and } \quad \bar{q} = -iq\,; \tag{5.14}$$

establish the correspondence up to the linear perturbations level, not only for background fields.

5.1.2 Power Spectra

Having been understood the above correspondence, we will find expressions for the power spectra of cosmological ($\tilde{\eta} = -1$) perturbations which will be useful to relate to QFT dual quantities later on. The power spectra can be written as, see [4]:

$$\Delta_s^2 = \frac{q^3}{2\pi^2}\mathcal{P}_\zeta(q) = \frac{q^3}{2\pi^2}|\zeta_{q(0)}|^2\,; \tag{5.15}$$

$$\Delta_t^2 = 2\Delta_\gamma^2 = \frac{q^3}{\pi^2}\mathcal{P}_\gamma(q) = \frac{2q^3}{\pi^2}|\gamma_{q(0)}|^2\,; \tag{5.16}$$

where γ_{ij} plays the role of h_{ij} of chapter 1 thus we have changed notation $\Delta_h^2 \to \Delta_\gamma^2$ to follow it. Recall that $\mathcal{P}_\zeta \propto \langle\zeta_k\zeta_{-k}\rangle$ and $\mathcal{P}_\gamma \propto \langle\gamma_k\gamma_{-k}\rangle$, see chapter 1. Moreover, $\zeta_{q(0)}$ and $\gamma_{q(0)}$ are the late time values of the perturbations $\zeta_q(z)$ and $\gamma_q(z)$ (defined from $\gamma_{ij} = \gamma_q e_{ij}$, where e_{ij} is some polarization tensor which is constant in the cosmological time), which are frozen outside the horizon, see [4].

One might notice the definition of ζ is distinct from the one of chapter 1. This raises the question of whether we are calculating the same scalar power spectrum as the one introduced in the first chapter or not. Fortunately, the definitions are related indeed, such that the Δ_s^2 above is approximately the same as the one obtained for slow-roll inflation. The definition of chapter 1 was inherited from [33]. There, it is also stated that under slow-roll conditions the gauge-invariant scalar (1.85) of chapter 1, here denoted by $\tilde{\zeta}$ to avoid confusion with (5.11), can be written as:

$$-\tilde{\zeta} \simeq \psi + \frac{H}{\dot{\phi}}\delta\phi \equiv \zeta\,; \tag{5.17}$$

and since Δ_s^2 is proportional to the 2-point function of the momentum space scalar perturbation ζ_q, the overall minus sign does not make a difference.

To canonically quantize the perturbations ζ_q and γ_q one has to find the conjugate momenta associated with them and then impose the canonical commutation relations. The momenta are found in [5, 4] to be:

$$\Pi_q^{(\zeta)} = \frac{2}{\kappa^2}\varepsilon_H a^3\dot{\zeta}_q \quad ; \text{and} \quad \Pi_q^{(\gamma)} = \frac{1}{4\kappa^2}a^3\dot{\gamma}_q \tag{5.18}$$

Imposing the canonical commutation relations then means:

$$\zeta_q \Pi_q^{(\zeta)*} - \Pi_q^{(\zeta)} \zeta_q^* = i \quad ; \text{and} \quad \gamma_q \Pi_q^{(\gamma)*} - \Pi_q^{(\gamma)} \gamma_q^* = \frac{i}{2}\,; \tag{5.19}$$

where the factor of 1/2 in the second relation above comes from the convention used in [5, 4].

The relevant quantities that will be associated with some QFT dual counterparts later are the response functions Ω and E defined from:

$$\Pi_q^{(\zeta)} \equiv \Omega \zeta_q \quad ; \text{and} \quad \Pi_q^{(\gamma)} \equiv E \gamma_q \tag{5.20}$$

The definitions of (5.20) allows us to rewrite the power spectra of equations (5.15) and (5.16) is a simple form. We start by plugging the definitions (5.20) into the canonical commutation relations (5.19):

$$\Omega^* \zeta_q \zeta_q^* - \Omega \zeta_q \zeta_q^* = i \therefore -2i\,\mathrm{Im}(\Omega)|\zeta_q|^2 = i \Rightarrow \Delta_s^2(q) = -\frac{q^3}{4\pi^2\,\mathrm{Im}[\Omega_{(0)}(q)]}\,; \tag{5.21}$$

$$E^* \gamma_q \gamma_q^* - E \gamma_q \gamma_q^* = \frac{i}{2} \therefore -4i\,\mathrm{Im}(E)|\gamma_q|^2 = i \Rightarrow \Delta_t^2(q) = -\frac{q^3}{2\pi^2\,\mathrm{Im}[E_{(0)}(q)]}\,; \tag{5.22}$$

where $\Omega_{(0)}$ and $E_{(0)}$ are the response functions values associated to $\zeta_{q(0)}$ and $\gamma_{q(0)}$ respectively.

We have computed the power spectra in FLRW cosmology. It is shown in [4] that one can rewrite these power spectra in terms of domain wall response functions $\bar{\Omega}$ and $\bar{E}$ in the same fashion:

$$\Delta_s^2(q) = -\frac{q^3}{4\pi^2\,\mathrm{Im}[\bar{\Omega}_{(0)}(-iq)]} \quad ; \text{and} \quad \Delta_t^2(q) = -\frac{q^3}{2\pi^2\,\mathrm{Im}[\bar{E}_{(0)}(-iq)]}\,; \tag{5.23}$$

as long as we define them as:

$$\bar{\Pi}_{\bar{q}}^{(\zeta)} = \frac{2}{\bar{\kappa}^2}\varepsilon_H a^3 \dot{\zeta}_{\bar{q}} \equiv -\bar{\Omega}\zeta_{\bar{q}} \quad ; \text{and} \quad \bar{\Pi}_{\bar{q}}^{(\gamma)} = \frac{1}{4\bar{\kappa}^2} a^3 \dot{\gamma}_{\bar{q}} \equiv -\bar{E}\gamma_{\bar{q}}\,; \tag{5.24}$$

where the minus signs in the definitions of $\bar{\Omega}$ and $\bar{E}$ ensure $\bar{\Omega}(-iq) = \Omega(q)$ and $\bar{E}(-iq) = E(q)$ through the analytic continuation from FLRW cosmology to the domain wall theory.

We will aim to compute the above power spectra from QFT dual quantities. Therefore, since the holography is well-understood for asymptotically AdS or power-law domain walls (corresponding to asymptotically dS or power-law cosmologies respectively), one needs to rewrite $\bar{\Omega}_{(0)}$ and $\bar{E}_{(0)}$ in terms of variables of the QFTs which are dual to those domain walls. It will not be proven here, see [4] and its references for a complete discussion, but it turns out that the relevant quantity to be related to the domain wall response functions is the boundary theory stress-energy tensor 2-point function. This correlator can be written in term of generic $A = A(\bar{q})$ and $B = B(\bar{q})$ functions as:

$$\langle T_{ij}(\bar{q}) T_{kl}(-\bar{q}) \rangle = A(\bar{q}) \Pi_{ijkl} + B(\bar{q}) \pi_{ij} \pi_{kl} \,; \tag{5.25}$$

where $\Pi_{ijkl} \equiv \frac{1}{2}(\pi_{ik}\pi_{lj} + \pi_{il}\pi_{kj} - \pi_{ij}\pi_{kl})$ is the transverse traceless projection operator and $\pi_{ij} \equiv \delta_{ij} - \bar{q}_i \bar{q}_j / \bar{q}^2$.

The fact that the energy-momentum tensor 2-point function is the relevant quantity to be calculated in the QFT dual is not a surprise, since the bulk metric couples to it via the holographic map. However, a highly non-trivial issue is how the above A and B functions relate to the domain wall response functions. Again, these relations are found in [4] to be:

$$A(\bar{q}) = 4\bar{E}_{(0)}(\bar{q}) \quad ; \text{and} \quad B(\bar{q}) = \frac{1}{4}\bar{\Omega}_{(0)}(\bar{q}) \tag{5.26}$$

Finally, we are able to write the power spectra in terms of the holographic functions $A(\bar{q})$ and $B(\bar{q})$. From (5.23), we get:

$$\Delta_s^2(\bar{q}) = -\frac{(i\bar{q})^3}{4\pi^2 \,\mathrm{Im}[\bar{\Omega}_{(0)}(\bar{q})]} = -\frac{(i\bar{q})^3}{16\pi^2 \,\mathrm{Im}[B(\bar{q})]} \,; \tag{5.27}$$

$$\Delta_t^2(\bar{q}) = -\frac{(i\bar{q})^3}{2\pi^2 \,\mathrm{Im}[\bar{E}_{(0)}(\bar{q})]} = -\frac{2(i\bar{q})^3}{\pi^2 \,\mathrm{Im}[A(\bar{q})]} \,; \tag{5.28}$$

which should be Wick rotated back to the FLRW momentum $q = i\bar{q}$ to compare against CMBR observations.

A crucial point to be made is that we have been assuming a reliable description for cosmology in terms of the metric and a scalar field at early times (before the start of the standard radiation-dominated period), in close analogy to the inflationary case. However, while this is a needed assumption for inflation, it can be dropped in this holographic construction. At early times the gravity theory may be strongly coupled and therefore not admit a description in terms of low energy fields. In this case, the QFT dual would be weakly coupled at early times. This means one would be able to calculate $A(\bar{q})$ and $B(\bar{q})$ in the field theory dual to the domain wall and then analytically continue the result to obtain the cosmological power spectra above. This means the holographic cosmology paradigm is broader than the inflationary one.

5.1.3 Phenomenological (Pseudo-)QFT Dual

We are ready to define the QFT dual to the domain wall. But first, allow us to justify the title of this subsection. Since the holography which is actually well-defined is between the domain wall theory and a QFT at its boundary, to get the field theory dual to FLRW cosmology we need to perform the analytic continuation defined in the last section on both sides of the duality. The issue is that the analytic continuation on the QFT side of it corresponds to a change not only in momentum $\bar{q} = -iq$ but also in the rank of the QFT gauge group in the following fashion: $\bar{N}^2 = -N^2$. This means that, considering a real rank $\bar{N}$ for the QFT dual to the domain wall, the gauge group of the field theory dual to our cosmological theory has an imaginary rank. Although it is certainly a weird feature, this can simply be understood as an operational definition of the field theory dual to FLRW cosmology, in the sense that all calculations are done in the well-defined QFT dual to the domain wall theory and, only after, we perform an analytic continuation defined by $\bar{q} = -iq$ and $\bar{N}^2 = -N^2$.

The goal of the phenomenological approach taken by McFadden and Skenderis was to construct a generic superrenormalizable gauge theory on $d = 3$ dimensions to the dual to the domain wall. The theory also has to admit a large $\bar{N}$ limit. They proposed [4, 77, 78] a $SU(\bar{N})$ Yang-Mills theory with $\mathcal{N}_s$ scalars ϕ^M ($M = 1, \ldots, \mathcal{N}_s$), $\mathcal{N}_f$ spin-1/2 fermions ψ^L ($L = 1, \ldots, \mathcal{N}_f$) and a gauge field $A_i = A_i^a T_a$ (with field strength F_{ij}). All the fields transform in the adjoint of $SU(\bar{N})$. For the interactions, both Yukawa and ϕ^4 interactions were considered for reasons to be clear shortly. In summary, the full phenomenological action is written as:

$$
\begin{aligned}
S_{QFT} = &\int d^3x \, \mathrm{Tr}\left(\frac{1}{2}\tilde{F}_{ij}\tilde{F}^{ij} + \delta_{M_1M_2}\tilde{D}_i\tilde{\phi}^{M_1}\tilde{D}^i\tilde{\phi}^{M_2} + 2\delta_{L_1L_2}\bar{\tilde{\psi}}^{L_1}\tilde{\not{D}}\tilde{\psi}^{L_2}\right. \\
&\left. + \sqrt{2}\tilde{\mu}_{ML_1L_2}\tilde{\phi}^M\bar{\tilde{\psi}}^{L_1}\tilde{\psi}^{L_2} + \frac{1}{6}\tilde{\lambda}_{M_1M_2M_3M_4}\tilde{\phi}^{M_1}\tilde{\phi}^{M_2}\tilde{\phi}^{M_3}\tilde{\phi}^{M_4}\right) \\
= &\frac{1}{g_{YM}^2}\int d^3x \, \mathrm{Tr}\left(\frac{1}{2}F_{ij}F^{ij} + \delta_{M_1M_2}D_i\phi^{M_1}D^i\phi^{M_2} + 2\delta_{L_1L_2}\bar{\psi}^{L_1}\not{D}\psi^{L_2}\right. \\
&\left. + \sqrt{2}\mu_{ML_1L_2}\phi^M\bar{\psi}^{L_1}\psi^{L_2} + \frac{1}{6}\lambda_{M_1M_2M_3M_4}\phi^{M_1}\phi^{M_2}\phi^{M_3}\phi^{M_4}\right);
\end{aligned}
\tag{5.29}
$$

where the trace is over gauge group indices and we consider the normalization convention $\mathrm{Tr}\big(T^aT^b\big) = \frac{1}{2}\delta^{ab}$. From the first to the second equality we have redefined the tilde fields plus the coupling constants $\tilde{\mu}_{ML_1L_2}$ and $\tilde{\lambda}_{M_1M_2M_3M_4}$ such that the Yang-Mills coupling $1/g_{YM}^2$ could become a common factor.

The action (5.29) has a so-called generalized conformal structure, which means that one can promote the Yang-Mills coupling to another field of the theory in such a way as to construct a conformal field theory. This is closely related to the usefulness of assigning the usual 4-dimensional mass dimensions for the fields of the theory, yielding $[g_{YM}] = 1/2$ in the second expression above, see [77]. Note that, to write a generalized conformal invariant action, one is restricted to the Yukawa and ϕ^4 interactions since their coefficients are dimensionless in 4 dimensions, hence explaining the absence of mass terms, for instance.

In the AdS/CFT correspondence section of chapter 3 it was claimed that, in the large N limit of a $SU(N)$ gauge theory, the effective coupling is actually $g^2_{YM}N$, see [63]. In the current case, the Yang-Mills coupling is not dimensionless, therefore the dimensionless effective coupling is found to be:

$$g^2_{eff} \equiv \frac{g^2_{YM}\bar{N}}{\bar{q}}\,; \tag{5.30}$$

where $\bar{q}$ is the relevant quantity in the domain wall/phenomenological QFT duality, the field theory energy scale, which has unity dimension of mass and hence is the natural factor to be introduced to neutralize the Yang-Mills coupling dimensionality.

We can then further rewrite the power spectra of (5.27) and (5.28) in terms of functions of g^2_{eff}, $f(g^2_{eff})$ and $f_T(g^2_{eff})$, by noticing how the momentum space stress-energy tensor 2-point correlator depends on $\bar{q}$ and the rank of the gauge group. First, note that $[\langle T_{ij}(\bar{q})T_{kl}(-\bar{q})\rangle] = 3$, since $T_{ij}(x) \propto \frac{\partial\mathcal{L}}{\partial g^{ij}} \Rightarrow [T_{ij}(x)] = 3$ in 3 dimensions and $\langle T_{ij}(\bar{q})T_{kl}(-\bar{q})\rangle \propto \int d^3r \langle T_{ij}(x)T_{kl}(x+r)\rangle e^{ik\cdot x} \Rightarrow [\langle T_{ij}(\bar{q})T_{kl}(-\bar{q})\rangle] = [\langle T_{ij}(x)T_{kl}(x+r)\rangle] - 3 = 3$, therefore both A and B functions have to be proportional to $\bar{q}^3$, generically. Furthermore, the leading contribution to $\langle T_{ij}(\bar{q})T_{kl}(-\bar{q})\rangle$ comes from 1-loop diagrams [4], hence it should be proportional to $\bar{N}^2$. Finally, one can define the dimensionless functions as:

$$A(\bar{q}) \equiv \bar{q}^3\bar{N}^2 f_T(g^2_{eff}) \quad ;\text{ and}\quad B(\bar{q}) \equiv \frac{1}{4}\bar{q}^3\bar{N}^2 f(g^2_{eff}) \tag{5.31}$$

Plugging the above results into equations (5.27) and (5.28) we get:

$$\Delta^2_s(\bar{q}) = -\frac{(i\bar{q})^3}{4\pi^2\,\mathrm{Im}[\bar{q}^3\bar{N}^2 f(g^2_{eff})]} \quad ;\text{ and}\quad \Delta^2_t(\bar{q}) = -\frac{2(i\bar{q})^3}{\pi^2\,\mathrm{Im}[\bar{q}^3\bar{N}^2 f_T(g^2_{eff})]} \tag{5.32}$$

The last step is to Wick rotate back to cosmological variables q and N. From the previous relations $\bar{q} = -iq$ and $\bar{N}^2 = -N^2$, one gets that $\bar{q}^3\bar{N}^2 = -iq^3N^2$. Moreover, by choosing $\bar{N} = -iN$, from $\bar{N}^2 = -N^2$, we ensure g^2_{eff} is invariant under the analytic continuation. For real functions f and f_T, the dimensionless power spectra become:

$$\Delta_s^2(q) = \frac{1}{4\pi^2 N^2 f(g_{eff}^2)} \quad ; \text{ and } \quad \Delta_t^2(q) = \frac{2}{\pi^2 N^2 f_T(g_{eff}^2)} \tag{5.33}$$

If we consider a non-geometric (strongly coupled) cosmological theory, unlike the standard case of inflation plus ΛCDM cosmology, then we can calculate f and f_T using perturbation theory in the QFT dual. Up to 2-loops, these functions assume the generic form, see [77]:

$$f(g_{eff}^2) = f_0 \left[1 - f_1 g_{eff}^2 \log\left(g_{eff}^2\right) + f_2 g_{eff}^2 + \mathcal{O}(g_{eff}^4)\right] \; ; \tag{5.34}$$

$$f_T(g_{eff}^2) = f_{T0} \left[1 - f_{T1} g_{eff}^2 \log\left(g_{eff}^2\right) + f_{T2} g_{eff}^2 + \mathcal{O}(g_{eff}^4)\right] \; ; \tag{5.35}$$

where f_0 and f_{T0} come from 1-loop diagrams while the rest of the prefactors are obtained from 2-loop calculations.

By setting the renormalization group (RG) scale μ to be equal to a scale q_* of reference, a pivot scale, and defining some new variables g, β, g_T and β_T:

$$g q_* \equiv f_1 g_{YM}^2 N = f_1 g_{eff}^2 q \quad ; \quad \log(\beta) \equiv -\frac{f_2}{f_1} - \log|f_1| \, ; \tag{5.36}$$

$$g_T q_* \equiv f_{T1} g_{YM}^2 N = f_{T1} g_{eff}^2 q \; ; \; \log(\beta_T) \equiv -\frac{f_{T2}}{f_{T1}} - \log|f_{T1}| \, ; \tag{5.37}$$

to make the power spectra momentum dependence more evident, one finds the expressions:

$$\begin{aligned}
\Delta_s^2(q) &= \left(\frac{1}{4\pi^2 N^2 f_0}\right) \left[1 - f_1 g_{eff}^2 \log\left(g_{eff}^2\right) - f_1 g_{eff}^2 \log(\beta) - f_1 g_{eff}^2 \log|f_1| + \mathcal{O}(g_{eff}^4)\right]^{-1} \\
&= \left(\frac{1}{4\pi^2 N^2 f_0}\right) \left[1 - f_1 g_{eff}^2 \log\left(\frac{g q_*}{f_1 q}\right) - f_1 g_{eff}^2 \log|\beta f_1| + \mathcal{O}(g_{eff}^4)\right]^{-1} \\
&= \left(\frac{1}{4\pi^2 N^2 f_0}\right) \left\{1 + \left(g\frac{q_*}{q}\right) \log\left|\frac{1}{\beta g}\frac{q}{q_*}\right| + \mathcal{O}\left[\left(g\frac{q_*}{q}\right)^2\right]\right\}^{-1} \; ;
\end{aligned} \tag{5.38}$$

and the analogous for $\Delta_t^2(q)$:

$$\Delta_t^2(q) = \left(\frac{2}{\pi^2 N^2 f_{T0}}\right) \left\{1 + \left(g_T \frac{q_*}{q}\right) \log\left|\frac{1}{\beta_T g_T}\frac{q}{q_*}\right| + \mathcal{O}\left[\left(g_T \frac{q_*}{q}\right)^2\right]\right\}^{-1} \tag{5.39}$$

Notice that the above power spectra have a different form than the usually assumed power-law spectra (1.92), for a (nearly) constant tilt, and (1.95), which are approximately reproduced by slow-roll inflation, see [12] chapter 11. Therefore, one has to check against experimental data whether the obtained scalar power spectrum can reproduce observations as well as slow-roll inflation or not. It turns out, see next subsection, that holographic cosmology, assuming a perturbative QFT dual, can indeed reproduce observations with some caveats. Was it not the case, this type of holographic model would have to be dropped.

5.1.4 Fitting Observations: An Alternative to Inflation

In this subsection, we will briefly highlight a few relevant observational implications on the model introduced above. A complete analysis is found in [77, 78].

First, recall from chapter 1 that the inflation plus ΛCDM model fits CMBR data with only six parameters. Two of those, $\Delta_s^2(k_*)$ and n_s from equations (1.91) and (1.92), constrain the form of the scalar power spectrum. In the holographic model studied here, we need three parameters, $\Delta_0^2 \equiv (4\pi^2 N^2 f_0)^{-1}$, g and $\log(\beta)$ (or equivalently f_0, g and β), to constrain the spectrum form. Therefore, in order to compare both models, [78] promoted the running of the spectral index, the α_s of (1.91), to a seventh parameter of the inflation plus ΛCDM model. It was found that at low multipoles $l \lesssim 30$, where l was defined in equation (1.93), the QFT dual becomes non-perturbative and the power spectrum expression (5.38) cannot be trusted anymore. Nevertheless, for $l > 30$, the holographic form for the power spectrum fits CMBR data as well as the power law form. Therefore, holographic models of this type can be seen as an alternative to inflation.

So far we have been just analyzing the power spectra forms. Nothing was said about ruling out specific realizations of the phenomenological QFT proposed by McFadden and

Skenderis, equation (5.29). Through a QFT calculation, the parameters f_0, g and β where calculated in [78] in terms of $\mathcal{N}_s$ and $\mathcal{N}_f$. Experimental constraints demand a non-minimal gravity coupling, ξ_M, with the scalars ϕ^M. It turns out that models with just fermions and the gauge field do not yield a good experimental fit and thus have to be dropped. On the other hand, for a large enough number of scalars ($\mathcal{N}_s \gg \mathcal{N}_f$), yet $N^2 \gg \mathcal{N}_s$ as required for large N calculations, one can indeed fit CMBR data.

5.2 Solutions to Pre-Inflationary Problems

Last section we have seen that the phenomenological holographic approach by McFadden and Skenderis, but also more generally the whole holographic cosmology setup, can reproduce the power spectrum of scalar perturbations obtained from CMBR observations. Reproducing these observations is perhaps the main triumph of inflation. However, the original motivations for inflation were the pre-inflationary problems of Hot Big Bang cosmology as we saw in chapter 1. Therefore, to fully be considered an alternative to inflation, holographic cosmology should solve the same problems. How the holographic scenario solves such problems will be outlined in this section based on [79, 80].

First, recall that inflation solves most of the problems by the same mechanism: a fast (nearly exponential) expansion of the fabric of space which sent originally interacting scales outside the horizon only to get back inside it just recently, during the standard radiation or matter-dominated eras. Three concrete examples of pre-inflationary problems were discussed in chapter 1: the flatness problem, the horizon problem, and the monopole/relic problem. The first of them lies in the weird fact that not only we observe a nearly flat space today, i.e. $|\Omega^{(0)} - 1| \ll 1$, but this flatness would have been even more severe in the past, i.e. $|\Omega(t) - 1| \lll 1$ for $t < t_0$. This implies a fine-tuning of initial conditions. Inflation solves this problem since the exponential expansion of space stretches its fabric making any local region nearly flat. Thus, it doesn't matter how curved the space at the beginning of the universe could have been, inflation can always make it as flat as needed with an appropriate number of e-folds. The horizon problem is the fact that the space is nearly homogeneous and isotropic even taking into account regions that are observed in opposite directions in the sky. Such regions were not supposed to ever have been in causal contact within pre-inflationary cosmology, so we wouldn't have a mechanism to explain such correlation over large distances. Inflation's abrupt expansion allows for those regions to had been in causal contact before inflation started, thus solving the problem. The monopole/relic problem lies in the fact that we don't observe cosmological relics in the Hot Big Bang cosmology expected concentration. Some mechanism must have diluted the density of relics per density of nucleons ratio to match observed bounds. Inflation plus reheating compose such a mechanism. The rapid inflationary expansion dilutes both the relics' and nucleons' densities but reheating produces radiation which ends up creating more nucleons. The net behavior is the expected decrease in the relic density per nucleon density ratio.

Our goal is to analyze how these problems are solved in a non-geometrical gravity regime within holographic cosmology, more specifically considering the phenomenological model introduced in the first section. In a weakly coupled gravity regime, the setup would be the same as the inflationary one and we already know the solutions to pre-inflationary problems in this case. That is why we focus on the new case of interest made viable (calculable) by holography.

First, to understand the new holographic solutions to the above-mentioned problems one needs to recall from the last section of chapter 3 that, in general, in a gauge/gravity duality the field theory energy scale E relates to the radial (holographic) coordinate r of the gravity dual as $E \propto r/\alpha'$. In the case of domain wall/phenomenological QFT duality this means:

$$\bar{q} \simeq \frac{z}{\alpha'} \, ; \tag{5.40}$$

for $\tilde{\eta} = +1$ in the metric (5.7). From equation (5.30), this means that as the radial coordinate z increases the QFT becomes weakly coupled, $g_{eff}^2 \ll 1$. The important information to be kept in mind is the analytically continued version of the last sentence: time evolution in cosmology means going from infrared (IR) to ultraviolet (UV) in the dual QFT, or equivalently going from strongly to weakly coupled QFT regime.

Following the pattern of chapter 1, we will discuss the solution to each problem separately:

1. The flatness problem

   In the solution to the flatness problem we hope to find a mechanism, in the field theory side of the duality, which can explain how one gets $\Omega - 1 \lll 1$ at late times, in the UV of the QFT, starting from $\Omega - 1 \sim \mathcal{O}(1)$ in the field theory IR. As usual, one can write the spatial metric as $g_{ij} = \delta_{ij} + h_{ij}$ and the field h_{ij} quantifies the deviation from flat space. In holography, this field is supposed to couple to the stress-energy tensor T_{ij} of the field theory dual. Therefore, from the QFT point of view, h_{ij} is a coupling that should run with the running of the QFT energy scale, i.e. run with the RG flow. We expect h_{ij} to get suppressed from IR to UV (as time evolves) to solve the problem. This is obtained if T_{ij} is a relevant operator. Recall that the renormalization procedure relates a theory in a given scale to its version at another scale. A relevant operator of a microscopic theory is one that is important to macroscopic physics or, in other words, one whose effects are enhanced by renormalization as the analyzed length scale grows large (energy scale decreases), see [35]. Thus having a relevant operator T_{ij}, whose effects become irrelevant from IR to UV of the field theory, means that the coupling h_{ij} is suppressed over time, which explains the apparent fine-tuning.

   The implications of requiring a relevant T_{ij} on the phenomenological QFT are yet to be discussed. First, note that, as already discussed, the classical dimension of T_{ij} makes it a marginal operator (mass dimension 3 in $d = 3$ space dimensions). Therefore, we need to analyze its 2-point function to figure out whether it is marginally relevant (which still solves the problem) or not. This analysis is done in [79, 80], where it was found that if $f_1 < 0$ in (5.34) then T_{ij} is marginally relevant. The bound $f_1 < 0$ plays the role of a constraint on proposed models of the type (5.29).

2. The horizon problem

At first sight, the fact that we are considering a non-geometrical phase of the universe seems to make it difficult to define this problem. It is hard to define causal contact without a precise definition of past light-cones, coming from the nonexistence of a low energy effective metric. Nevertheless, one must look at the QFT dual and seek an explanation for the CMBR temperature correlation between two far apart points in terms of field theory operators.

In this case, again, the observable (correlator over the sky) of interest comes from h_{ij}, as defined in the flatness problem discussion. Recall that, after some redefinitions, the observed temperature correlation is connected to 2-point functions of the metric perturbation h_{ij} through the so-called transfer function, see chapter 1. Thus, the QFT quantity related to $\langle h_{ij}h_{kl}\rangle$ is again the energy-momentum tensor correlator $\langle T_{ij}T_{kl}\rangle$. The solution to the problem then lies in the fact that the dual QFT correlation functions in different scales are related through the running of its energy scale or, in other words, they are related by the RG flow. Therefore, $\langle T_{ij}T_{kl}\rangle$ is non-zero at small scales it will be non-zero at large scales too, explaining the non-vanishing h_{ij} correlators at large scales. The only issue one must check is whether the QFT is well-defined in all the range of energy scales that are covered, which leads to a large range of g^2_{eff} values. Fortunately, the class of theories considered in (5.29) is IR finite, see [79, 80] and the references therein, and therefore one can always solve the horizon problem.

Recall that both the flatness and the horizon problems led to the same constraint on the number of e-folds N_e in the inflationary case. In the current non-geometric cosmology case, one must figure out which QFT quantity will be analogously constrained, since cosmological observables are simply inherited from dual field theory calculations. It is clear from the analysis of the last paragraph that a minimum amount of change in scales, which is what the N_e bound really means, should translate into a minimum amount of RG flow. Moreover, the anticipated IR finiteness of the phenomenological model ensures such bound can always be respected, unlike in inflation whose formulation breaks down for a sufficiently large number of e-folds, see chapter 4's introductory remarks.

3. The monopole/relic problem

   We saw in chapter 1 that the monopole problem was conceptually the same as other relics problems. Only the constraint was tighter for monopoles due to direct searches for it on Earth, so we could effectively forget about other relics since the monopole dilution mechanism automatically satisfies other bounds. Within holographic cosmology, this scenario changes and the monopole problem deserves to be treated separately.

   For general relics, about which we don't assume anything other than it has gravitational effects that must be diluted, the problem is similar to the one already solved in the last two problems. One considers these relics as perturbations in the bulk metric which is related to $\langle T_{ij} T_{kl} \rangle$ via holography. The desired dilution effect is realized in the QFT by the fact that T_{ij} should be (marginally) relevant. Furthermore, since the experimental constraint for the flatness and horizon problems is tighter than the bound on relics densities (see chapter 1), by requiring the former to be respected this one will also automatically be, see [80].

   Monopoles also generate gravitational effects which should be diluted in the same manner as other relics. Again, this implies a lower bound on the amount of RG flow needed (actually inverse RG flow, i.e. from IR to UV) which is less stringent than in the one from flatness and horizon problems. However, we must also dilute the magnetic effects of monopoles. These effects are seen as fluctuations of a bulk gauge field A^a_μ which is associated (coupled) to a global current j^a_i of the boundary theory. Recall that local bulk symmetries must correspond to global symmetries at the boundary, in the same fashion as it was shown in chapter 3 for the original AdS/CFT setup. Thus, diluting magnetic effects of monopoles in the bulk corresponds to having a (marginally) relevant magnetic current $\tilde{j}^a_i$, which is the electric/magnetic dual of the Noether electric current

j_i^a, for the same reason we have required T_{ij} to be marginally relevant in the above discussions.

It has been shown in [80] that for a simple toy model of the (5.29) type, quite generically, one obtains a marginally irrelevant electric current which implies a marginally relevant magnetic $\tilde{j}_i^a$ as desired. The toy model was proposed since one cannot define j_i^a generally, i.e. a specific gauge symmetry in the bulk has to be proposed first, and then one can calculate correlators of its gauge field source j_i^a. This is not proof that one always obtains a relevant magnetic current, since the calculations were based on the specific toy model, but rather an example that such construction is possible and quite independent of the scalar potential form. In the end, requiring $\tilde{j}_i^a$ to be marginally relevant is a restriction to the set of QFT models aiming to reproduce observations, just like in the T_{ij} case.

Let us give a taste of the calculations found in [80] which were used to show that the magnetic current is marginally relevant for the given toy model. The model is defined to be composed of 6 complex scalars and the gauge field only, recall that to fit CMBR data one needs more scalars than fermions so, for simplicity, the model admits no fermions. The six scalars are represented as ϕ_i^a, where a is an index in the adjoint of $SO(3)$ $(a = 1, 2, 3)$ and $i = 1, 2$. The only allowed scalar interaction was seen to be of the ϕ^4 type to preserve generalized conformal symmetry. Then, the toy model considers a potential $V = 4\lambda \operatorname{Tr}\left(|\vec{\phi}_1 \times \vec{\phi}_2|^2\right)$. In summary, the full QFT action in Minkowski space is:

$$S = \int d^3x \operatorname{Tr}\left(-\frac{1}{2}F_{\mu\nu}F^{\mu\nu} - 2\sum_{i=1}^{2}|D_\mu\vec{\phi}_i|^2 - 4\lambda|\vec{\phi}_1 \times \vec{\phi}_2|^2\right) ; \tag{5.41}$$

where the trace is over $SU(N)$ adjoint indices. Note that in this discussion we will follow the notation of the original paper and consider greek indices running from 1 to 3, to avoid confusion with middle-alphabet roman indices, e.g. i, which stand for one of the two sets of scalars.

After writing the Euclidean version of the above action, one can show that Noether current associated with the global $SO(3)$ symmetry can be written as:

$$j_\mu^a = \sum_{i=1}^{2} \vec{\phi}_i^* T^a D_\mu \vec{\phi}_i + h.c. = \sum_{i=1}^{2} i\epsilon^{abc} \phi_{b,i}^* D_\mu \phi_{c,i} + h.c.\,; \tag{5.42}$$

where in the second equality we have used the fact that the scalars are in the adjoint of $SO(3)$ and $(T_a)_{bc} = i f_{abc} = i\epsilon_{abc}$.

With the above expression for the electric current at hand, one can calculate $\langle j_\mu^a j_\nu^b \rangle$. In the large N limit, this is done up to 2-loops, i.e. up to $\sim \mathcal{O}(g_{eff}^2)$, in [80]. One finds out whether the operator is relevant by calculating its anomalous dimension δ. Strictly speaking, such quantity is defined for conformal field theories, but as we approach the UV of the theory, i.e. as $g_{eff}^2 \to 0$, the theory is approximately conformal (see [79]) so we will define it anyway by the proportionality:

$$\langle j_\mu^a(p) j_\nu^b(-p) \rangle \propto N^2 \pi_{\mu\nu} p^{1+2\delta} \simeq N^2 \pi_{\mu\nu} p \left[1 + 2\delta \log(p) + \mathcal{O}(\delta^2)\right] \;; \tag{5.43}$$

where $\pi_{\mu\nu} \equiv \delta_{\mu\nu} - \frac{p_\mu p_\nu}{p^2}$ is the transverse projector.

By matching against the Feynman diagrams calculations, up to 2-loops, it was found that $\delta \simeq \frac{2}{\pi^2} g_{eff}^2 > 0$, which implies an irrelevant electric current. After some rather involving argumentation, [80] concludes that the magnetic current $\tilde{j}_\mu^a$ shows the opposite behavior of its S-dual related counterpart j_μ^a, i.e. $\delta(\tilde{j}) < 0$ leading to a relevant magnetic current operator. As anticipated, this solves the monopole problem.

To end the discussion about the monopole problem, it is important to remark that a recent work [81] has shown that the desired behavior for the magnetic current (more precisely $\delta(j) > 0$ for the electric current, which is argued to yield $\delta(\tilde{j}) = -\delta(j) < 0$) is obtained for a broader set of models. It is obtained for any purely bosonic phenomenological model, not only ones of the toy model type (5.41).

The solution to other pre-inflationary problems are also found in [79, 80]. Here, it was enough to explain the above ones to compare with their inflationary solutions described in chapter 1.

5.3 Insight on the Cosmological Constant Problem

In the first section, the holographic cosmology setup was introduced as an alternative to the inflationary scenario. The last section only reinforced this point. We have even argued that using holography we can drop the usual inflationary assumption of having a perturbative gravitational theory at early times such that one obtains a low energy effective theory in terms of the metric plus a scalar field. Nonetheless, the insights this setup gives us on dark energy were not yet discussed. In this section, we will show that within holographic cosmology one can map the cosmological constant problem to a simple field theory property, namely the dimensional RG flow, see [6].

The scenario that will be considered here is the following: at very early times, analogous to the beginning of inflation, we will assume the approximate cosmological constant is indeed as high as the theoretically expected value for it and just later, at times analogous to the end of inflation, we wind up with the low observed value for it ($\sim 10^{-120}$ of the original value in the most severe case of the cosmological constant problem, see chapter 2). After that, the universe is supposed to go through some transition phase to start the usual radiation-dominated era, analogous to reheating for inflation. This transition phase is studied in [82], but it will not be our focus here. The important point is that the approximate cosmological constant must remain the same from the beginning of this transition phase until today when we observe the low dark energy density value discussed in chapter 2.

First, let us start our analysis by considering a geometric phase for the bulk cosmological theory. Under this circumstance, the theory is described by the metric $g_{\mu\nu}$ plus a scalar field.

This corresponds to very early times in the cosmological evolution since we saw that the dual QFT energy scale increases over time, which leads to a decreasing g_{eff}^2, so the field theory is strongly coupled for early times while the gravity dual is perturbative. As pointed out in the last paragraph, we assume an as-large-as-expected cosmological constant at early times to try to dilute its effects later on, by a similar dilution mechanism as the one described in the last section. Therefore, the equations of motion are the Einstein equations (1.6) with the cosmological constant term. For a generic d-dimensional theory, the trace of such equation is:

$$R - \frac{d}{2}R + d\Lambda = \frac{T}{M_{Pl}^2} \therefore \frac{\Lambda}{M_{Pl}^2} = \frac{1}{d}\frac{T}{M_{Pl}^4} + \frac{d-2}{2d}\frac{R}{M_{Pl}^2} \therefore \frac{\Lambda}{M_{Pl}^2} \simeq \frac{d-2}{2d}\frac{R}{M_{Pl}^2}\,; \tag{5.44}$$

where $T \equiv g^{\mu\nu}T_{\mu\nu}$ and it was dropped in the last equation since at very early times the cosmological constant term is expected to dominate. For our case of interest, $d = 4$, we have $\Lambda \lesssim R$.

Recall now that holographic cosmology can generate a nearly flat late time universe from a natural $\Omega - 1 \sim \mathcal{O}(1)$ at very early times by requiring T_{ij} to be (marginally) relevant or, equivalently, $f_1 < 0$. This means $R/M_{Pl}^2 \to 0$ as we go to the UV of the field theory or, in other words, as $g_{eff}^2 \to 0$. Therefore, it is assumed in [6] a relation $R/M_{Pl}^2 \sim g_{eff}^p$, with $p > 0$. This is assumed to be an approximate relation rather than just a proportionality since in the AdS/CFT original case this is the relation one gets. However, in that case, $p = -1/2$, which is not allowed in the domain wall/phenomenological QFT duality case by the argument of this paragraph. Combining this assumption with the result $\Lambda \lesssim R$ one gets:

$$\frac{\Lambda}{M_{Pl}^2} \lesssim g_{eff}^p \equiv \left(\frac{g_{YM}^2 N}{q}\right)^p \tag{5.45}$$

Therefore, as time evolves and the QFT scale flows from the IR to the UV, the cosmological constant has to shrink down to small values. Since the phenomenological QFT seems to be well-defined in the IR, as previously discussed, one can always force a decrease in the cosmological constant value (even an order 10^{-120} decrease) by requiring a deep enough (inverse) RG flow.

In [6], two concrete examples which yields a bound like (5.45) with $p > 0$ are studied, in order to show that the assumption $R/M_{Pl}^2 \sim g_{eff}^p$ is viable.

It is important to point out that, although the cosmological constant problem would be holographically solved in this way, i.e. we would be able to explain the current low observed value for the dark energy density, this is not a full solution to the problem. The precise mechanism in the gravitational theory which would dilute the cosmological constant effects is yet unknown. Another observation is that the $\Lambda \lesssim R$ relation was obtained by neglecting the trace of the energy-momentum tensor. As Λ decreases one has to be careful to analyze whether this assumption is still valid. Moreover, as we approach the UV of the field theory, corresponding to a non-geometric bulk theory, we would be extrapolating the region where Λ and most certainly the Ricci scalar R are well-defined. Thus we would be assuming $\Lambda \lesssim R$ still makes sense in the unknown non-geometric gravity theory.

Chapter 6

Chameleon Cosmology

The last type of cosmology in which dark energy will be analyzed in this thesis is the chameleon cosmology. Chameleon scalars were introduced by Khoury and Weltman [7, 8] as fields that would effectively have a large mass in regions with high local matter densities but would be light at cosmological scales. This property allows us to avoid violating current fifth force experimental constraints on Earth. The chameleon setup is actually string-independent, but it is very useful to models coming from string theory since massless scalars, i.e. moduli, naturally arise in them, as discussed in chapter 3, and the chameleon mechanism represents a way of making them consistent with observations without having to stabilize them, i.e. give them a nonperturbative potential, which can be challenging.

6.1 General Chameleon Idea

The way such scalar fields acquire an effective mass that is dependent on local matter density is through a non-canonical coupling of matter with the metric. Namely:

$$\tilde{g}_{\mu\nu} = A^2(\phi) g_{\mu\nu} \, ; \tag{6.1}$$

where $g_{\mu\nu}$ is the Einstein frame metric and $A = A(\phi)$ is the conformal factor (actually its square root but from now on we will refer to A as the conformal factor) needed to achieve the desired chameleon properties.

The full action for a generic chameleon model is:

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{Pl}^2}{2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right] + \int d^4x\mathcal{L}_m(\tilde{g}, \psi) \tag{6.2}$$

where R is the Ricci scalar, as usual, and $\mathcal{L}_m$ is the Lagrangean density for matter fields ψ.

The crucial fact for the chameleon model to work is that matter fields ψ couple to the Einstein frame metric through $\tilde{g}_{\mu\nu}$. The notation here follows the one of [83] since it is more general than the original papers' notation, which starts with a specific phenomenological realization of the conformal factor, namely $A(\phi) = e^{g\phi/M_{Pl}}$, where g is a dimensionless constant. It should be pointed out that one can consider different couplings $\tilde{g}^{(i)}_{\mu\nu} = A^2_{(i)}g_{\mu\nu}$ for each type of matter $\psi^{(i)}$. Nevertheless, for simplicity, we will consider a unique conformal factor in the following analysis. The scalar potential $V(\phi)$ must respect specific constraints to give the scalar field the desired features for a chameleon, as we will see shortly.

The presence of a function of ϕ in the matter Lagrangean affects the scalar field equation of motion in the following way:

$$\begin{aligned}\delta S &= \int d^4x\sqrt{-g}\left[-g^{\mu\nu}\partial_\mu\phi\partial_\nu(\delta\phi) - V'(\phi)\delta\phi\right] + \int d^4x\frac{\partial\mathcal{L}_m}{\partial\tilde{g}^{\mu\nu}}\frac{\partial\tilde{g}^{\mu\nu}}{\partial\phi}\delta\phi \\ &= \int d^4x\delta\phi\left\{\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right) - \sqrt{-g}V'(\phi) - \frac{\sqrt{-\tilde{g}}}{2}\tilde{T}_{\mu\nu}[-2A^{-3}A'(\phi)g^{\mu\nu}]\right\} \\ &= \int d^4x\sqrt{-g}\delta\phi\left[\Box\phi - V'(\phi) + AA'(\phi)\tilde{T}_{\mu\nu}(A^2\tilde{g}^{\mu\nu})\right] = 0\,;\end{aligned} \tag{6.3}$$

where in the second equality we have partially integrated the first term, considering a vanishing boundary integral, and used the definition of energy-momentum tensor $\tilde{T}_{\mu\nu} \equiv -\frac{2}{\sqrt{-\tilde{g}}}\frac{\delta S_m}{\delta\tilde{g}^{\mu\nu}} = -\frac{2}{\sqrt{-\tilde{g}}}\frac{\partial\mathcal{L}_m}{\partial\tilde{g}^{\mu\nu}}$ in the so-called Jordan frame. In the third equality we have used the fact $\partial_\nu(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi) = \sqrt{-g}\nabla_\nu(g^{\mu\nu}\partial_\mu\phi) = \sqrt{-g}g^{\mu\nu}\nabla_\nu\partial_\mu\phi \equiv \Box\phi$, derived from equation (A.11) plus $\nabla_\nu g^{\mu\nu} = 0$, and just rewritten the last term using equation (6.1) such that $\sqrt{-g}$ becomes a common factor.

Defining $\tilde{T} \equiv \tilde{T}_{\mu\nu}\tilde{g}^{\mu\nu}$, the equation of motion for the scalar field becomes:

$$\Box\phi - V'(\phi) + A^3A'(\phi)\tilde{T} = 0 \tag{6.4}$$

Note that the last term in the above equation only exists because of a non-trivial conformal factor, i.e. $A(\phi) \neq 1$. It is already clear from equation (6.4) that this additional term generates an effective potential whose first derivative has the form:

$$V'_{eff}(\phi) = V'(\phi) - A^3A'(\phi)\tilde{T}\,; \tag{6.5}$$

but, to make our point, allow us to first rewrite $\tilde{T}$ in terms of a more familiar quantity, the local matter energy density ρ.

Note from the general energy-momentum tensor expression for a perfect fluid, equation (1.20), that $\tilde{g}^{\mu\nu}\tilde{T}_{\mu\nu} = -\tilde{\rho} + 3\tilde{P}$ for a generic metric $\tilde{g}_{\mu\nu}$. For non-relativistic (dust) matter, we saw in chapter 1 that $\tilde{P} \simeq 0$, thus $\tilde{g}^{\mu\nu}\tilde{T}_{\mu\nu} \equiv \tilde{T} \simeq -\tilde{\rho}$. However, this is the energy density in the Jordan frame and it is ϕ-dependent. It will be useful to rewrite that in terms of a ϕ-independent quantity. In order to do this, let us consider $g_{\mu\nu} \simeq \eta_{\mu\nu}$, the validity of this approximation is justified in [8], leading to an approximate stress-energy tensor for dust matter of the form $\tilde{T}_{\mu\nu} \simeq \tilde{\rho}u_\mu u_\nu$ which implies $\tilde{T}^0{}_0 \simeq -\tilde{\rho}$ while the rest of the components vanish, in a frame where $u^\mu = (-1, 0, 0, 0)$. Note that, since $\tilde{T}_{\mu\nu}$ should be conserved in the Jordan frame:

$$\begin{aligned}
\tilde{\nabla}_\mu\tilde{T}^\mu{}_\nu &= \partial_\mu\tilde{T}^\mu{}_\nu + \tilde{\Gamma}^\mu{}_{\mu\rho}\tilde{T}^\rho{}_\nu - \tilde{\Gamma}^\rho{}_{\mu\nu}\tilde{T}^\mu{}_\rho = 0 \Rightarrow \\
\tilde{\nabla}_\mu\tilde{T}^\mu{}_0 &= \partial_\mu\tilde{T}^\mu{}_0 + \tilde{\Gamma}^\mu{}_{\mu 0}\tilde{T}^0{}_0 - \tilde{\Gamma}^0{}_{00}\tilde{T}^0{}_0 = \partial_0\tilde{T}^0{}_0 + \tilde{\Gamma}^i{}_{i0}\tilde{T}^0{}_0 = \partial_0\tilde{T}^0{}_0 + \Gamma^i{}_{i0}\tilde{T}^0{}_0 + 3A^{-1}A'\partial_0\phi\tilde{T}^0{}_0 \\
&\simeq (\partial_t + 3A^{-1}A'\partial_t\phi)(-\tilde{\rho}) = 0 \therefore \partial_t\tilde{\rho} = -3\frac{A'(\phi)}{A(\phi)}\partial_t\phi\tilde{\rho} \Rightarrow \frac{\partial\tilde{\rho}}{\partial\phi} = -3\frac{A'(\phi)}{A(\phi)}\tilde{\rho}\,;
\end{aligned} \tag{6.6}$$

where, in the third line, we have used the approximation $g_{\mu\nu} \simeq \eta_{\mu\nu} \therefore \Gamma^{\mu}{}_{\nu\rho} \simeq 0$ to make the second term in the previous expression vanish. Moreover, in the third equality of the second line the following relation was used:

$$\begin{aligned}
\tilde{\Gamma}^{\mu}{}_{\nu\rho} &= \frac{\tilde{g}^{\mu\sigma}}{2}\left(\partial_\nu \tilde{g}_{\sigma\rho} + \partial_\rho \tilde{g}_{\sigma\nu} - \partial_\sigma \tilde{g}_{\nu\rho}\right) \\
&= \frac{A^{-2} g^{\mu\sigma}}{2}\left[A^2\left(\partial_\nu g_{\sigma\rho} + \partial_\rho g_{\sigma\nu} - \partial_\sigma g_{\nu\rho}\right) + 2AA'(\phi)\left(\partial_\nu \phi g_{\sigma\rho} + \partial_\rho \phi g_{\sigma\nu} - \partial_\sigma \phi g_{\nu\rho}\right)\right] \Rightarrow \\
\tilde{\Gamma}^{i}{}_{i0} &= \Gamma^{i}{}_{i0} + 3A^{-1}A'(\phi)\partial_0 \phi
\end{aligned} \tag{6.7}$$

From equation (6.6), notice that $\rho \equiv A^3 \tilde{\rho}$ is automatically a ϕ-independent quantity:

$$\frac{\partial \rho}{\partial \phi} = A^3 \frac{\partial \tilde{\rho}}{\partial \phi} + 3A^2 A'(\phi)\tilde{\rho} = A^3 \left[\frac{\partial \tilde{\rho}}{\partial \phi} + 3\frac{A'(\phi)}{A(\phi)}\tilde{\rho}\right] = 0 \tag{6.8}$$

Furthermore, it is proportional to the trace of the energy-momentum tensor in the Einstein frame:

$$\begin{aligned}
-\tilde{\rho} \simeq \tilde{T} = \tilde{g}^{\mu\nu}\left(\frac{-2}{\sqrt{-\tilde{g}}}\frac{\partial \mathcal{L}_m}{\partial \tilde{g}^{\mu\nu}}\right) = A^{-2} g^{\mu\nu}\left(\frac{-2}{A^4\sqrt{-g}}\frac{\partial \mathcal{L}_m}{\partial g^{\mu\nu}} A^2\right) = A^{-4} g^{\mu\nu} T_{\mu\nu} \equiv A^{-4} T \\
\Rightarrow -\rho A^{-3} \simeq A^{-4} T \therefore T \simeq -A\rho\,;
\end{aligned} \tag{6.9}$$

therefore, ρ earns the name matter energy density in the Einstein frame.

Finally, back to equation (6.5), the effective potential is found to be:

$$V'_{eff}(\phi) \simeq V'(\phi) + A^3 A'(\phi)\tilde{\rho} = V'(\phi) + A'(\phi)\rho \Rightarrow V_{eff}(\phi) = V(\phi) + \rho A(\phi)\,; \tag{6.10}$$

where since, at least classically, a constant off-set in the potential is negligible, the constant of integration after the $\Rightarrow$ was dropped.

The point of the model is to generate an effective potential that has a minimum at $\phi = \phi_{min}$. Note that the potential $V(\phi)$ does not need to have a minimum itself, only its sum with the ρ-dependent term needs. Therefore, since $V'_{eff}(\phi_{min}) = 0$ by construction:

$$V'(\phi_{min}) + \rho A'(\phi_{min}) = V_{,\phi}(\phi_{min}) + \rho A_{,\phi}(\phi_{min}) = 0 \tag{6.11}$$

Moreover, the square of the scalar field's effective mass is the second derivative of the potential with respect to ϕ, as usual. This effective mass is ϕ-dependent. In particular, the effective mass at the minimum of V_{eff}, i.e. at $\phi = \phi_{min}$, is:

$$m^2_{min} \equiv V''(\phi_{min}) + \rho A''(\phi_{min}) = V_{,\phi\phi}(\phi_{\rm min}) + \rho A_{,\phi\phi}(\phi_{min}) \tag{6.12}$$

It is usual to get, from a specific model, an exponential conformal factor $A(\phi)$. Therefore, we must now constrain the form of the potential $V(\phi)$ so that the desired conditions for a chameleon scalar are achieved. This will be useful later when analyzing the chameleon string-motivated model of Hinterbichler, Khoury, and Nastase [83] and the chameleon Kaluza-Klein model.

6.1.1 Conditions on the Chameleon Potential

The conditions on $V(\phi)$ depend on whether $A(\phi)$ is a monotonically increasing (positive exponential) or a monotonically decreasing (negative exponential) function, therefore we will split the two cases in this analysis. All three constraints on $V(\phi)$ come from equations (6.11) and (6.12).

- For a positive exponential $A(\phi)$ one must demand:

  (i). $V_{,\phi}(\phi_{min}) < 0$ so that $V_{eff}(\phi)$ has a minimum, see equation (6.11);

  (ii). $V_{,\phi\phi}(\phi_{min}) > 0$ so that the mass squared is positive even when $\rho \to 0$ in equation (6.12);

(iii). $\frac{dm^2_{min}}{d\rho} > 0$ since the main idea behind chameleon scalars is to allow for a large scalar mass in very dense regions, such as the Earth. This will lead to the following condition on $V_{,\phi\phi\phi}$.

From equation (6.11), considering ϕ_{min} as a function of ρ and taking the derivative with respect to ρ on both sides:

$$\begin{aligned}\frac{d\phi_{min}}{d\rho} &= -A_{,\phi}(\phi_{min})\left[\frac{dV_{,\phi}(\phi_{min})}{d\phi_{min}} + \rho\frac{dA_{,\phi}(\phi_{min})}{d\phi_{min}}\right]^{-1} \\ &= -A_{,\phi}(\phi_{min})[V_{,\phi\phi}(\phi_{min}) + \rho A_{,\phi\phi}(\phi_{min})]^{-1} < 0 \end{aligned} \tag{6.13}$$

Then, taking the derivative with respect to ρ of equation (6.12):

$$\begin{aligned}\frac{dm^2_{min}}{d\rho} &= V_{,\phi\phi\phi}(\phi_{min})\frac{d\phi_{min}}{d\rho} + A_{,\phi\phi}(\phi_{min}) + \rho A_{,\phi\phi\phi}(\phi_{min})\frac{d\phi_{min}}{d\rho} > 0 \\ &\Rightarrow V_{,\phi\phi\phi}(\phi_{min}) < -\rho A_{,\phi\phi\phi}(\phi_{min}) < 0 \end{aligned} \tag{6.14}$$

To summarize, in a region of interest around $\phi = \phi_{min}$, the chameleon potential must obey the following constraints:

$$V_{,\phi} < 0 \;\; ; \;\; V_{,\phi\phi} > 0 \;\; ; \text{and} \;\; V_{,\phi\phi\phi} < -\rho A_{,\phi\phi\phi} < 0 \tag{6.15}$$

- For a negative exponential $A(\phi)$ one must demand:

(i). $V_{,\phi}(\phi_{min}) > 0$ to cancel against $\rho A_{,\phi}(\phi_{min})$, equation (6.11), so that the effective potential has a minimum;

(ii). $V_{,\phi\phi}(\phi_{min}) > 0$ to ensure the positiveness of the mass squared in (6.12);

(iii). $\frac{dm^2_{min}}{d\rho} > 0$ again, which leads to the following condition on $V_{,\phi\phi\phi}$.

From equation (6.11), considering ϕ_{min} as a function of ρ and taking the derivative with respect to ρ on both sides:

$$\frac{d\phi_{min}}{d\rho} = -A_{,\phi}(\phi_{min})\left[\frac{dV_{,\phi}(\phi_{min})}{d\phi_{min}} + \rho\frac{dA_{,\phi}(\phi_{min})}{d\phi_{min}}\right]^{-1}$$
$$= -A_{,\phi}(\phi_{min})[V_{,\phi\phi}(\phi_{min}) + \rho A_{,\phi\phi}(\phi_{min})]^{-1} > 0 \qquad (6.16)$$

Then, taking the derivative with respect to ρ of equation (6.12):

$$\frac{dm^2_{min}}{d\rho} = V_{,\phi\phi\phi}(\phi_{min})\frac{d\phi_{min}}{d\rho} + A_{,\phi\phi}(\phi_{min}) + \rho A_{,\phi\phi\phi}(\phi_{min})\frac{d\phi_{min}}{d\rho} > 0$$
$$\Rightarrow V_{,\phi\phi\phi}(\phi_{min}) > -\rho A_{,\phi\phi\phi}(\phi_{min}) > 0 \qquad (6.17)$$

All in all, around $\phi = \phi_{min}$ the scalar potential must behave like:

$$V_{,\phi} > 0 \ ; \ V_{,\phi\phi} > 0 \ ; \text{and} \ V_{,\phi\phi\phi} > -\rho A_{,\phi\phi\phi} > 0 \qquad (6.18)$$

Requiring that $\frac{dm^2_{min}}{d\rho} > 0$ in both cases is crucial to the model not only because we need a massive chameleon on Earth to avoid violating experimental bounds, as will become clear in the next subsection, but also since, if it is intended to construct a quintessence model from the chameleon field, the scalar mass should be very low on cosmological scales. Therefore, a simpler massive field at any scale is not acceptable. The reason for requiring a very low mass from a quintessence model field lies in the heuristic argument that goes as follows: Since m^2 gives us information about the flatness of the effective potential (note that on cosmological scales $\rho_{\rm cosm} \simeq 10^{-29}{\rm g/cm}^3 \sim 10^{-46}\,{\rm GeV}^4 \Rightarrow V_{eff} \simeq V$ so the argument also holds true for the potential itself), considering a low mass means having an almost flat potential which dominates over the kinetic term over a region of interest, i.e. $V_{eff}(\phi) \simeq V(\phi) \gg \dot{\phi}^2$. Recall, from chapter 2, that Planck data [18] tell us that, in order to fit observations, the dark energy equation of state parameter needs to be $w \simeq -1$ (an approximate cosmological constant). Moreover, it was seen in chapter 1 that a scalar field has $w \simeq -1$ when $V(\phi) \gg \dot{\phi}^2$, which is the case obtained for a nearly massless scalar field on cosmological scales.

On the other hand, it is argued in [84] that the mass of the chameleon on cosmological scales is expected to be $\gg H$ (the Hubble parameter) in a quintessence model since it indicates that if $\phi = \phi_{min}$ at some point in time it will remain at the minimum later.

6.1.2 The Thin-Shell Effect

There is one last important feature of chameleon models which is essential to avoid violating equivalence principle tests, see [8], on Earth which is the so-called thin-shell effect. To show this feature, allow us to consider $A(\phi) = e^{g\phi/M_{Pl}}$, the phenomenological conformal factor of [8], for concreteness. For $g > 0$ (positive exponential), a generic chameleon potential that follows the constraints derived in the last subsection is considered. We want to describe the ϕ profile in the static case of having a sphere with homogeneous density ρ_{in} and radius R at the origin in a background with homogeneous density ρ_{out} outside it. This setup models, for instance, the Earth embedded in the solar system, here described as a homogeneous background. Deep inside the Earth, at $r \to 0$, ϕ is expected to tend to ϕ_{min-in}, the value which minimizes $V_{eff}(\phi)$ inside the sphere. We will see that this is only the case because the Earth is a large enough object (our definition of large enough will be clear later). At infinity, $r \to \infty$, one should expect $\phi \to \phi_{min-out}$, the value which minimizes the effective potential outside the sphere, since the effect of it should be negligible at this distance.

From equation 6.4, the equation of motion in this static and spherically symmetric case becomes:

$$\Box\phi = V'(\phi) + \frac{g}{M_{Pl}}\rho e^{g\phi/M_{Pl}} \therefore \frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr} = V'(\phi) + \frac{g}{M_{Pl}}\rho(r)e^{g\phi/M_{Pl}} = V'_{eff}(\phi)\,; \quad (6.19)$$

where:

$$\rho(r) = \begin{cases} \rho_{in} & \text{for} \quad r < R\,; \\ \rho_{out} & \text{for} \quad r \geq R \end{cases} \tag{6.20}$$

Equation (6.19) demands two boundary conditions which are going to be considered $\frac{d\phi}{dr} = 0$ at $r = 0$, so that the solution is stable at the origin, and $\phi \to \phi_{min-out}$ as $r \to \infty$. The solution has to be a continuous function that interpolates $\phi \simeq \phi_{min-in}$ at the origin and $\phi \to \phi_{min-out}$ at infinity. Notice that since $V'_{eff}(\phi) \simeq 0$ close to the origin, the field will tend to stay at its initial value. However, it has to come to a point $r = R_{roll}$ where it starts to change since at some point it will have to continuously connect with the behavior of $\phi(r)$ outside the sphere. Thus, there are three different regions to be analyzed: the first one is $0 \leq r < R_{roll}$, the second $R_{roll} \leq r < R$ and the third $r \geq R$.

1. First region: $0 \leq r < R_{roll}$.

   As we argued, the potential has to start close to the point that minimizes V_{eff} inside the sphere, namely ϕ_{min-in}. Therefore $V'_{eff}(\phi) \simeq 0$ – at least compared to the left hand side terms in the equation of motion (6.19) – in this region, so:

$$\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr} \simeq 0 \therefore \frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right) \simeq 0 \therefore \frac{d\phi}{dr} \simeq \frac{C_1}{r^2} \therefore \phi(r) \simeq -\frac{C_1}{r} + C_2 \tag{6.21}$$

   Applying the boundary condition $\frac{d\phi}{dr}\big|_{r=0} = 0$, one gets $C_1 = 0$, and by our initial consideration $C_2 \simeq \phi_{min-in}$. All in all:

$$\phi(r) \simeq \phi_{min-in} \quad ; \text{for} \quad 0 \leq r < R_{roll} \tag{6.22}$$

2. Second region: $R_{roll} \leq r < R$.

   Note that since $V_{eff} = V + \rho A$ and $V_{,\phi} < 0$ close to ϕ_{min-in}, as derived in the last subsection, regions with lower ρs have larger ϕ_{min}s, thus $\phi_{min-in} < \phi_{min-out}$. Since

$\phi(r)$ should be a continuous interpolation between ϕ_{min-in} and $\phi_{min-out}$, the values of ϕ in this region should be $> \phi_{min-in}$ and thus the effective potential and its derivative should be dominated by the conformal factor term, i.e. $|V'(\phi)| \ll (g\rho_{in}/M_{Pl})e^{g\phi/M_{Pl}}$. Furthermore, one can approximate $g\phi/M_{Pl} \ll 1$ since ϕ is not far from the minimum, supposedly $\ll M_{Pl}$, and g is a dimensionless coupling which might be considered $\sim \mathcal{O}(1)$, see [7]. Finally, the equation of motion becomes:

$$\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr} \simeq \frac{g\rho_{in}}{M_{Pl}} \therefore \frac{d}{dr}\left(r^2\frac{d\phi}{dr}\right) \simeq \frac{g\rho_{in}}{M_{Pl}}r^2 \therefore r^2\frac{d\phi}{dr} \simeq \frac{g\rho_{in}}{M_{Pl}}\frac{r^3}{3} + C_1$$
$$\phi(r) \simeq \frac{g\rho_{in}}{M_{Pl}}\frac{r^2}{6} - \frac{C_1}{r} + C_2 \tag{6.23}$$

To ensure continuity of $\phi(r)$ with what was derived in the first region one must demand $\phi(R_{roll}) \simeq \phi_{min-in}$ and $\frac{d\phi}{dr}\big|_{r=R_{roll}} = 0$, determining the value of the constants C_1 and C_2. After a little bit of algebra one finds:

$$C_1 = -\frac{g\rho_{in}}{M_{Pl}}\frac{R_{roll}^3}{3} \;\; ; \text{and} \;\; C_2 \simeq \phi_{min-in} - \frac{g\rho_{in}}{M_{Pl}}\frac{R_{roll}^2}{2} \Rightarrow$$
$$\phi(r) \simeq \frac{g\rho_{in}}{3M_{Pl}}\left(\frac{r^2}{2} + \frac{R_{roll}^3}{r}\right) - \frac{g\rho_{in}}{M_{Pl}}\frac{R_{roll}^2}{2} + \phi_{min-in} \;\; ; \text{for} \quad R_{roll} \leq r < R \tag{6.24}$$

3. Third region: $r \geq R$.

   In this region, since $\rho(r)$ suddenly drops from $\rho_{in} = \frac{3M}{4\pi R^3}$ to $\rho_{out} \ll \rho_{in}$, the derivative of the potential $V(\phi)$ dominates $V'_{eff}(\phi)$, i.e. $|V_{,\phi}| \gg (g\rho_{out}/M_{Pl})e^{g\phi/M_{Pl}}$. Then, a priori, to solve the equation of motion in this region one should know the form of the potential. However, since it must smoothly approach $\phi_{min-out}$ as $r \to \infty$, one might approximate the potential outside the sphere by $V(\phi) \simeq \frac{m^2_{min-out}}{2}(\phi - \phi_{min-out})^2$, where the coefficient had to be a mass squared and outside the sphere it is reasonable for it to be $\propto m^2_{min-out}$. Be careful, in our notation M is the total mass of the sphere, while m stands for the scalar mass in different regions, e.g. $m^2_{min-in} \equiv V''_{eff}(\phi_{min-in})$ and analogously for $m_{min-out}$. The equation of motion becomes:

$$\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr} \simeq m^2_{min-out}(\phi - \phi_{min-out}) \therefore \frac{d^2\tilde{\phi}}{dr^2} + \frac{2}{r}\frac{d\tilde{\phi}}{dr} \simeq m^2_{min-out}\tilde{\phi}\,; \tag{6.25}$$

where $\tilde{\phi} \equiv \phi - \phi_{min-out}$.

To put the above differential equation in a form in which it will be more easily solved we further redefine the field $\varphi \equiv r\tilde{\phi}$.

$$\begin{aligned}
\frac{d^2}{dr^2}\left(\frac{\varphi}{r}\right) + \frac{2}{r}\frac{d}{dr}\left(\frac{\varphi}{r}\right) - m^2_{min-out}\left(\frac{\varphi}{r}\right) \simeq 0 \\
-\frac{2}{r^2}\frac{d\varphi}{dr} + \frac{1}{r}\frac{d^2\varphi}{dr^2} + \frac{2}{r^3}\varphi + \frac{2}{r^2}\frac{d\varphi}{dr} - \frac{2}{r^3}\varphi - m^2_{min-out}\left(\frac{\varphi}{r}\right) \simeq 0 \\
\frac{d^2\varphi}{dr^2} - m^2_{min-out}\varphi \simeq 0 \Rightarrow \varphi(r) \simeq C_1 e^{m_{min-out}r} + C_2 e^{-m_{min-out}r} \Rightarrow \\
\phi(r) \simeq C_1\frac{e^{m_{min-out}r}}{r} + C_2\frac{e^{-m_{min-out}r}}{r} + \phi_{min-out}
\end{aligned} \tag{6.26}$$

From the boundary condition $\phi \to \phi_{min-out}$ as $r \to \infty$, one gets $C_1 = 0$. The values of C_2 and R_{roll} must be found by matching the $\phi(r)$ and its first derivative at $r = R$ against the equation (6.24). Considering $R - R_{roll} \ll R$, one can show that:

$$\phi(r) \simeq -\left(\frac{g}{4\pi M_{Pl}}\right)\left(\frac{3\Delta R}{R}\right)\frac{Me^{-m_{min-out}(r-R)}}{r} + \phi_{min-out} \;\; ; \text{for} \quad r \geq R\,; \tag{6.27}$$

where:

$$\frac{\Delta R}{R} \equiv \frac{\phi_{min-out} - \phi_{min-in}}{6gM_{Pl}\Phi} \simeq \frac{R - R_{roll}}{R} \ll 1 \;\; ; \text{and} \;\; \Phi \equiv \frac{M}{8\pi M^2_{Pl}R} = \frac{R^2\rho_{in}}{6M^2_{Pl}} \tag{6.28}$$

Note that Φ is the Newtonian potential at the surface of the sphere when the potential at infinity is set to zero as usual.

The proper definition of a "large enough" object comes from equation (6.28). This is an object whose Newtonian potential at the surface, i.e. Φ, is much bigger than $(\phi_{min-out} - \phi_{min-in})/6gM_{Pl}$, leading to $(\Delta R/R) \ll 1$. Such objects are large enough for ϕ to continuously approach ϕ_{min-in} at the origin coming from $\phi_{min-out}$ far outside the sphere. The factor $(3\Delta R/R)$ in equation (6.27) implies that only the outer piece, the thin-shell piece, of the sphere contributes to the potential outside it. The interpretation of this is that in the core of the sphere, i.e. the region where $0 \leq r < R_{roll}$, the scalar field has a high mass m_{min-in} which Yukawa-suppresses the influence of this region on $\phi(r)$ outside the sphere, see [83]. The reason for the need for a high chameleon mass on Earth, compared to its value on cosmological scales, is precisely to allow this suppression to occur. Small objects, in the sense that $\Delta R/R \sim 1$, have a thick-shell. It is shown in [8] that the factor $(3\Delta R/R)$ does not appear in the analogous of equation (6.27) in this case. Therefore, the effective coupling g^2_{eff} in the interaction of two large objects is $\ll g^2$ due to two screening factors $(3\Delta R/R)$, one from each object:

$$g^2_{eff} = g^2 \left(\frac{3\Delta R}{R}\right)_1 \left(\frac{3\Delta R}{R}\right)_2 ; \tag{6.29}$$

while the interaction between two small objects is simply governed by g^2, since the two screening factors are of order one leading to $g^2_{eff} \sim g^2$.

The thin-shell effect is important to avoid violating experimental constraints while keeping $g \sim \mathcal{O}(1)$. On the other hand, small satellites might have a thick-shell enabling us to look for experimental evidence of a chameleon scalar in satellite experiments.

6.2 Chameleon in Superstring Theory

As anticipated, although the general chameleon setup is string-independent, the idea is quite useful in trying to obtain 4-dimensional physics from 10-dimensional string theory compactification. The paper [83] presents a concrete chameleon model based on the KKLT scenario.

6.2.1 General Remarks on the KKLT Scenario

The Kachru, Kallosh, Linde, and Trivedi (KKLT) construction [85] came to answer an important question: Is it possible to obtain de Sitter backgrounds in string theory? The answer given by the authors was affirmative. The nonexistence of such backgrounds would perhaps be a problem to the theory since observations tell us our world has a de Sitter background, $\Lambda > 0$ as discussed in chapter 1. It must be pointed out though that the KKLT construction has been contested. This ended up leading to the swampland conjecture [86], but it will not be discussed here.

To our purposes, it is enough to say that the KKLT scenario yields a potential, coming from equation (3.43), whose shape (with a modification in the exponential piece of the super-potential, see [83]) is the following: first, the potential starts at infinity when the scalar field is at $R = 0$ – we will shortly define R and its relation to the canonical scalar ϕ of equation (6.2) – then it drops steeply to a positive minimum, after that, it reaches a maximum and then finally it smoothly approaches zero as $R \to \infty$, see figure 6.1. Inspired by this result, the general form of a potential $V(R)$ will be guessed after the conformal factor $A = A(\phi)$ is figured out.

First, one needs to discuss the dimensional reduction from the 10-dimensional string theory to our 4D world. In chapter 3 we saw that the dimensional reduction ansatz may bring inconsistencies to the theory, i.e. some equations of motion of the higher dimensional theory may not be satisfied. It turns out that to dimensionally reduce a metric it is convenient

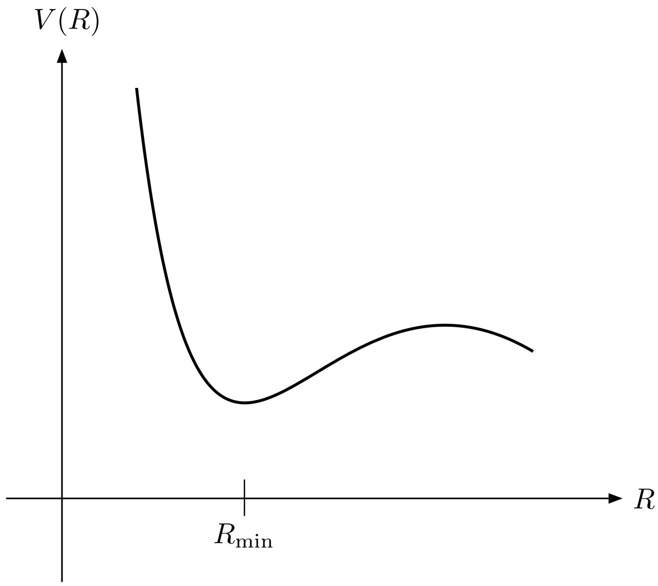

Figure 6.1: Representation of the shape of the KKLT potential for a superpotential $W = W_0 + Ae^{ia\varrho}$ with $a < 0$. The relevant region for our analysis will be $R \leq R_{min}$.

to use a non-linear KK expansion before keeping only the zero modes, in the same fashion it was done for the original KK theory in appendix D, see [12] chapter 13. For a dimensional reduction using a non-linear ansatz from D to d dimensions, one generically gets:

$$ds_D^2 \equiv \hat{g}_{MN} dx^M dx^N = R^2 ds_d^2 + 2g_{\mu n} dx^\mu dx^n + g_{mn} dx^m dx^n \, ; \tag{6.30}$$

where $R = \left[\sqrt{\det(g_{mn})}\right]^{-\frac{1}{d-2}}$ and $ds_d^2 = g_{\mu\nu} dx^\mu dx^\nu$ is the d-dimensional Einstein frame line element.

By making use of this conformal factor R one goes from a D-dimensional gravity action in the Einstein frame to its counterpart in d dimensions after the reduction, again this is what was done in appendix D for the simple KK theory. That is why it is this useful to consider an ansatz that includes R, other conformal factors would not yield a lower dimension gravity action in the Einstein frame. However, the way 4-dimensional matter fields will couple to gravity is not clear in this reduction process. One possibility is that such matter fields will couple to the D-dimensional Jordan frame metric. In other words, the conformal factor of equation (6.1) is:

$$\tilde{g}_{\mu\nu} = R^2 g_{\mu\nu} \Rightarrow A^2(\phi) = R^2 = [\det(g_{mn})]^{-\frac{1}{d-2}} \tag{6.31}$$

The exact relation between R and the canonical scalar field ϕ of equation (6.2) depend on the compactification, but in general, one can write:

$$\phi = \frac{M_{Pl}}{g} \log\left(\frac{R}{R_*}\right) ; \tag{6.32}$$

where g is the coupling between the scalar and matter fields, it has the same role as in the thin-shell analysis, and R_* is just a scale of reference fixing a conformal factor reference $A(R = R_*) = A(\phi = 0) = 1$. Therefore, we can rewrite the conformal factor as:

$$A(\phi) = \frac{A(\phi(R))}{A(\phi = 0)} = \frac{R}{R_*} = e^{g\phi/M_{Pl}} ; \tag{6.33}$$

this procedure of replacing $A(\phi) = R \rightarrow R/R_*$ can be done without loss of generality, since the R_* can always be absorbed as a rescaling of ρ in equation (6.10), see [83].

Up until now, we have discussed how gravity couples to matter fields in 4D, which gave us the conformal factor $A(\phi)$ of the model and the function $\phi = \phi(R)$ was described. The last thing we need to provide the model with is the KKLT inspired potential. The shape of $V(R)$ is constructed to mimic the behavior of the KKLT scalar potential itself. This is done by considering a potential defined by parts:

$$V(R) = \begin{cases} M_{Pl}^4 v \left[e^{\gamma\left(R^{-k} - R_*^{-k}\right)} - 1\right] & ;\text{ for } R < R_* ; \\ M_{Pl}^4 \left[-\alpha(R - R_*) + \beta(R - R_*)^2\right] & ;\text{ for } R \geq R_* ; \end{cases} \tag{6.34}$$

where $\alpha, \beta, \gamma, k, v > 0$ are dimensionless parameters.

The first line describes the steep potential of the KKLT construction close to $R \rightarrow 0$, while the second line is a generic quadratic potential to model the minimum of the modified KKLT potential as described before. The behavior of $V(R)$ at large R doesn't match the one of the original KKLT scenario, but the region $R \gg R_{min}$ (where R_{min} corresponds to the minimum of the potential) will not be relevant to our analysis anyway.

Notice from equation (6.34) that the minimum of the potential is given by:

$$V'(R_{min}) = M_{Pl}^4[-\alpha + 2\beta(R_{min} - R_*)] = 0 \therefore R_{min} = R_* + \frac{\alpha}{2\beta} \tag{6.35}$$

The last relevant point to be made here is that, since the potential has to be continuous at $R = R_*$, one must impose $V(R \to R_*^-) = V(R \to R_*^+)$ and $\frac{dV}{dR}\big|_{R\to R_*^-} = \frac{dV}{dR}\big|_{R\to R_*^+}$. The first equation is trivially satisfied while the second yields a constraint among the parameters of the model:

$$V'(R^-) = M_{Pl}^4 v e^{\gamma\left(R^{-k} - R_*^{-k}\right)}(-\gamma k)R^{-k-1} \quad ; \text{and} \quad V'(R^+) = M_{Pl}^4[-\alpha + 2\beta(R - R_*)] \Rightarrow$$
$$\frac{dV}{dR}\Big|_{R\to R_*^-} = \frac{dV}{dR}\Big|_{R\to R_*^+} \therefore v(-\gamma k)R_*^{-k-1} = -\alpha \therefore \alpha = \frac{\gamma k v}{R_*^{k+1}} \tag{6.36}$$

6.2.2 Experimental Bounds on a KKLT Motivated Model

We shall review how experiments constrain the above KKLT inspired chameleon model parameters, see [83] for more details. The most stringent constraint comes from fifth force searches which are performed in vacuum chambers where the chameleon mass is set by the size of the chamber to be $m_{vac} \sim 10^{-1}\,\text{cm}^{-1} \sim 10^{-33} M_{Pl}$. For such experiments, the constraint on fifth force interactions leads to [87]:

$$g_{eff}^2 \lesssim 10^{-3}\,; \tag{6.37}$$

Throughout this subsection m_X (ϕ_X) will always stand for m_{min-X} (ϕ_{min-X}) in the X medium, e.g. $m_{vac} \equiv m_{min-vac}$ ($\phi_{vac} \equiv \phi_{min-vac}$). The *min* part of the subscript will be hidden for simplicity of the notation.

Considering $g \sim \mathcal{O}(1)$, one must have as the screening factor for a test mass:

$$\left(\frac{3\Delta R}{R}\right)_{tm} = \frac{\phi_{vac} - \phi_{tm}}{2gM_{Pl}\Phi_{tm}} \lesssim 10^{-3/2}\,; \tag{6.38}$$

assuming identical test masses, i.e. equal screening factors for both. The tm subscript stands for test mass. Both ϕ_{vac} and ϕ_{tm} are the field values that minimize V_{eff} out and inside the test mass respectively.

The Newtonian potential is estimated to be $\Phi_{tm} \sim 10^{-27}$, see [7]. Again, assuming a coupling of order one, one gets:

$$\phi_{vac} - \phi_{tm} \lesssim 10^{-28} M_{Pl} \tag{6.39}$$

We must now find a way to translate the inequality (6.39) into a constraint on the parameters of the model. Let us assume the field is rolling down the steep drop of the potential, i.e. $R < R_*$. In this case, we have $R_* - R \ll R_* \Rightarrow (R - R_*)/R_* \to 0$ thus the following approximation is valid:

$$\begin{aligned} \phi = \frac{M_{Pl}}{g}\log\left(\frac{R}{R_*}\right) = \frac{M_{Pl}}{g}\log\left(\frac{R-R_*}{R_*}+1\right) \Rightarrow \\ \phi \simeq \frac{M_{Pl}}{g}\left(\frac{R-R_*}{R_*}\right) \therefore R \simeq gR_*\frac{\phi}{M_{Pl}} + R_*\,; \end{aligned} \tag{6.40}$$

since $\log(x+1) \simeq x$ for $x \to 0$.

Since we know the chameleon mass in the vacuum of the chamber, it would be nice to write an expression for R_{vac} in terms of it. First, note that the chameleon mass at any field value is:

$$\begin{aligned} m^2 \equiv \frac{d^2V_{eff}}{d\phi^2} \simeq \frac{d^2V}{d\phi^2} = \frac{dR}{d\phi}\frac{d}{dR}\left(\frac{dR}{d\phi}\frac{dV}{dR}\right) = \left(\frac{gR_*}{M_{Pl}}\right)^2 \frac{d}{dR}\left[-M_{Pl}^4 vk\gamma R^{-k-1}e^{\gamma\left(R^{-k}-R_*^{-k}\right)}\right] \\ \simeq M_{Pl}^2 R_*^2 g^2 k^2 \gamma^2 v R^{-2k-2} e^{\gamma\left(R^{-k}-R_*^{-k}\right)} \simeq g^2 M_{Pl}^2 v \left(\frac{\gamma k}{R^k}\right)^2 e^{\gamma\left(R^{-k}-R_*^{-k}\right)}\,; \end{aligned} \tag{6.41}$$

where, in the second and fifth equalities above, we only considered derivatives of the exponentials since they dominate over derivatives of polynomials. Moreover, in the last equality the fact that $R_*/R \to 1$ was used.

Note that one can simplify the exponential factor above since:

$$\begin{aligned}\gamma(R^{-k} - R_*^{-k}) &= -\gamma R_*^{-k}\left[1 - \left(\frac{R}{R_*}\right)^{-k}\right] = -\gamma R_*^{-k}\left[1 - \left(\frac{R-R_*}{R_*} + 1\right)^{-k}\right] \\ &\simeq -\gamma R_*^{-k}\left\{1 - \left[1 - k\left(\frac{R-R_*}{R_*}\right)\right]\right\} = -\frac{\gamma k}{R_*^k}\left(\frac{R-R_*}{R_*}\right); \qquad (6.42)\end{aligned}$$

leading to $e^{\gamma\left(R^{-k}-R_*^{-k}\right)} \simeq e^{-\frac{\gamma k}{R_*^k}\left(\frac{R-R_*}{R_*}\right)}$.

Finally, from equation (6.41), the expression for R_{vac} in terms of m_{vac} is:

$$\begin{aligned}m_{vac}^2 &\simeq g^2 M_{Pl}^2 v \left(\frac{\gamma k}{R_*^k}\right)^2 \left(\frac{R_*}{R_{vac}}\right)^{2k} e^{-\frac{\gamma k}{R_*^k}\left(\frac{R_{vac}-R_*}{R_*}\right)} \simeq g^2 M_{Pl}^2 v \left(\frac{\gamma k}{R_*^k}\right)^2 e^{-\frac{\gamma k}{R_*^k}\left(\frac{R_{vac}-R_*}{R_*}\right)} \Rightarrow \\ \frac{R_{vac}-R_*}{R_*} &\simeq -\frac{R_*^k}{\gamma k}\log\left[\frac{m_{vac}^2}{g^2 M_{Pl}^2 v}\left(\frac{R_*^k}{\gamma k}\right)^2\right] = \frac{R_*^k}{\gamma k}\log\left[\frac{g^2 M_{Pl}^2 v}{m_{vac}^2}\left(\frac{\gamma k}{R_*^k}\right)^2\right] \qquad (6.43)\end{aligned}$$

Now, we need to find a similar expression for the value of the field inside the test mass R_{tm}. However, this time ρ_{tm} is the known quantity instead of the mass of the chameleon inside the object m_{tm}. So we will start by setting $\frac{dV_{eff}}{d\phi}(\phi_{tm}) = 0 \Rightarrow \frac{dV_{eff}}{dR}(R_{tm}) = 0$ rather than by the expression of m^2, as in equation (6.41).

$$\begin{aligned}V_{eff}(R) &= M_{Pl}^4 v\left[e^{\gamma\left(R^{-k}-R_*^{-k}\right)} - 1\right] + \rho\frac{R}{R_*} \Rightarrow \\ \frac{dV_{eff}}{dR}\Big|_{R=R_{tm}} &= -M_{Pl}^4 v\gamma k R_{tm}^{-k-1} e^{\gamma\left(R^{-k}-R_*^{-k}\right)} + \frac{\rho_{tm}}{R_*} = 0 \Rightarrow \\ \rho_{tm} &= M_{Pl}^4 v\gamma k\left(\frac{R_*^{k+1}}{R_{tm}^{k+1}}\right)\frac{1}{R_*^k}e^{\gamma\left(R^{-k}-R_*^{-k}\right)} \simeq M_{Pl}^4 v\gamma k\frac{1}{R_*^k}e^{-\frac{\gamma k}{R_*^k}\left(\frac{R-R_*}{R_*}\right)} \Rightarrow \\ \frac{R_{tm}-R_*}{R_*} &\simeq -\frac{R_*^k}{\gamma k}\log\left(\frac{\rho_{tm}}{M_{Pl}^4 v}\frac{R_*^k}{\gamma k}\right) = \frac{R_*^k}{\gamma k}\log\left(\frac{M_{Pl}^4 v}{\rho_{tm}}\frac{\gamma k}{R_*^k}\right) \qquad (6.44)\end{aligned}$$

Finally, back to the inequality (6.39):

$$\begin{aligned}\phi_{vac} - \phi_{tm} &\simeq \frac{M_{Pl}}{g}\left(\frac{R_{vac}-R_*}{R_*} - \frac{R_{tm}-R_*}{R_*}\right) = \frac{M_{Pl}}{g}\left(\frac{R_{vac}-R_{tm}}{R_*}\right) \\ &\simeq \left(\frac{M_{Pl}}{g}\right)\frac{R_*^k}{\gamma k}\log\left(\frac{g^2\rho_{tm}}{m_{vac}^2 M_{Pl}^2}\frac{\gamma k}{R_*^k}\right) \lesssim 10^{-28} M_{Pl}\end{aligned} \tag{6.45}$$

Considering $g \sim \mathcal{O}(1)$, as anticipated, and substituting $\rho_{tm} \sim 10\,\mathrm{g/cm}^3 \sim 10^{-89} M_{Pl}^4$ and $m_{vac} \sim 10^{-33} M_{Pl}$ we get:

$$\begin{aligned}\frac{R_*^k}{\gamma k}\log\left(\frac{10^{-89}}{10^{-66}}\frac{\gamma k}{R_*^k}\right) = \frac{R_*^k}{\gamma k}\left[\log\left(\frac{\gamma k}{R_*^k}\right) - 23\log 10\right] \lesssim 10^{-28} \Rightarrow \\ \frac{R_*^k}{\gamma k} \lesssim 10^{-30}\end{aligned} \tag{6.46}$$

That is the desired constraint one obtains assuming $g \sim \mathcal{O}(1)$ for the parameters of the steep drop piece of the potential. We can further derive the maximum range of the scalar interaction by combining equation (6.41) with the analogous of the third line of equation (6.44) for any medium, not necessarily the test mass, as long as ρ is large enough such that the value of R which minimizes the effective potential lies in the steep exponential part of $V(R)$. Dividing the two expressions one gets:

$$m^2 \simeq g^2 \frac{\gamma k}{R_*^k}\frac{\rho}{M_{Pl}^2} \gtrsim g^2 10^{30}\frac{\rho}{M_{Pl}^2} \therefore m \gtrsim 10^{15}\frac{\sqrt{\rho}}{M_{Pl}} \tag{6.47}$$

From the above constraint, one can tell for instance that the range on Earth is $m_{\mathrm{Earth}}^{-1} \lesssim 0.1\,\mathrm{mm}$, which is another way of realizing, other than directly calculating $\Delta R/R$ for Earth, that the Earth has a thin-shell.

Equation (6.47) shows us that the higher the energy density of a region of interest is, the higher the minimum mass the chameleon can assume is. This is the desired feature one expects from a chameleon scalar in order to avoid violating experimental bounds on Earth.

We have not considered low energy density regions yet. For regions whose $\rho < \rho_*$ (ρ_* will be defined next) the value of R which minimizes the effective potential lies in the quadratic part of (6.34) rather than in the steep exponential part of it. Therefore, all calculations made up until this point in this subsection are not valid for low ρ regions, since we were assuming R fell into the steep drop piece of $V(R)$. To be precise, ρ_* is defined to be the density for which the effective potential has a minimum at $R = R_*$. On one hand, ρ_* is found to be:

$$\frac{dV_{eff}(R_*)}{dR} = \frac{dV(R_*)}{dR} + \rho_* \frac{dA(R_*)}{dR} = 0 \therefore M_{Pl}^4 \upsilon\gamma k \left[e^{\gamma\left(R^{-k} - R_*^{-k}\right)} R^{-k-1} \right] \Big|_{R=R_*} = \frac{\rho_*}{R_*} \Rightarrow$$
$$\rho_* = M_{Pl}^4 \upsilon\gamma k R_*^{-k}\,; \tag{6.48}$$

while using the other part of the potential:

$$\frac{dV_{eff}(R_*)}{dR} = \frac{dV(R_*)}{dR} + \rho_* \frac{dA(R_*)}{dR} = 0 \therefore M_{Pl}^4 \alpha = \frac{\rho_*}{R_*} \therefore \rho_* = M_{Pl}^4 \alpha R_*\,; \tag{6.49}$$

and both expressions are certainly compatible due to (6.36), as they are supposed to be.

On cosmological scales we know that $\rho_{cosm} \simeq 10^{-29}\,\mathrm{g/cm}^3 \sim 10^{-46}\,\mathrm{GeV}^4 \Rightarrow V_{eff} \simeq V$ and thus we expect the minimum of the effective potential to lie in the quadratic part of (6.34), close to R_{min}. Therefore, it turns out that in order to constrain the parameters α and β, one needs to consider experimental bounds coming from the solar system data and beyond.

Demanding that our galaxy is screened, see [83], one gets:

$$\left(\frac{3\Delta R}{R}\right)_G = \frac{\phi_{cosm} - \phi_{ss}}{2gM_{Pl}\Phi_G} < 1\,; \tag{6.50}$$

where $\Phi_G \sim 10^{-6}$ and the subscript ss is short for solar system.

Considering $R_{cosm} \simeq R_{min}$, by the discussion of the second to last paragraph, and R_{ss} very close to R_{min}, more precisely $|R_{min} - R_{ss}| \ll R_*$, which itself is expected to be close to

R_* such that (6.40) is still valid, the above inequality translates to:

$$\phi_{cosm} - \phi_{ss} \simeq \frac{M_{Pl}}{g}\left(\frac{R_{min} - R_{ss}}{R_*}\right) \lesssim 2gM_{Pl}10^{-6} \therefore \frac{R_{min} - R_{ss}}{R_*} \lesssim 10^{-6}\,; \tag{6.51}$$

where, after the $\therefore$ symbol, we have considered $g \sim \mathcal{O}(1)$, as before.

Approximating $R_{ss} \simeq R_*$ yields:

$$\frac{\alpha}{\beta} \lesssim 10^{-6} R_*\,; \tag{6.52}$$

since $R_{min} = R_* + \frac{\alpha}{2\beta}$, equation (6.35).

The constraint (6.52) gives $R_{min} \simeq R_*$, as previously supposed, and since R_{ss} is expected to be in the middle of these two values the approximation $|R_{min} - R_{ss}| \ll R_*$ is also justified.

As discussed before, the above consideration of R_{cosm} falling into the quadratic part of the potential, means one must make sure $\rho_* > \rho_{cosm}$, thus:

$$\rho_* = M_{Pl}^4 \alpha R_* > \rho_{cosm} \sim 10^{-46}\text{GeV}^4 \sim 10^{-120} M_{Pl}^4 \therefore \alpha \gtrsim 10^{-120} R_*^{-1} \tag{6.53}$$

Combining the constraints (6.52) and (6.53), one gets the following bound on β:

$$\beta \gtrsim \alpha 10^6 R_*^{-1} \gtrsim 10^{-114} R_*^{-2} \tag{6.54}$$

Finally, from the constraint (6.54), we arrive at a lower bound on the scalar field mass on cosmological scales:

$$m_{cosm}^2 \simeq \left.\frac{d^2V}{d\phi^2}\right|_{\phi=\phi_{min}} = \left.\left[\frac{dR}{d\phi}\frac{d}{dR}\left(\frac{dR}{d\phi}\frac{dV}{dR}\right)\right]\right|_{R=R_{min}} \simeq 2M_{Pl}^2 R_*^2 g^2 \beta \sim M_{Pl}^2 R_*^2 \beta \gtrsim 10^{-114} M_{Pl}^4$$

$$m_{cosm} \gtrsim 10^{-57} M_{Pl} \sim 10^3 H_0\,; \tag{6.55}$$

where, in the third equality of the first line, we have used the quadratic piece of the potential (6.34) and the approximation (6.40). After that, the previously obtained bound on β was used along with the consideration $g \sim \mathcal{O}(1)$ made throughout this subsection.

The above constraint on m_{cosm} agrees with the expected $m_{\rm cosm} \gg H_0$ relation discussed on [84], as anticipated. More importantly, the maximal range of the scalar interaction on cosmological scales is $\sim$ Mpc, allowing it to contribute to structure formation, see [83] and the references therein.

The bounds calculated thus far take the potential (6.34) phenomenologically and constrain its parameters. However, the actual KKLT potential studied in [83] does not yield the bound (6.55) on m_{cosm}. Instead one gets:

$$m_{cosm} \gtrsim 10^{15} H_0\,; \tag{6.56}$$

which is not suitable for a dark energy candidate since the scalar interaction range becomes too small. The difference between the above lower bounds on m_{cosm} is due to the quadratic part of (6.34) not being a good model for the KKLT potential. Nevertheless, the analysis made on the steep exponential part of (6.34) is still valid.

As a final remark, the string-theory-motivated model studied in this section naturally has $g \sim \mathcal{O}(1)$ coming from string compactification. Nevertheless, one could allow any value for g and analyze the constraints one gets from it and their consequences. This was done in [88], where insights on both why the observed cosmological constant is too small, yet non-zero, compared to its theoretically expected value and why it is comparable to other energy densities nowadays, the coincidence problem, were given.

6.3 Chameleon in the Original Kaluza-Klein Theory

Moving on from the KKLT motivated model, another model that would naturally benefit from the chameleon idea is one coming from the Kaluza-Klein original theory. By the time the theory was proposed, the only known forces of nature were the gravitational and electromagnetic ones. Today, there are the two nuclear forces that are experimentally known and the possibility of the existence of a fifth force explaining dark matter features in quintessence models which is still hypothetical, of course. So even if one makes the KK theory work, it would not be a complete theory of physics as imagined by physicists at that time. Nevertheless, at large distances where the nuclear forces are not relevant, if we were able to get rid of the inconsistency brought to the theory by requiring $\varphi = 0 \Leftrightarrow \phi = 1$ in the action (3.17), Kaluza and Klein's idea could generate an effective model of the universe. The previously mentioned inconsistency is known not to be the only problem within the theory, see [41], but making the theory consistent would already be a progress. The chameleon idea can be applied to serve this purpose.

The reduction of the 5-dimensional Einstein-Hilbert action does not give us a potential for the scalar field, which is necessary for a chameleon model. Therefore, if one wants to motivate a potential coming from the 5-dimensional theory, the simplest possibility is to consider a cosmological constant in 5D, after all, there is nothing simpler than a constant as the Lagrangean density. The dimensional reduction of this piece of the action leads to:

$$\begin{aligned} \hat{S}_\Lambda &= -\frac{1}{16\pi\hat{G}}\int d^5x\sqrt{-\hat{g}}(2\Lambda) = -\left(\frac{1}{16\pi\hat{G}}\int dy\right)\int d^4x\sqrt{-g}\phi^{-1/3}(2\Lambda) \\ &= -\frac{1}{16\pi G}\int d^4x\sqrt{-g}\phi^{-1/3}(2\Lambda) \end{aligned} \tag{6.57}$$

where we recall that $G = \hat{G}/\int dy$, see chapter 3.

The full action is the sum of the 5D Einstein-Hilbert action plus $\hat{S}_\Lambda$. Since the boundary integral in equation (D.32) does not contribute to the equation of motion, we will define $S_{KK} \equiv \hat{S}_{E-H} + \hat{S}_\Lambda - \oint_{\partial\Omega} d\Sigma_\mu \sqrt{-g}X^\mu$ for short. Therefore:

$$S_{KK} = \int d^4x\sqrt{-g}\left[\frac{R}{16\pi G} - \frac{1}{4}\phi F_{\mu\nu}F^{\mu\nu} - \frac{1}{6(16\pi G)}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} - \frac{\Lambda}{8\pi G}\phi^{-1/3}\right] \tag{6.58}$$

It is clear from the above action that not putting $\phi = 1$ does not yield exactly the action for gravity plus electromagnetism plus the chameleon scalar in 4 dimensions. The $F_{\mu\nu}F^{\mu\nu}$ term has an extra ϕ factor multiplying it. One would have to explain then why ϕ does not vary considerably from place to place or in time nowadays so that it is effectively a constant. This problem will not be the focus of our analysis here though.

From equation (6.58), one can see that the cosmological constant action in 5 dimensions becomes a potential for the scalar in the dimensionally reduced theory. To rewrite S_{KK} in the form of the action (6.2), except for the matter action, one needs to redefine ϕ to get a canonical kinetic term for the scalar field. Only then all the analysis of the first section will be valid. One has two alternatives then:

- Define $\varphi \equiv -\frac{\log\phi}{4\sqrt{3\pi G}} \Rightarrow \phi = e^{-4\sqrt{3\pi G}\varphi} = e^{-\sqrt{6}\varphi/M_{Pl}}$, as it was done in equation (3.17). Then, the full action becomes:

$$S_{KK} = \int d^4x\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{1}{4}e^{-\sqrt{6}\frac{\varphi}{M_{Pl}}}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{\Lambda}{8\pi G}e^{\frac{\sqrt{6}}{3}\varphi/M_{Pl}}\right)$$
$$\Rightarrow V(\varphi) = \frac{\Lambda}{8\pi G}e^{\frac{\sqrt{6}}{3}\varphi/M_{Pl}} \equiv Ce^{k\varphi/M_{Pl}} \;\; ; \text{ where } \; k > 0 \tag{6.59}$$

- Define $\varphi \equiv \frac{\log\phi}{4\sqrt{3\pi G}} \Rightarrow \phi = e^{4\sqrt{3\pi G}\varphi} = e^{\sqrt{6}\varphi/M_{Pl}}$, leading to:

$$S_{KK} = \int d^4x\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{1}{4}e^{\sqrt{6}\frac{\varphi}{M_{Pl}}}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{\Lambda}{8\pi G}e^{-\frac{\sqrt{6}}{3}\varphi/M_{Pl}}\right)$$
$$\Rightarrow V(\varphi) = \frac{\Lambda}{8\pi G}e^{-\frac{\sqrt{6}}{3}\varphi/M_{Pl}} \equiv Ce^{-k\varphi/M_{Pl}} \;\; ; \text{ where } \; k > 0 \tag{6.60}$$

6.3.1 Problems with the Initial Idea

Up until now, our idea of saving Kaluza-Klein theory from inconsistencies using the chameleon setup seems promising. However, the matter action still needs to be described if we wish φ to be a chameleon. Now, we must face the same problem as in the KKLT inspired model. In both cases, dimensional reductions have been considered, this time from 5 to 4 dimensions and from 10 to 4 dimensions in the previous case. The issue is that we have no guidelines on how the 4-dimensional matter action should look like. Adopting the same ansatz as before, one assumes the 4-dimensional matter fields couple to the 5-dimensional Jordan frame metric. Then, from equation (3.16), the conformal factor can be figured out to be:

$$\tilde{g}_{\mu\nu} \equiv \phi^{-1/3} g_{\mu\nu} \Rightarrow A^2(\phi) = \phi^{-1/3} \tag{6.61}$$

The conformal factor for both cases analyzed before will be different. If one considers $\varphi \equiv -\frac{\log\phi}{4\sqrt{3\pi G}}$, then $A(\varphi) = e^{\frac{1}{\sqrt{6}}\varphi/M_{Pl}} = e^{\frac{k}{2}\varphi/M_{Pl}}$, i.e. a positive exponential. On the other hand, if $\varphi \equiv \frac{\log\phi}{4\sqrt{3\pi G}}$ is being considered, then $A(\varphi) = e^{-\frac{1}{\sqrt{6}}\varphi/M_{Pl}} = e^{-\frac{k}{2}\varphi/M_{Pl}}$, i.e. a negative exponential.

One needs to check whether our model fulfills the potential $V(\varphi)$ requirements for a chameleon or not. This is where problems arise. Let us analyze each case separately.

1. In the first case, we have:

$$V(\varphi) = Ce^{k\varphi/M_{Pl}} \;\; ; \text{and} \;\; A(\varphi) = e^{\frac{k}{2}\varphi/M_{Pl}} \;\; ; \text{for} \;\; k > 0 \tag{6.62}$$

   In the first section of this chapter, the conditions on $V(\varphi)$ when $A(\varphi)$ is a positive exponential were shown to be $V_{,\varphi} < 0$, $V_{,\varphi\varphi} > 0$ and $V_{,\varphi\varphi\varphi} < -\rho A_{,\varphi\varphi\varphi} < 0$ in a region around the field value φ_{min} which minimizes the effective potential, see equation (6.15). In our case, if $C \equiv \frac{\Lambda}{8\pi G} > 0$, it is straight forward to see that $V_{,\varphi} = \frac{Ck}{M_{Pl}} e^{k\varphi/M_{Pl}} > 0$

anywhere, violating the first requirement. On the other hand, if C is taken to be < 0, then $V_{,\varphi} = \frac{Ck}{M_{Pl}} e^{k\varphi/M_{Pl}} < 0$ anywhere, fulfilling the first condition. But $V_{,\varphi\varphi} = \frac{Ck^2}{M_{Pl}^2} e^{k\varphi/M_{Pl}} < 0$ violates the second.

2. In the second case, we have:

$$V(\varphi) = Ce^{-k\varphi/M_{Pl}} \;\; ; \text{and} \;\; A(\varphi) = e^{-\frac{k}{2}\varphi/M_{Pl}} \;\; ; \text{for} \;\; k > 0 \tag{6.63}$$

The conditions on $V(\varphi)$ when $A(\varphi)$ is a negative exponential were $V_{,\varphi} > 0$, $V_{,\varphi\varphi} > 0$ and $V_{,\varphi\varphi\varphi} > -\rho A_{,\varphi\varphi\varphi} > 0$ around φ_{min} again, see equation (6.18). If $C > 0$, then $V_{,\varphi} = -\frac{Ck}{M_{Pl}} e^{-k\varphi/M_{Pl}} < 0$, violating the first condition. Conversely, if $C < 0$, then $V_{,\varphi} = -\frac{Ck}{M_{Pl}} e^{-k\varphi/M_{Pl}} > 0$, which agrees with the first requirement. Nonetheless, $V_{,\varphi\varphi} = \frac{Ck^2}{M_{Pl}^2} e^{-k\varphi/M_{Pl}} < 0$ and the second condition is violated.

In summary, $V(\varphi)$ does not have the desired form for a chameleon potential in any of the two cases. There are two possibilities to fix this problem: either the potential has to be different, i.e. it must come from a different 5-dimensional action other than the cosmological constant one, or our ansatz to how 4-dimensional matter fields couple to the metric must change, yielding a different conformal factor $A = A(\varphi)$. Both possibilities are going to be explored below.

6.3.2 Attempts to Fix the Problems

First, one could drop the $\hat{S}_\Lambda$ piece of the action from S_{KK} and come up with another potential. It could be considered, for instance, a potential which has the shape of the steep drop part of the KKLT inspired potential, $V(\varphi) = M_{Pl}^4 C e^{\alpha e^{-\beta\varphi/M_{Pl}}}$ where $C, \alpha, \beta > 0$, since this is well-motivated from string theory as mentioned before. Unfortunately, this potential is not naturally motivated from a reduction from 5 dimensions as we were able to do with the 5D cosmological constant action. Even if one finds a 5D action that yields such a potential when dimensionally reduced, the 5-dimensional theory will certainly not be as simple as just the extension of the well-known 4-dimensional Einstein-Hilbert plus cosmological constant action to 5 dimensions. Nevertheless, let us consider this potential and show that when considering the redefinition of the scalar field $\varphi \equiv -\frac{\log \phi}{4\sqrt{3\pi G}} \Rightarrow A(\varphi) = e^{\frac{1}{\sqrt{6}}\varphi/M_{Pl}} \equiv e^{g\varphi/M_{Pl}}$, for $g \equiv \frac{1}{\sqrt{6}} > 0$, this potential fulfills all the requisites for a chameleon potential.

(i). The first derivative is:

$$V_{,\varphi} = -M_{Pl}^3 C\alpha\beta e^{\alpha e^{-\beta\varphi/M_{Pl}}} e^{-\beta\varphi/M_{Pl}} < 0\,; \tag{6.64}$$

valid for any value of φ, including the region around φ_{min}.

(ii). The second derivative now:

$$\begin{aligned} V_{,\varphi\varphi} &= -M_{Pl}^3 C\alpha\beta e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} \left(-\frac{\beta}{M_{Pl}} - \frac{\alpha\beta}{M_{Pl}} e^{-\beta\varphi/M_{Pl}}\right) \\ &= M_{Pl}^2 C\alpha\beta^2 e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} (1 + \alpha e^{-\beta\varphi/M_{Pl}}) > 0\,; \end{aligned} \tag{6.65}$$

also valid for any φ.

(iii). And the third one:

$$\begin{aligned} V_{,\varphi\varphi\varphi} =& M_{Pl}^2 C\alpha\beta^2 e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} \Big[-\frac{\beta}{M_{Pl}}(1 + \alpha e^{-\beta\varphi/M_{Pl}}) \\ & - \frac{\alpha\beta}{M_{Pl}} e^{-\beta\varphi/M_{Pl}} (1 + \alpha e^{-\beta\varphi/M_{Pl}}) - \frac{\alpha\beta}{M_{Pl}} e^{-\beta\varphi/M_{Pl}} \Big] \\ =& - M_{Pl} C\alpha\beta^3 e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} (1 + 3\alpha e^{-\beta\varphi/M_{Pl}} + \alpha^2 e^{-2\beta\varphi/M_{Pl}}) < 0 \end{aligned} \tag{6.66}$$

While we have $A_{,\varphi\varphi\varphi} = (g/M_{Pl}^3)\, e^{g\varphi/M_{Pl}}$ for the conformal factor. Thus, the third condition on the potential requires that:

$$V_{,\varphi\varphi\varphi} < -\rho A_{,\varphi\varphi\varphi} \Rightarrow M_{Pl} C \alpha \beta^3 e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} \times$$
$$\times \left(1 + 3\alpha e^{-\beta\varphi/M_{Pl}} + \alpha^2 e^{-2\beta\varphi/M_{Pl}}\right) > \rho \left(\frac{g}{M_{Pl}}\right)^3 e^{g\varphi/M_{Pl}}\,; \tag{6.67}$$

which is a messy constraint on the parameters that must be respected around φ_{min}. If one only considers derivatives of the double exponential, as it was done for the KKLT inspired model, this becomes a little bit less messy:

$$M_{Pl} C \alpha^3 \beta^3 e^{-3\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} > \rho \left(\frac{g}{M_{Pl}}\right)^3 e^{g\varphi/M_{Pl}} \tag{6.68}$$

The above constraint can be simplified if one notes that:

$$\frac{dV_{eff}(\varphi_{min})}{d\varphi} = 0 \Rightarrow M_{Pl}^3 C \alpha \beta e^{-\beta\varphi_{min}/M_{Pl}} e^{\alpha e^{-\beta\varphi_{min}/M_{Pl}}} = \rho \left(\frac{g}{M_{Pl}}\right) e^{g\varphi_{min}/M_{Pl}}\,; \tag{6.69}$$

which is only valid, strictly speaking, at $\varphi = \varphi_{min}$, but it is expected to hold approximately true in a region close to it. Thus, back to (6.68) one gets:

$$M_{Pl} C \alpha^3 \beta^3 e^{-3\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}} \gtrsim \left(\frac{g}{M_{Pl}}\right)^2 C \alpha \beta e^{-\beta\varphi/M_{Pl}} e^{\alpha e^{-\beta\varphi/M_{Pl}}}$$
$$\alpha^2 \beta^2 e^{-2\beta\varphi/M_{Pl}} \gtrsim g^2\,; \tag{6.70}$$

where $\varphi \simeq \varphi_{min}$.

Equations (6.64) and (6.65) plus the imposition of the constraint (6.70) close to φ_{min} on the model's parameters perfectly match the conditions (6.15). This is therefore an adequate chameleon potential.

The other possibility to fix the problem we have faced last subsection is to change our ansatz of how the 4-dimensional matter action should look like. Consider the potential obtained via dimensional reduction of the cosmological constant action with the redefinition $\varphi \equiv \frac{\log\phi}{4\sqrt{3\pi G}} \Rightarrow V(\varphi) = \frac{\Lambda}{8\pi G} e^{-\frac{\sqrt{6}}{3}\varphi/M_{Pl}} \equiv M_{Pl}^4 C e^{-k\varphi/M_{Pl}}$, with $\Lambda > 0$ (thus $C, k > 0$). In this case, one originally got a negative exponential as the conformal factor and the potential was not appropriate for a chameleon. However, if we instead consider a positive exponential conformal factor, $A(\varphi) = e^{g\varphi/M_{Pl}}$ for $g > 0$, the potential fits the requirements.

(i). The first derivative is:

$$V_{,\varphi} = -M_{Pl}^3 C k e^{-k\varphi/M_{Pl}} < 0\,; \tag{6.71}$$

valid for any value of φ.

(ii). The second derivative is:

$$V_{,\varphi\varphi} = M_{Pl}^2 C k^2 e^{-k\varphi/M_{Pl}} > 0\,; \tag{6.72}$$

also valid for any φ.

(iii). And the third:

$$V_{,\varphi\varphi\varphi} = -M_{Pl} C k^3 e^{-k\varphi/M_{Pl}} < 0 \tag{6.73}$$

While the third derivative of the conformal factor is $A_{,\varphi\varphi\varphi} = (g/M_{Pl})^3 e^{g\varphi/M_{Pl}}$. One can show that $V'_{eff}(\varphi_{min}) = 0 \Rightarrow e^{(g+k)\varphi_{min}/M_{Pl}} = M_{Pl}^4 \frac{Ck}{\rho g}$, leading to:

$$V_{,\varphi\varphi\varphi} < -\rho A_{,\varphi\varphi\varphi} \therefore M_{Pl} C k^3 e^{-k\varphi/M_{Pl}} > \rho\left(\frac{g}{M_{Pl}}\right)^3 e^{g\varphi/M_{Pl}} \Rightarrow$$
$$e^{(g+k)\varphi/M_{Pl}} < M_{Pl}^4 \frac{Ck^3}{\rho g^3}\,; \tag{6.74}$$

which should apply to a region close to the minimum, i.e. $\varphi \simeq \varphi_{min}$, thus:

$$e^{(g+k)\varphi/M_{Pl}} \simeq e^{(g+k)\varphi_{min}/M_{Pl}} = M_{Pl}^4 \frac{Ck}{\rho g} \lesssim M_{Pl}^4 \frac{Ck^3}{\rho g^3} \therefore \frac{k}{g} \gtrsim 1 \, ; \tag{6.75}$$

one ends up with a simple constraint on the parameters.

Equations (6.71) and (6.72) plus the imposition of the constraint (6.75) fit the requisites (6.15), and hence this is also a satisfactory chameleon potential.

6.3.3 Experimental Constraints on the Models

Even though the above $A(\varphi)$-$V(\varphi)$ pairs are not as well-motivated as in the unsuccessful original idea of deriving a chameleon model just from gravity plus the cosmological constant in 5 dimensions, they match the conditions with just a few free parameters. Thus, they can be seen as phenomenological models and one can find out the experimental constraints on their parameters similarly to what was done for the KKLT inspired model.

Equation (6.39) translates naturally to these models, since the shape of the conformal factor in both cases is precisely the one assumed to get to this equation, namely $A(\varphi) = e^{g\varphi/M_{Pl}}$ with $g \sim \mathcal{O}(1)$, therefore:

$$\varphi_{vac} - \varphi_{tm} \lesssim 10^{-28} M_{Pl} \tag{6.76}$$

We will calculate the laboratory constraints for each model individually:

1. First, consider the model where $V(\varphi) = M_{Pl}^4 C e^{\alpha e^{-\beta\varphi/M_{Pl}}}$ and $A(\varphi) = e^{g\varphi/M_{Pl}}$ with $g \equiv 1/\sqrt{6} \sim \mathcal{O}(1)$.

   Note that the above potential was constructed to have the same shape as the steep exponential part of the KKLT inspired potential. Besides, the conformal factor also has the same shape. Therefore, one can avoid doing the calculating from the beginning by rewriting $V(\varphi)$ and $A(\varphi)$ in a more familiar form. First, we define $R/R_* \equiv e^{g\varphi/M_{Pl}}$

thus $A(R) = R/R_*$, where again R_* is the field value of reference such that $A(R = R_*) = A(\varphi = 0) = 1$.

The potential in terms of R becomes:

$$\begin{aligned} V(R) &= M_{Pl}^4 C e^{\alpha R^{-\frac{\beta}{g}} R_*^{\frac{\beta}{g}}} \equiv M_{Pl}^4 C e^{\tilde{\alpha} R^{-\tilde{\beta}}} = M_{Pl}^4 C e^{\tilde{\alpha} R^{-\tilde{\beta}}} e^{\tilde{\alpha} R_*^{-\tilde{\beta}}} e^{-\tilde{\alpha} R_*^{-\tilde{\beta}}} \\ &\equiv M_{Pl}^4 \tilde{C} e^{\tilde{\alpha}\left(R^{-\tilde{\beta}} - R_*^{-\tilde{\beta}}\right)}; \end{aligned} \tag{6.77}$$

where we have redefined the constants $\tilde{\beta} \equiv \beta/g$, $\tilde{\alpha} \equiv \alpha R_*^{\beta/g}$ and $\tilde{C} \equiv C e^{\tilde{\alpha} R_*^{-\tilde{\beta}}} = C e^{\alpha}$ in order to make $V(R)$ look exactly like that of equation (6.34) for $R < R_*$. Comparing against the KKLT inspired model, it is straight forward to see that now the variables v, γ and k become $v \to \tilde{C}$, $\gamma \to \tilde{\alpha}$ and $k \to \tilde{\beta}$. The -1 term inside the square brackets of equation (6.34) was introduced in that case to enforce $V(R_*) = 0$ to connect with the quadratic piece of the potential. However, the potential is not defined by parts now and we can drop this constant since it doesn't change the derivation of the constraint on the model's parameters. The analog of equation (6.46) is, using the fact that $g \sim \mathcal{O}(1)$:

$$\frac{R_*^{\tilde{\beta}}}{\tilde{\alpha}\tilde{\beta}} \lesssim 10^{-30} \therefore \frac{1}{\alpha\beta} \lesssim 10^{-30}; \tag{6.78}$$

where we just plugged in the definitions of $\tilde{\alpha}$ and $\tilde{\beta}$.

Similarly, equation (6.47) becomes:

$$m^2 \simeq g^2 \frac{\tilde{\alpha}\tilde{\beta}}{R_*^{\tilde{\beta}}} \frac{\rho}{M_{Pl}^2} = g^2 \alpha\beta \frac{\rho}{M_{Pl}^2} \gtrsim g^2 10^{30} \frac{\rho}{M_{Pl}^2} \therefore m \gtrsim 10^{15} \frac{\sqrt{\rho}}{M_{Pl}} \tag{6.79}$$

In summary, since the shape of the potential is identical to the one we have seen before, we could jump the steps of the calculation (since they would be exactly the same) all the way to the final constraint $\alpha\beta \gtrsim 10^{30}$ which, unsurprisingly, leads to the same lower bound to the scalar mass in terms of the energy density ρ of a high-density region taken into consideration. Recall that this bound coming from lab experiments was valid for $\rho > \rho_*$ in the KKLT toy model. To get the analogous of this bound valid for cosmological scales, we will need to rederive (6.55) but now without considering the quadratic part of the potential, since the currently analyzed potential is not defined by parts anymore.

2. The second KK model has $V(\varphi) = M_{Pl}^4 C e^{-k\varphi/M_{Pl}}$ and $A(\varphi) = e^{g\varphi/M_{Pl}}$, where $C, k, g > 0$. Originally, we have obtained $k = \sqrt{6}/3$ by compactifying the 5-dimensional KK theory (from the cosmological constant 5D action). However, let us allow any value for k which will be constrained by experimental data. Indeed, since we have forced the conformal factor to have the desired form, this can be considered a phenomenological model which was just motivated by the original Kaluza-Klein theory. Recall that one must require $k \gtrsim g$ from the third condition imposed on $V(\varphi)$, see the last subsection. This time the shape of the potential cannot be rewritten as the steep drop piece of the KKLT inspired potential, therefore one must do the calculations from scratch, following the same steps of the last section.

   We shall start by calculating m_{vac}^2 in terms of φ_{vac}, the field value which minimizes the effective potential inside the vacuum chamber, and then inverting the expression to get $\varphi_{vac} = \varphi_{vac}(m_{vac}^2)$. But, first, allow us to calculate this expression for any medium, as it will be useful later.

$$
\begin{aligned}
m^2 \equiv \frac{d^2 V_{eff}}{d\varphi^2} &= \frac{d}{d\varphi}\left(-M_{Pl}^3 C k e^{-k\varphi/M_{Pl}} + \frac{\rho}{M_{Pl}} g e^{g\varphi/M_{Pl}}\right) \\
&= M_{Pl}^2 C k^2 e^{-k\varphi/M_{Pl}} + \frac{\rho}{M_{Pl}^2} g^2 e^{g\varphi/M_{Pl}} \qquad (6.80)
\end{aligned}
$$

   Applying this result to the vacuum chamber case:

$$m_{vac}^2 \simeq M_{Pl}^2 C k^2 e^{-k\varphi_{vac}/M_{Pl}}\,; \tag{6.81}$$

where, since $\rho_{vac}/M_{Pl}^4 \ll 1$, we have considered only the dominant, ρ_{vac}-independent, term.

Inverting the above relation, one gets:

$$\varphi_{vac} \simeq -\frac{M_{Pl}}{k}\log\left(\frac{m_{vac}^2}{M_{Pl}^2 C k^2}\right) = \frac{M_{Pl}}{k}\log\left(\frac{M_{Pl}^2 C k^2}{m_{vac}^2}\right) \tag{6.82}$$

The next step is to find an expression for φ_{tm}, the field value which minimizes the effective potential inside the test mass, in terms of the known density ρ_{tm}. Again, we will do the calculation for a generic medium, for later purposes, and then apply it to the test mass case.

$$\frac{dV_{eff}}{d\varphi} = 0 \therefore -M_{Pl}^3 C k e^{-k\varphi/M_{Pl}} + \frac{\rho}{M_{Pl}} g e^{g\varphi/M_{Pl}} = 0 \Rightarrow$$
$$\rho = \frac{M_{Pl}^4 C k}{g} e^{-(k+g)\varphi/M_{Pl}} \tag{6.83}$$

Then, inside the test mass, we obtain:

$$\varphi_{tm} = -\frac{M_{Pl}}{k+g}\log\left(\frac{\rho_{tm} g}{M_{Pl}^4 C k}\right) \tag{6.84}$$

Equations (6.76), (6.82) and (6.84) lead to:

$$\varphi_{vac} - \varphi_{tm} = \frac{M_{Pl}}{k}\log\left(\frac{M_{Pl}^2 C k^2}{m_{vac}^2}\right) + \frac{M_{Pl}}{k+g}\log\left(\frac{\rho_{tm} g}{M_{Pl}^4 C k}\right) \lesssim 10^{-28} M_{Pl} \tag{6.85}$$

To get a simpler analytical constraint, we can further consider $k \gg g \sim \mathcal{O}(1)$ (instead of just $k \gtrsim g$ which is already required). Then, since $k + g \simeq k$, the above constraint becomes:

$$\varphi_{vac} - \varphi_{tm} = \frac{M_{Pl}}{k} \log\left(kg\frac{\rho_{tm}}{m_{vac}^2 M_{Pl}^2}\right) \lesssim 10^{-28} M_{Pl} \Rightarrow$$
$$\frac{1}{k}\log\left(kg\frac{\rho_{tm}}{m_{vac}^2 M_{Pl}^2}\right) \sim \frac{1}{k}\left[\log k - 23\log 10\right] \lesssim 10^{-28} \Rightarrow \frac{1}{k} \lesssim 10^{-30} \therefore k \gtrsim 10^{30}\,; \quad (6.86)$$

where we have used the previously given values $\rho_{tm} \sim 10^{-89} M_{Pl}^4$ and $m_{vac} \sim 10^{-33} M_{Pl}$.

Note that the obtained bound on k is consistent with the previously used assumption $k \gg g$.

Moreover, from equations (6.80) and (6.83), we get that for a generic medium whose ρ is still $\ll M_{Pl}^4$ (similar to the vacuum chamber case):

$$\frac{m^2}{\rho} \simeq \frac{kg}{M_{Pl}^2} \sim \frac{k}{M_{Pl}^2} \gtrsim \frac{10^{30}}{M_{Pl}^2} \therefore m \gtrsim 10^{15}\frac{\sqrt{\rho}}{M_{Pl}}\,; \quad (6.87)$$

which is the same bound obtained in the KKLT inspired model and in the first KK motivated model.

We must point out that having $\rho/M_{PL}^4 \ll 1$ may actually not be enough to drop the second term in equation (6.80). Notice that the first term is $\propto e^{-k\varphi/M_{Pl}}$, so it might be considerably suppressed by the exponential. Moreover, we didn't arrive at a constraint to C, which affects our assumption. Nevertheless, we still consider the first term as the dominant one and the complete assumption we are making is:

$$\frac{\rho}{M_{Pl}^4} \ll \frac{Ck^2}{g^2} e^{-(k+g)\varphi/M_{Pl}}\,; \quad (6.88)$$

which is also assumed in the vacuum chamber case.

The above constraints were obtained from lab experiments, as anticipated. We must now take into consideration the constraint coming from the screening of the galaxy which was seen, equation (6.51), to translate to:

$$\varphi_{cosm} - \varphi_{ss} \lesssim 10^{-6} M_{Pl} \tag{6.89}$$

Our goal is to arrive at a lower bound on m_{cosm} analogous to (6.55) for each KK motivated model to check whether they can indeed have relevant implications on large scales or not.

1. It was seen that the first model could be rewritten as $V(R) = M_{Pl}^4 \tilde{C} e^{\tilde{\alpha}\left(R^{-\tilde{\beta}} - R_*^{-\tilde{\beta}}\right)}$ and $A(R) = R/R_*$ in order to match exactly the behavior of the steep exponential part of the potential (6.34). The difference between the two models is that the KKLT motivated one was defined by parts and, since R_{cosm} and R_{ss} fell in the quadratic part of the potential, the constraint coming from (6.51) was derived using this part of the potential. In the current case, the potential is the same for any value of R, so the derivation has to be done from scratch.

   Note, though, that the approximation (6.40) is still valid since the potential drops very steeply and we can assume all relevant ($\leq R_{cosm}$) values of R are close together and therefore close to R_*, i.e. $|R - R_*| \ll R_*$. This means:

$$\varphi_{cosm} - \varphi_{ss} \simeq \frac{M_{Pl}}{g}\left(\frac{R_{cosm} - R_*}{R_*} - \frac{R_{ss} - R_*}{R_*}\right) = \frac{M_{Pl}}{g}\left(\frac{R_{cosm} - R_{ss}}{R_*}\right) \tag{6.90}$$

   Recall that R_{cosm} (R_{ss}) is related to φ_{cosm} (φ_{ss}) which is the value that minimizes the effective potential on cosmological scales (on the solar system). Therefore, one finds an expression for R_{cosm} (R_{ss}) by requiring $\left.\frac{dV_{eff}}{d\varphi}\right|_{\varphi=\varphi_{cosm}} = 0 \therefore \left.\frac{dV_{eff}}{dR}\right|_{R=R_{cosm}} = 0$ (the analogous holds true for the solar system case) and solving the equation for R_{cosm} (R_{ss}). This is exactly what was done in equation (6.44) for the case of the test mass, considering $R_* - R_{tm} \ll R_*$. The same calculation is also valid for cosmological scales and the solar system. So, recalling that in the current case it was found that $v \to \tilde{C} = Ce^{\alpha}$, $\gamma \to \tilde{\alpha} = \alpha R_*^{\beta/g}$ and $k \to \tilde{\beta} = \beta/g$, one gets:

$$\frac{R_{cosm}-R_*}{R_*} \simeq \frac{R_*^{\tilde{\beta}}}{\tilde{\alpha}\tilde{\beta}} \log\left(\frac{M_{Pl}^4 \tilde{C}}{\rho_{cosm}} \frac{\tilde{\alpha}\tilde{\beta}}{R_*^{\tilde{\beta}}}\right) \sim \frac{1}{\alpha\beta} \log\left(\frac{M_{Pl}^4}{\rho_{cosm}} C e^{\alpha} \alpha\beta\right) ; \tag{6.91}$$

$$\frac{R_{ss}-R_*}{R_*} \simeq \frac{R_*^{\tilde{\beta}}}{\tilde{\alpha}\tilde{\beta}} \log\left(\frac{M_{Pl}^4 \tilde{C}}{\rho_{ss}} \frac{\tilde{\alpha}\tilde{\beta}}{R_*^{\tilde{\beta}}}\right) \sim \frac{1}{\alpha\beta} \log\left(\frac{M_{Pl}^4}{\rho_{ss}} C e^{\alpha} \alpha\beta\right) ; \tag{6.92}$$

where after the $\sim$ symbol on both lines we have already considered $g \sim \mathcal{O}(1)$.

Combining equations (6.90), (6.91) and (6.92) with the constraint (6.89), it is straight forward to get:

$$\varphi_{cosm} - \varphi_{ss} \sim \frac{M_{Pl}}{\alpha\beta} \log\left(\frac{\rho_{ss}}{\rho_{cosm}}\right) \sim \frac{M_{Pl}}{\alpha\beta}(5\log 10) \lesssim 10^{-6} M_{Pl} \therefore \alpha\beta \gtrsim 10^{7} \tag{6.93}$$

Note that, since the potential is not defined by parts as in the KKLT case, the above bound constrains the same parameters as (6.78). Moreover, since (6.93) is less stringent than the lab constraint, we must require that $\alpha\beta \gtrsim 10^{30}$ and the galaxy constraint will automatically be respected.

The final step is to figure out the bound $\alpha\beta \gtrsim 10^{30}$ implications on the scalar mass on cosmological scales. Note that deriving an expression for m^2_{cosm} follows the same steps as the derivation of m^2_{vac} from (6.41) to (6.43). Therefore, we can write:

$$\begin{aligned} m^2_{cosm} &\simeq g^2 M_{Pl}^2 \tilde{C} \left(\frac{\tilde{\alpha}\tilde{\beta}}{R_*^{\tilde{\beta}}}\right)^2 e^{-\frac{\tilde{\alpha}\tilde{\beta}}{R_*^{\tilde{\beta}}}\left(\frac{R_{cosm}-R_*}{R_*}\right)} \sim M_{Pl}^2 C e^{\alpha} \alpha^2\beta^2 \left(\frac{\rho_{cosm}}{M_{Pl}^4 C e^{\alpha}\alpha\beta}\right) \\ &\sim \alpha\beta \frac{\rho_{cosm}}{M_{Pl}^2} \sim \alpha\beta H_0^2 ; \end{aligned} \tag{6.94}$$

where, after the $\sim$ symbol in the first line, the expression (6.91) was plugged in and in the second line it was used the fact $\rho_{cosm} \sim H_0^2 M_{Pl}^2$ coming from the first Friedmann equation (1.21), neglecting the curvature energy density.

Combining the above expression for m_{cosm} with the constraint (6.78), one realizes that:

$$m_{cosm} \gtrsim 10^{15} H_0\,; \tag{6.95}$$

which shows us that, although this model's scalar field seemed suitable as a dark energy candidate, it cannot have relevant implications for structure formation.

2. The second KK motivated model has a defining potential $V(\varphi) = M_{Pl}^4 C e^{-k\varphi/M_{Pl}}$ and a conformal factor $A(\varphi) = e^{g\varphi/M_{Pl}}$. Therefore, one finds φ_{cosm} to be:

$$\begin{gathered} \left.\frac{dV_{eff}}{d\varphi}\right|_{\varphi=\varphi_{cosm}} = \left.\frac{dV}{d\varphi}\right|_{\varphi=\varphi_{cosm}} + \rho_{cosm}\left.\frac{dA}{d\varphi}\right|_{\varphi=\varphi_{cosm}} = 0 \Rightarrow \\ M_{Pl}^3 C k e^{-k\varphi_{cosm}/M_{Pl}} = \rho_{cosm}\frac{g}{M_{Pl}} e^{g\varphi_{cosm}/M_{Pl}} \therefore e^{k\varphi_{cosm}/M_{Pl}} \sim Ck\frac{M_{Pl}^4}{\rho_{cosm}} \Rightarrow \\ \varphi_{cosm} \sim \frac{M_{Pl}}{k}\log\left(Ck\frac{M_{Pl}^4}{\rho_{cosm}}\right)\,; \end{gathered} \tag{6.96}$$

where in the second line, after the $\therefore$ symbol, we have used the facts $k \gg g$ and $g \sim \mathcal{O}(1)$.

A completely analogous calculation could have been done to get the field value which minimizes the effective potential on solar system scales, φ_{ss}, as:

$$\varphi_{ss} \sim \frac{M_{Pl}}{k}\log\left(Ck\frac{M_{Pl}^4}{\rho_{ss}}\right) \tag{6.97}$$

Plugging (6.96) and (6.97) into the galaxy constraint (6.89) we get:

$$\varphi_{cosm} - \varphi_{ss} \sim \frac{M_{Pl}}{k}\log\left(\frac{\rho_{ss}}{\rho_{cosm}}\right) \sim \frac{M_{Pl}}{k}(5\log 10) \lesssim 10^{-6} M_{Pl} \therefore k \gtrsim 10^{7} \tag{6.98}$$

Also in this model, since the potential is not defined by parts, we have obtained a constraint on the same parameter as in the lab bound case, equation (6.86). Once again, the laboratory constraint prevails and we must require $k \gtrsim 10^{30}$, in order not to violate any experimental bound. The chameleon mass on cosmological scales is:

$$m_{cosm}^2 \equiv \left.\frac{d^2 V_{eff}}{d\varphi^2}\right|_{\varphi=\varphi_{cosm}} = M_{Pl}^2 C k^2 e^{-k\varphi_{cosm}/M_{Pl}} + \rho_{cosm}\left(\frac{g}{M_{Pl}}\right)^2 e^{g\varphi_{cosm}/M_{Pl}}$$
$$\simeq M_{Pl}^2 C k^2 e^{-k\varphi_{cosm}/M_{Pl}} \sim M_{Pl}^2 C k^2 \frac{\rho_{cosm}}{C k M_{Pl}^4} \sim k H_0^2\,; \tag{6.99}$$

where to get to the second line we have made the same assumption on the value of C as in the laboratory experiment case, but now for cosmological scales:

$$\frac{\rho_{cosm}}{M_{Pl}^4} \ll \frac{Ck^2}{g^2} e^{-(k+g)\varphi_{cosm}/M_{Pl}}\,. \tag{6.100}$$

Furthermore, it was considered $g \sim \mathcal{O}(1)$ and equation (6.96) was used in the second line.

Substituting the bound on k found in equation (6.86) in the expression for m_{cosm}, equation (6.99), one finds:

$$m_{cosm} \gtrsim 10^{15} H_0\,; \tag{6.101}$$

which is again not the expected characteristic of a quintessence field on cosmological scales.

To sum up, we have calculated the experimental bounds on both KK inspired models' parameters. In the first case, the most stringent bound was found to be $\alpha\beta \gtrsim 10^{30}$ while in the second we found $k \gtrsim 10^{30}$. In both cases, these bounds ended up spoiling the desired behavior of the field on cosmological scales since the scalar interaction range, m_{cosm}^{-1}, would be too short to generate important effects on the cosmological evolution. This is the same problem obtained in [83] for the actual KKLT modified potential – not to be confused with the phenomenological KKLT motivated model of (6.34) – see equation (6.56). Actually, the bounds (6.95) and (6.101) are exactly the same as (6.56).

It was shown in [89] that adding a second exponential to the KKLT superpotential allows us to reproduce the bound $m_{cosm} \gtrsim 10^3 H_0$ obtained for the phenomenological model. However, this would be realized by sort of fine-tuning the parameters of the superpotential. If one intends to change the above KK motivated models looking for cosmologically interesting chameleons, the addition of new free parameters may be necessary to make the behavior of the potential considerably different from the φ-region relevant for lab experiments to the relevant region for galaxy and cosmological densities, as in the case of the phenomenological KKLT potential (6.34).

Chapter 7

Conclusion

The three cosmological models analyzed in this thesis showed dark energy features in quite distinct fashions. The first of them, string (brane) gas cosmology, is the only intrinsically stringy model out of the three. It was shown that this model has many interesting new features concerning the early universe like, for instance, the emergence of three large spatial dimensions (along with a hierarchy of dimensions in the brane case). As for the late universe, there are not many changes though. A dark energy feature is found in this model only in the Hagedorn phase and if one considers a D3-brane in a non-critical $d = 3$ number of dimensions. Therefore, no real insights are gained on the dark energy we observe nowadays. To approach the cosmological constant problem within string gas cosmology (SGC) one might motivate a non-perturbative scalar potential for the 4-dimensional effective theory and follow the steps of [83] to construct a chameleon model. In fact, [70] already suggests this possible implementation of chameleon scalars within SGC.

Holographic cosmology has been growing as an alternative to inflation. Even if we consider a non-geometrical cosmological phase at early times, CMBR observations were shown to be reproduced via calculations on the weakly coupled dual field theory. Although this model is concerned mostly about the early universe, it was shown it also has important implications to late time cosmology. The most relevant implication to this thesis is the shrinking of the cosmological constant value. It was shown that, even if we consider a cosmological constant as high as its expected value coming from QFT calculations (see chapter 2), one can dilute its effect over time to match the currently observed dark energy density, from the observation that time evolution means inverse RG flow (from low to high energy scales) in the dual QFT. This holographic solution to the cosmological constant problem has its caveats though, as discussed at the end of chapter 5.

The chameleon model is string-independent by construction. However, it is very useful to string theory models since one avoids having to stabilize all the moduli of the dimensionally reduced theory. An attempt of embedding the chameleon idea in string theory was proposed in [83]. It was found that a phenomenological potential inspired by the KKLT construction could satisfy experimental constraints while having the desired features of a quintessence model. However, the scenario was not the same for the KKLT potential itself. A later modification of the potential was shown to work by adding another free parameter that effectively changes the form of the potential from small to large scales [89].

Based on the chameleon model, we have presented some new calculations trying to interpret the scalar field obtained in the original Kaluza-Klein theory as a chameleon. Some problems were found and possible solutions to them were proposed. In the end, two phenomenological models were considered on which we have imposed experimental constraints in analogy to what was done in [83]. Unfortunately, the scalar fields were shown not to present all the desired features to have had important implications in structure formation. The undesired bound found on the mass of the chameleon on cosmological distances was the same as the one obtained for the KKLT potential. It seems like the proposed potentials do not present two needed different behaviors for lab experiment scales and cosmological scales like the phenomenological KKLT inspired potential of equation (6.34).

To conclude, it is hard to find common ground for all the models. It starts by their distinct origins: while SGC considers a simplified gas of non-interacting strings, the other two models are string independent. Holographic cosmology is based on a duality first realized in string theory context but since, as it is, its UV completion is yet unknown, the setup is string independent. As far as the cosmological constant problem is concerned it was already discussed that SGC by itself has little to tell us. The solution to the problem within holographic cosmology is in a sense more complete since one does not have to assume there is some unknown mechanism that makes the vacuum energy density vanish, the cosmological constant itself would shrink down to the currently observed value. Nonetheless, we don't understand the precise mechanism by which the cosmological constant is shrunk in the bulk. On the other hand, the chameleon solution to the problem is realized via a quintessence model which is well-defined in the bulk, but one has to assume a vanishing vacuum energy density.

Appendix A

Deriving the Einstein Equations

In chapter 1 we stated that the variation of the metric:

$$S = S_{E-H} + S_\Lambda + S_m = \int d^4x\sqrt{-g}\left[\frac{1}{2\kappa}(R - 2\Lambda) + \mathcal{L}_m\right]; \tag{A.1}$$

with respect to a change in the inverse metric $g^{\mu\nu}$ gives us the Einstein equations of equation (1.6). Here we will derive that in full detail. We will present several pieces of calculation that will be necessary in order to calculate the variation of S.

(i). We will start by varying the matter action:

$$\delta S_m = \delta\left(\int d^4x\sqrt{-g}\mathcal{L}_m\right) = \int d^4x\frac{\delta S_m}{\delta g^{\mu\nu}}\delta g^{\mu\nu} = \int d^4x\left(-\frac{\sqrt{-g}}{2}T_{\mu\nu}\right)\delta g^{\mu\nu}; \tag{A.2}$$

where we recall that $T_{\mu\nu} \equiv \frac{-2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}$.

(ii). Now, recalling that we can rewrite the determinant of the metric as $g \equiv \det(g_{\mu\nu}) = e^{\mathrm{Tr}[\log(g_{\mu\nu})]}$, thus:

$$\delta\log g = \frac{1}{g}\delta g = \delta\,\mathrm{Tr}[\log(g_{\mu\nu})] = \mathrm{Tr}\big[(g_{\rho\mu})^{-1}\delta g_{\mu\nu}\big] = g^{\mu\nu}\delta g_{\mu\nu} \therefore \delta g = gg^{\mu\nu}\delta g_{\mu\nu}; \tag{A.3}$$

where we have used the Jacobi's formula to calculate the derivative of a determinant.

(iii). Since we want to calculate the variation of the action with respect to a change in $g^{\mu\nu}$, not exactly in the metric, we need to relate $\delta g_{\mu\nu}$ of the last item to $\delta g^{\mu\nu}$.

$$\begin{aligned}\delta g_{\mu\nu} = \delta(g^{\rho\sigma} g_{\rho\mu} g_{\sigma\nu}) &= \delta g^{\rho\sigma} g_{\rho\mu} g_{\sigma\nu} + \delta g_{\rho\mu} g^{\rho\sigma} g_{\sigma\nu} + \delta g_{\sigma\nu} g^{\rho\sigma} g_{\rho\mu} \\ &= \delta g^{\rho\sigma} g_{\rho\mu} g_{\sigma\nu} + \delta g_{\rho\mu} \delta^{\rho}{}_{\nu} + \delta g_{\sigma\nu} \delta^{\sigma}{}_{\mu} = \delta g^{\rho\sigma} g_{\rho\mu} g_{\sigma\nu} + 2\delta g_{\mu\nu} \\ &\Rightarrow \delta g_{\mu\nu} = -\delta g^{\rho\sigma} g_{\rho\mu} g_{\sigma\nu} \end{aligned} \tag{A.4}$$

(iv). From the results of equations (A.3) and (A.4) it is easy to calculate:

$$\delta(\sqrt{-g}) = -\frac{1}{2\sqrt{-g}}\delta g = \frac{\sqrt{-g}}{2} g^{\mu\nu}\delta g_{\mu\nu} = -\frac{\sqrt{-g}}{2} g_{\mu\nu}\delta g^{\mu\nu} \tag{A.5}$$

(v). We can calculate the variation of the cosmological constant action now using the result obtained in equation (A.5):

$$\delta S_\Lambda = -\frac{1}{\kappa}\delta\left(\int d^4x\sqrt{-g}\Lambda\right) = \frac{1}{2\kappa}\int d^4x\sqrt{-g}\Lambda g_{\mu\nu}\delta g^{\mu\nu} \tag{A.6}$$

(vi). In order to calculate the variation of the Einstein-Hilbert action, we will see that we need to calculate the variation of the Ricci tensor.

$$\begin{aligned}\delta R_{\mu\nu} &= \partial_\rho \delta\Gamma^{\rho}{}_{\mu\nu} - \partial_\nu \delta\Gamma^{\rho}{}_{\mu\rho} + \Gamma^{\rho}{}_{\sigma\rho}\delta\Gamma^{\sigma}{}_{\mu\nu} + \Gamma^{\sigma}{}_{\mu\nu}\delta\Gamma^{\rho}{}_{\sigma\rho} - \Gamma^{\sigma}{}_{\rho\nu}\delta\Gamma^{\rho}{}_{\mu\sigma} - \Gamma^{\sigma}{}_{\mu\rho}\delta\Gamma^{\rho}{}_{\sigma\nu} \\ &= (\partial_\rho \delta\Gamma^{\rho}{}_{\mu\nu} + \Gamma^{\rho}{}_{\sigma\rho}\delta\Gamma^{\sigma}{}_{\mu\nu} - \Gamma^{\sigma}{}_{\mu\rho}\delta\Gamma^{\rho}{}_{\sigma\nu} - \Gamma^{\sigma}{}_{\rho\nu}\delta\Gamma^{\rho}{}_{\mu\sigma}) - (\partial_\nu \delta\Gamma^{\rho}{}_{\mu\rho} - \Gamma^{\sigma}{}_{\mu\nu}\delta\Gamma^{\rho}{}_{\sigma\rho}) \\ &= \nabla_\rho \delta\Gamma^{\rho}{}_{\mu\nu} - \nabla_\nu \delta\Gamma^{\rho}{}_{\mu\rho} \end{aligned} \tag{A.7}$$

Thus, using the metric compatibility of the covariant derivative equation $\nabla_\rho g^{\mu\nu} = 0$:

$$g^{\mu\nu}\delta R_{\mu\nu} = \nabla_\rho(g^{\mu\nu}\delta\Gamma^{\rho}{}_{\mu\nu}) - \nabla_\nu(g^{\mu\nu}\delta\Gamma^{\rho}{}_{\mu\rho}) = \nabla_\rho(g^{\mu\nu}\delta\Gamma^{\rho}{}_{\mu\nu} - g^{\mu\rho}\delta\Gamma^{\nu}{}_{\mu\nu}) \equiv \nabla_\rho X^\rho\,; \tag{A.8}$$

where we have defined the field X^ρ in the last equality just to simplify further expressions.

(vii). We can now, using the result of equation (A.8), calculate the variation of the Ricci scalar:

$$\delta R = \delta(g^{\mu\nu}R_{\mu\nu}) = \delta g^{\mu\nu}R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} = \delta g^{\mu\nu}R_{\mu\nu} + \nabla_\mu X^\mu \tag{A.9}$$

(viii). However, the full term that we need to calculate the variation of in the Einstein-Hilbert action is:

$$\begin{aligned}\delta(\sqrt{-g}R) &= \delta(\sqrt{-g})R + \sqrt{-g}\delta R \\ &= -\frac{\sqrt{-g}}{2}g_{\mu\nu}\delta g^{\mu\nu}R + \sqrt{-g}\delta g^{\mu\nu}R_{\mu\nu} + \sqrt{-g}\nabla_\mu X^\mu \\ &= -\frac{\sqrt{-g}}{2}g_{\mu\nu}\delta g^{\mu\nu}R + \sqrt{-g}\delta g^{\mu\nu}R_{\mu\nu} + \partial_\mu(\sqrt{-g}X^\mu)\,;\end{aligned} \tag{A.10}$$

where in the first equality we just applied the product rule for a differentiation, in the second equality we used the results of equations (A.5) and (A.9) and in the third equality we used the following fact:

$$\begin{aligned}\sqrt{-g}\nabla_\mu X^\mu = \sqrt{-g}(\partial_\mu X^\mu + \Gamma^\mu{}_{\mu\nu}X^\nu) &= \sqrt{-g}\left(\partial_\mu X^\mu + \frac{g^{\mu\rho}}{2}\partial_\nu g_{\mu\rho}X^\nu\right) \\ &= \sqrt{-g}\partial_\mu X^\mu + \partial_\mu(\sqrt{-g})X^\mu = \partial_\mu(\sqrt{-g}X^\mu)\end{aligned} \tag{A.11}$$

Note that we have used the relation $\sqrt{-g}g^{\mu\rho}\partial_\nu g_{\mu\rho} = 2\partial_\nu(\sqrt{-g})$ which is derived exactly in the same fashion as equation (A.5).

(ix). Finally, we can calculate the variation of the Einstein-Hilbert action based on the results we have already found.

$$\begin{aligned}\delta S_{E-H} &= \frac{1}{2\kappa}\int d^4x\delta(\sqrt{-g}R) \\ &= \frac{1}{2\kappa}\int_\Omega d^4x\sqrt{-g}\delta g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) + \frac{1}{2\kappa}\int_\Omega d^4x\partial_\mu(\sqrt{-g}X^\mu) \\ &= \frac{1}{2\kappa}\int_\Omega d^4x\sqrt{-g}\delta g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) + \frac{1}{2\kappa}\oint_{\partial\Omega} d\Sigma_\mu\sqrt{-g}X^\mu\,;\end{aligned} \tag{A.12}$$

where, in the first equality, we just plugged in the result of equation (A.10).

By the Stokes's theorem, we notice that the last term becomes an integral over the boundary of Ω, namely $\partial\Omega$, and although this boundary term does not automatically vanish, it only contributes as a boundary condition when we require $\delta S = 0$ not to the bulk equations. Therefore, we will only keep the first integral in the expression (A.12) and consider that the boundary conditions are satisfied.

Finally, we can get our results together, equations (A.2), (A.6) and (A.12), to find the equations of motion (Einstein equations).

$$\delta S = \delta S_{E-H} + \delta S_\Lambda + \delta S_m = \int_\Omega d^4x\sqrt{-g}\delta g^{\mu\nu}\left[-\frac{1}{2}T_{\mu\nu} + \frac{1}{2\kappa}\Lambda g_{\mu\nu} + \frac{1}{2\kappa}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)\right] = 0$$
$$\Rightarrow G_{\mu\nu} + \Lambda g_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} = 8\pi G T_{\mu\nu}\,; \tag{A.13}$$

where in equation (A.13), in order to expose the different versions of these equations, we just added the definition of the Einstein tensor $G_{\mu\nu}$ and the value of κ in terms of the Newton constant as discussed in chapter 1.

This result indeed matches that of equation (1.6) as we promised.

Appendix B

Friedmann and Continuity Equations

In order to derive the Friedmann equations we need to calculate the Ricci tensor and the Ricci scalar for the FLRW, equation (1.14). We start by calculating the Christoffel symbols from equation (1.5).

From:

$$g_{\mu\nu} = \mathrm{diag}\left(-1, \frac{a^2}{1-kr^2}, a^2r^2, a^2r^2\sin^2\theta\right)$$
$$\Rightarrow g^{\mu\nu} = \mathrm{diag}\left(-1, \frac{1-kr^2}{a^2}, \frac{1}{a^2r^2}, \frac{1}{a^2r^2\sin^2\theta}\right) \tag{B.1}$$

And:

$$\Gamma^{\rho}{}_{\mu\nu} = \frac{g^{\rho\sigma}}{2}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}) \tag{B.2}$$

We can see that:

(i). Any symbol with at least two time indices vanishes:

$$\Gamma^{\mu}{}_{00} = 0 \text{ and } \Gamma^{0}{}_{0i} = \Gamma^{0}{}_{i0} = 0 \tag{B.3}$$

(ii). A symbol with any set of unrepeated indices also vanishes:

$$\Gamma^{i}{}_{0j} = \Gamma^{i}{}_{j0} = 0\ ;\ \Gamma^{0}{}_{ij} = 0 \text{ and } \Gamma^{i}{}_{jk} = 0\ ;\ \text{for } i \neq i \neq k \text{ and } i \neq k \tag{B.4}$$

(iii). The only non-zero Christoffel symbols with a time index as its first (top) index are:

$$\begin{aligned}
\Gamma^0{}_{11} &= \frac{1}{2}\partial_t\left(\frac{a^2}{1-kr^2}\right) = \frac{a\dot{a}}{1-kr^2} \\
\Gamma^0{}_{22} &= \frac{1}{2}\partial_t(a^2r^2) = a\dot{a}r^2 \\
\Gamma^0{}_{33} &= \frac{1}{2}\partial_t(a^2r^2\sin^2\theta) = a\dot{a}r^2\sin^2\theta
\end{aligned} \tag{B.5}$$

(iv). To complete the Christoffel symbols with a time index we have:

$$\begin{aligned}
\Gamma^1{}_{01} = \Gamma^1{}_{10} &= \frac{1-kr^2}{2a^2}\partial_t\left(\frac{a^2}{1-kr^2}\right) = \frac{\dot{a}}{a} \\
\Gamma^2{}_{02} = \Gamma^2{}_{20} &= \frac{1}{2a^2r^2}\partial_t(a^2r^2) = \frac{\dot{a}}{a} \\
\Gamma^3{}_{03} = \Gamma^3{}_{30} &= \frac{1}{2a^2r^2\sin^2\theta}\partial_t(a^2r^2\sin^2\theta) = \frac{\dot{a}}{a}
\end{aligned} \tag{B.6}$$

(v). The symbols with only one type of spatial index are:

$$\begin{aligned}
\Gamma^1{}_{11} &= \frac{1-kr^2}{2a^2}\partial_r\left(\frac{a^2}{1-kr^2}\right) = \frac{kr}{1-kr^2} \\
\Gamma^2{}_{22} &= \Gamma^3{}_{33} = 0
\end{aligned} \tag{B.7}$$

(vi). Now, notice that any symbol with two radial indices and the other index being either a θ or a ϕ index vanishes:

$$\Gamma^1{}_{1i} = \Gamma^1{}_{i1} = 0 \text{ and } \Gamma^i{}_{11} = 0 \text{ where } i \in \{2,3\} \tag{B.8}$$

(vii). The other Christoffel symbols with a radial index are:

$$\begin{aligned}
\Gamma^2{}_{21} = \Gamma^2{}_{12} &= \frac{1}{2a^2r^2}\partial_r(a^2r^2) = \frac{1}{r} \\
\Gamma^3{}_{31} = \Gamma^3{}_{13} &= \frac{1}{2a^2r^2\sin^2\theta}\partial_r(a^2r^2\sin^2\theta) = \frac{1}{r} \\
\Gamma^1{}_{22} &= -\frac{1-kr^2}{2a^2}\partial_r(a^2r^2) = -r(1-kr^2) \\
\Gamma^1{}_{33} &= -\frac{1-kr^2}{2a^2}\partial_r(a^2r^2\sin^2\theta) = -r\sin^2\theta(1-kr^2)
\end{aligned} \tag{B.9}$$

(viii). Finally, the rest of the Christoffel symbols, involving only θ and ϕ indices, are:

$$\begin{aligned}
&\Gamma^2{}_{23} = \Gamma^2{}_{32} = 0 \text{ and } \Gamma^3{}_{22} = 0 \\
&\Gamma^3{}_{23} = \Gamma^3{}_{32} = \frac{1}{2a^2r^2\sin^2\theta}\partial_\theta(a^2r^2\sin^2\theta) = \frac{\cos\theta}{\sin\theta} \\
&\Gamma^2{}_{33} = -\frac{1}{2a^2r^2}\partial_\theta(a^2r^2\sin^2\theta) = -\sin\theta\cos\theta
\end{aligned} \tag{B.10}$$

We can check that we haven't forgotten any Christoffel symbol (nor counted any of them twice) by realizing we calculated exactly 64 symbols, the precise expected number since μ, ν and ρ can assume any value from 0 to 3 in equation (B.2).

We must calculate now, from the results obtained in equations (B.3, B.4, B.5, B.6, B.7, B.8, B.9, B.10), the Ricci tensor components:

$$R_{\mu\nu} = \partial_\rho\Gamma^\rho{}_{\mu\nu} - \partial_\nu\Gamma^\rho{}_{\mu\rho} + \Gamma^\rho{}_{\sigma\rho}\Gamma^\sigma{}_{\mu\nu} - \Gamma^\rho{}_{\sigma\nu}\Gamma^\sigma{}_{\mu\rho} \tag{B.11}$$

The only non-vanishing components are the diagonal ones. One can explicitly see that other components vanish from the above calculated Christoffel symbols so we will only explicitly calculate the non-vanishing (relevant) ones here:

(i). The time component is:

$$R_{00} = -\partial_t\Gamma^i{}_{0i} - \Gamma^i{}_{j0}\Gamma^j{}_{0i} = -3\partial_t\left(\frac{\dot a}{a}\right) - 3\left(\frac{\dot a}{a}\right)^2 = -3\frac{\ddot a}{a} \tag{B.12}$$

(ii). The radial component is:

$$\begin{aligned}
R_{11} =&\partial_t\left(\frac{a\dot a}{1-kr^2}\right) + \partial_r\left(\frac{kr}{1-kr^2}\right) - \partial_r\left(\frac{kr}{1-kr^2}\right) - 2\partial_r\left(\frac{1}{r}\right) + \Gamma^i{}_{0i}\Gamma^0{}_{11} \\
&+ \Gamma^i{}_{ji}\Gamma^j{}_{11} - \Gamma^0{}_{i1}\Gamma^i{}_{10} - \Gamma^i{}_{01}\Gamma^0{}_{1i} - \Gamma^i{}_{j1}\Gamma^j{}_{1i} \\
=&\frac{\dot a}{1-k^2} + \frac{a\ddot a}{1-kr^2} + \frac{2}{r^2} + 3\frac{\dot a^2}{1-kr^2} + \frac{k^2r^2}{(1-kr^2)^2} + \frac{2k}{1-kr^2} \\
&- 2\frac{\dot a^2}{1-kr^2} - \frac{k^2r^2}{(1-kr^2)^2} - \frac{2}{r^2} \\
=&\frac{a\ddot a + 2k + 2\dot a^2}{1-kr^2}
\end{aligned} \tag{B.13}$$

(iii). The θ component is:

$$
\begin{aligned}
R_{22} =& \partial_t(a\dot{a}r^2) + \partial_r[-r(1-kr^2)] - \partial_\theta\left(\frac{\cos\theta}{\sin\theta}\right) + \Gamma^i{}_{0i}\Gamma^0{}_{22} + \Gamma^i{}_{1i}\Gamma^1{}_{22} - \Gamma^0{}_{22}\Gamma^2{}_{20} \\
& - \Gamma^1{}_{22}\Gamma^2{}_{21} - \Gamma^2{}_{02}\Gamma^0{}_{22} - \Gamma^2{}_{12}\Gamma^1{}_{22} - \Gamma^3{}_{32}\Gamma^3{}_{23} \\
=& \dot{a}^2r^2 + a\ddot{a}r^2 - (1-kr^2) + 2kr^2 + 1 + \frac{\cos^2\theta}{\sin^2\theta} + 2\dot{a}^2r^2 - kr^2 - 2(1-kr^2) \\
& - \dot{a}^2r^2 + 2(1-kr^2) - \frac{\cos^2\theta}{\sin^2\theta} \\
=& a\ddot{a}r^2 + 2kr^2 + 2\dot{a}^2r^2
\end{aligned}
\tag{B.14}
$$

(iv). The ϕ component is:

$$
\begin{aligned}
R_{33} =& \partial_t(a\dot{a}r^2\sin^2\theta) - \partial_r[r\sin^2\theta(1-kr^2)] - \partial_\theta(\sin\theta\cos\theta) + \Gamma^i{}_{0i}\Gamma^0{}_{33} + \Gamma^i{}_{1i}\Gamma^1{}_{33} \\
& + \Gamma^i{}_{2i}\Gamma^2{}_{33} - \Gamma^0{}_{33}\Gamma^3{}_{30} - \Gamma^3{}_{03}\Gamma^0{}_{33} - \Gamma^1{}_{33}\Gamma^3{}_{31} - \Gamma^2{}_{33}\Gamma^3{}_{32} - \Gamma^3{}_{13}\Gamma^1{}_{33} - \Gamma^3{}_{23}\Gamma^2{}_{33} \\
=& \dot{a}^2r^2\sin^2\theta + a\ddot{a}r^2\sin^2\theta - \sin^2\theta + kr^2\sin^2\theta + 2kr^2\sin^2\theta - \cos^2\theta + \sin^2\theta \\
& + 3\dot{a}^2r^2\sin^2\theta - kr^2\sin^2\theta - 2\sin^2\theta + 2kr^2\sin^2\theta - \cos^2\theta - 2\dot{a}^2r^2\sin^2\theta \\
& + 2\sin^2\theta - 2kr^2\sin^2\theta + 2\cos^2\theta \\
=& (a\ddot{a}r^2 + 2kr^2 + 2\dot{a}^2r^2)\sin^2\theta
\end{aligned}
\tag{B.15}
$$

We are now able to calculate the Ricci scalar from the Ricci tensor components calculated above, equations (B.12, B.13, B.14, B.15).

$$
\begin{aligned}
R \equiv g^{\mu\nu}R_{\mu\nu} =& 3\frac{\ddot{a}}{a} + \left(\frac{1-kr^2}{a^2}\right)\left(\frac{a\ddot{a} + 2k + 2\dot{a}^2}{1-kr^2}\right) + \frac{2}{a^2r^2}(a\ddot{a}r^2 + 2\dot{a}^2r^2 + 2kr^2) \\
=& \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k)
\end{aligned}
\tag{B.16}
$$

Note that we can define $T^{(\Lambda)}_{\mu\nu} \equiv -(\Lambda/8\pi G)g_{\mu\nu}$ such that we hide the cosmological constant contribution in the Einstein equations, equation (1.6), as a type of energy content through $T_{\mu\nu} \to \tilde{T}_{\mu\nu} \equiv T_{\mu\nu} + T^{(\Lambda)}_{\mu\nu}$. Therefore, we will consider only $G_{\mu\nu} = 8\pi G T_{\mu\nu}$ as the Einstein equations and if we want to consider a cosmological constant term, its contribution will be hidden in the energy-momentum tensor. To extract the differential equations for $a(t)$ from the Einstein equations we note from equation (1.20) that if we raise one index of the stress-energy tensor and work in the CMBR frame, i.e. the frame in which the 4-velocity of the energy content fluid is $u^\mu = (1, 0, 0, 0)$ and thus $u_\mu \equiv g_{\mu\nu}u^\nu = (-1, 0, 0, 0)$, we can write a simpler version of $T_{\mu\nu}$:

$$T^\mu{}_\nu \equiv g^{\mu\rho}T_{\rho\nu} = \mathrm{diag}(-\rho, P, P, P) \tag{B.17}$$

In order to compare $G_{\mu\nu}$ with this simplified version of the stress-energy tensor, we have to raise one index of $R_{\mu\nu}$ also.

$$\begin{aligned} R^0{}_0 &\equiv g^{0\mu}R_{\mu 0} = g^{00}R_{00} = 3\frac{\ddot{a}}{a} \\ R^1{}_1 &= R^2{}_2 = R^3{}_3 = \frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2}\,; \end{aligned} \tag{B.18}$$

which leads to:

$$\begin{aligned} G^0{}_0 &\equiv g^{0\mu}G_{\mu 0} = g^{00}G_{00} = R^0{}_0 - \frac{1}{2}R = -\frac{3}{a^2}(\dot{a}^2 + k) \\ G^1{}_1 &= G^2{}_2 = G^3{}_3 = -2\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 - \frac{k}{a^2} \end{aligned} \tag{B.19}$$

Since all the spatial components, either radial, θ or ϕ, give us the same equation, we end up with two distinct differential equations, one from the time component of $G^\mu{}_\nu$ and one from its spatial components:

$$\frac{3}{a^2}(\dot{a}^2 + k) = 8\pi G\rho \therefore H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}\,; \tag{B.20}$$

and:

$$-2\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 - \frac{k}{a^2} = 8\pi GP \therefore \frac{\ddot{a}}{a} = -4\pi GP - \frac{1}{2}H^2 - \frac{k}{2a^2} \tag{B.21}$$

We can get rid of H^2 in equation (B.21) using equation (B.20), obtaining:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{B.22}$$

Equation (B.20) is the so-called first Friedmann equation and equation (B.22) is the second Friedmann equation.

Before we derive the continuity equation, we will first make explicit the cosmological constant terms in the Friedmann equations (recall we have hidden them in the energy-momentum tensor). First, notice that the would have came to the same equations if we considered more than one type of energy content. The only difference would be that $\rho \to \sum_i \rho_i$ in the first Friedmann equation and $\rho + 3P \to \sum_i(\rho_i + 3P_i)$ in the second, where i represents the different types of energy content in the universe. Second, note that from $T^{(\Lambda)}_{\mu\nu} \equiv -(\Lambda/8\pi G)g_{\mu\nu}$ we get $T^{(\Lambda)\mu}{}_{\nu} = -(\Lambda/8\pi G)\delta^{\mu}{}_{\nu} = \text{diag}(-\rho_\Lambda, P_\Lambda, P_\Lambda, P_\Lambda) \Rightarrow P_\Lambda = -\rho_\Lambda = -(\Lambda/8\pi G)$. Thus, the Friedmann equations become:

$$H^2 = \frac{8\pi G}{3}(\rho + \rho_\Lambda) - \frac{k}{a^2} = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}\,; \tag{B.23}$$

and:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + \rho_\Lambda + 3P + 3P_\Lambda) = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3} \tag{B.24}$$

Finally, we may derive the continuity equation. From GR, we know that the stress-energy tensor is the current associated with spacetime translations. In order to have a symmetry under general coordinate transformations we must have:

$$\begin{aligned}\nabla_\mu T^{\mu}{}_{\nu} &= \partial_\mu T^{\mu}{}_{\nu} + \Gamma^{\mu}{}_{\mu\rho}T^{\rho}{}_{\nu} - \Gamma^{\rho}{}_{\mu\nu}T^{\mu}{}_{\rho} \\ &= \partial_t T^{0}{}_{\nu} + \Gamma^{i}{}_{i0}T^{0}{}_{\nu} + \Gamma^{i}{}_{ij}T^{j}{}_{\nu} - \Gamma^{i}{}_{j\nu}T^{j}{}_{i} = 0\end{aligned} \tag{B.25}$$

We can check from equation (B.25) that the equations obtained for $\nu = 1, 2$ or 3 are identically equal to zero. On the other hand, the time component if it gives us the desired continuity equation:

$$\nabla_\mu T^\mu{}_0 = -\partial_t \rho - 3\frac{\dot{a}}{a}\rho - 3\frac{\dot{a}}{a}P = 0 \therefore \dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P) \tag{B.26}$$

Notice that $\nabla_\mu T^\mu{}_\nu = 0$ is a conservation of energy (and momentum) statement in GR. We consider that each type of matter content in the universe has its stress-energy tensor conserved individually, therefore equation (1.23) is valid for each type of fluid.

Appendix C

Thermodynamic Computations

In order to derive expressions for the number density, energy density and pressure of a gas following a given distribution $f(E)$ we must use the usual trick of imagining the system in a box before taking the continuum limit of it. Recall that $f(E)$ is the average number of particles in a certain energy level E. For a gas in a box, the total number of particles N is found by summing up $f(E_n)$ for all energy levels n taking into consideration internal degrees of freedom of the species g, i.e. its degeneracy.

$$N = g \sum_n f(E_n) = g \sum_n f(E_n) \frac{\Delta p_x \Delta p_y \Delta p_z}{\Delta p_x \Delta p_y \Delta p_z} \, ; \tag{C.1}$$

where in the second equality we just multiplied and divided by the momentum space minimum volume since we want to get to a momentum space integral expression later on when we take the continuum limit which will be rewritten as an energy integral.

If we assume periodic boundary conditions, as it is often done in statistical mechanics calculations, we will have the relation $\Delta k_i = \Delta p_i = \frac{2\pi}{L_i} \Delta \tilde{n}_i$ where k_i is the variable conjugated to position, via a Fourier transform, in the ith direction ($k_i = p_i$ since $\hbar = 1$), L_i is the size of the box in the same direction and $\tilde{n}_i$ is the quantum number dictating the energy of the state - there are three of them, one for each direction. From its definition, the minimum value for $\Delta \tilde{n}_i$ possible is 1, since $\tilde{n}_i \in \mathbb{Z}$, therefore the minimum value possible for Δp_i is $\frac{2\pi}{L_i}$.

Substituting this result in the denominator of equation (C.1) we get:

$$N = g\sum_n f(E_n)\frac{L_xL_yL_z}{(2\pi)^3}\Delta p_x\Delta p_y\Delta p_z \Rightarrow n \equiv \frac{N}{L_xL_yL_z} = \frac{N}{V} = \frac{g}{(2\pi)^3}\int d^3p f(E(p))\,; \quad \text{(C.2)}$$

where after the implication arrow we took the continuum limit. It was also explicitly written $E(p)$ instead of p because if we wanted to integrate the above expression as it is we would need to know the dispersion relation, e.g. for a gas of relativistic non-interacting particles the old-fashioned $E(p) = \sqrt{|\vec{p}|^2 + m^2}$. However, we will shortly change variables to convert the integral over momenta into an integral over energy.

To calculate the energy density of the gas the only difference to the above procedure would be the starting point. $f(E_n)$ in equation (C.1) works as a weight factor of an average value calculation. To compute the total number of particles in the box we just needed to sum it, but to calculate the total energy E we need to sum the energy of each level E_n multiplied by this weight factor, thus:

$$E = g\sum_n E_n f(E_n) \quad \text{(C.3)}$$

After following the same steps as in the last derivation we arrive at:

$$\rho \equiv \frac{E}{V} = \frac{g}{(2\pi)^3}\int d^3p E(p) f(E(p)) \quad \text{(C.4)}$$

The same principle of derivation will be used to find an integral expression for the pressure. However, in this case we need to figure out what should appear multiplying $f(E_n)$ inside the sum. Remember from the kinetic theory of gases that in order to calculate the pressure of a gas we must track down the behaviour of one of its components to compute the rate of momentum transfer to a wall of the box. Considering a non-interacting gas whose particles collide elastically against a wall perpendicular to the x-direction of the box, we know that a momentum $\delta p_x = 2p_x$, where p_x is the momentum of the particle in the x-direction, is transferred per hit. We also know that it takes a time $\delta t = 2L_x/v_x = 2mL_x/p_x$, where L_x is the size of the box in the x-direction, v_x is the component of the particle velocity in the same direction and m is the particle mass, between two consecutive hits against the same wall. Therefore:

$$\frac{\delta p_x}{\delta t} = \frac{p_x^2}{mL_x} \tag{C.5}$$

We realize that $\delta p_x/\delta t$ in equation (C.5) is a force and to get the pressure out of that we must divide by the area of the wall, which is in the yz-plane by construction, thus $A = L_y L_z$. Since our gas is isotropic, on average $p_x^2 = p_y^2 = p_z^2 = p^2/3 \equiv |\vec{p}|^2/3$. Therefore, the pressure generated by the collisions of each particle is $P = \frac{p^2}{3mV}$. Our analysis was purely non-relativistic up until now, so in order to consider the relativistic case we just need to replace $m \to E$ which leads to:

$$P = g\sum_n \frac{p_n^2}{3VE_n} f(E_n) \Rightarrow P = \frac{g}{(2\pi)^3}\int d^3p \frac{p^2}{3E(p)} f(p)\,; \tag{C.6}$$

where all the steps from the sum to the integral expression are the same as in the case of the number density, therefore we only presented the final result.

Again from isotropy, all the integrals in equations (C.2), (C.4) and (C.6) only depend on the modulus of the momentum, thus working in spherical coordinates we can integrate out the angles in momentum space as usual $\int d^3p\tilde{f}(p) = 4\pi\int dpp^2\tilde{f}(p)$, where $\tilde{f}(p)$ is a generic function of the modulus of $\vec{p}$. Furthermore, still considering a non-interacting relativistic gas, the dispersion relation is $E^2 = p^2 + m^2$, as anticipated, leading to $pdp = EdE$ and $E \to \infty$ as $p \to \infty$ while $E \to m$ as $p \to 0$. Finally, equations (C.2), (C.4) and (C.6) are rewritten in terms of integrals over energy as:

$$n = \frac{g}{2\pi^2}\int_m^\infty dE\sqrt{E^2 - m^2}Ef(E) \tag{C.7}$$

$$\rho = \frac{g}{2\pi^2}\int_m^\infty dE\sqrt{E^2 - m^2}E^2 f(E) \tag{C.8}$$

$$P = \frac{g}{2\pi^2}\int_m^\infty dE\frac{E^2 - m^2}{3E}E\sqrt{E^2 - m^2}f(E) = \frac{g}{6\pi^2}\int_m^\infty dE(E^2 - m^2)^{3/2}f(E) \tag{C.9}$$

Those are the expressions we were looking for.

Appendix D

Kaluza-Klein Original Theory

Our goal in this appendix is to derive the 4-dimensional action coming from the original 5-dimensional Kaluza-Klein theory. One starts with only one field in 5-dimensions, the metric $\hat{g}_{MN}$, which is going to parametrized as in equation (3.16):

$$\hat{g}_{MN}(x^\mu) = \phi^{-1/3} \begin{pmatrix} g_{\mu\nu} + \kappa^2 \phi A_\mu A_\nu & \kappa\phi A_\mu \\ \kappa\phi A_\nu & \phi \end{pmatrix}; \tag{D.1}$$

where $\kappa \equiv 4\sqrt{\pi G}$.

Recall that all fields depend only on x^μ and therefore $\hat{g}_{MN}$ is y-independent. The Einstein-Hilbert action in 5D is:

$$\hat{S}_{E-H} = \frac{1}{16\pi\hat{G}} \int d^5x \sqrt{-\hat{g}}\hat{R}; \tag{D.2}$$

where $\hat{R}$ is the 5-dimensional Ricci scalar.

Calculating the Ricci scalar in full detail for the metric in equation (D.1) is too long. Therefore, we will first compute the 4D action considering $A_\mu = 0$, which is a consistent reduction as it will be seen at the end, as a warm up and then show how to generalize to the full non-linear metric case. The warm up calculation gives us the same insights needed to the more complete case and since it is way shorter, although still a bit long, it is economical and not less insightful to present things this way. Thus, the metric and its inverse are:

$$\hat{g}_{MN} = \text{diag}(\phi^{-1/3} g_{\mu\nu}, \phi^{2/3}) \text{ ; and } \hat{g}^{MN} = \text{diag}(\phi^{1/3} g^{\mu\nu}, \phi^{-2/3}) \tag{D.3}$$

Note that, from the above expression for the metric, one gets:

$$\det(\hat{g}_{MN}) \equiv \hat{g} = \phi^{2/3}\phi^{-4/3}\det(g_{\mu\nu}) = \phi^{-2/3} g \Rightarrow \sqrt{-\hat{g}} = \phi^{-1/3}\sqrt{-g} \tag{D.4}$$

Furthermore, with the results of equation (D.3) one must calculate the Christoffel symbols, $\hat{\Gamma}$s, in terms of the 4D fields $g_{\mu\nu}$ and ϕ:

$$\hat{\Gamma}^{A}{}_{BC} = \frac{\hat{g}^{AD}}{2}\left(\partial_C \hat{g}_{DB} + \partial_B \hat{g}_{DC} - \partial_D \hat{g}_{BC}\right) \tag{D.5}$$

(i). First, since the metric and its inverse are block diagonal and no metric component depends on y, i.e. $\partial_4 \hat{g}_{MN} = 0$, note that:

$$\hat{\Gamma}^{4}{}_{44} = 0 \tag{D.6}$$

(ii). For $\hat{\Gamma}$s with one μ (running from 0 to 3, as usual) index downstairs:

$$\hat{\Gamma}^{4}{}_{4\mu} = \hat{\Gamma}^{4}{}_{\mu 4} = \frac{\phi^{-2/3}}{2}\left[\partial_\mu(\phi^{2/3})\right] = \frac{1}{3\phi}(\partial_\mu \phi) \tag{D.7}$$

(iii). For one upstairs index μ now:

$$\hat{\Gamma}^{\mu}{}_{44} = \frac{\phi^{1/3} g^{\mu\nu}}{2}\left[-\partial_\nu(\phi^{2/3})\right] = -\frac{1}{3} g^{\mu\nu}\partial_\nu \phi \tag{D.8}$$

(iv). For two indices of the type μ downstairs:

$$\hat{\Gamma}^{4}{}_{\mu\nu} = 0 \tag{D.9}$$

(v). Also, we get the same result for one large dimension index μ up and another one down:

$$\hat{\Gamma}^{\mu}{}_{\nu 4} = \hat{\Gamma}^{\mu}{}_{4\nu} = 0 \tag{D.10}$$

(vi). Finally, for all indices running from 0 to 3:

$$\begin{aligned}\hat{\Gamma}^{\mu}{}_{\nu\rho} &= \frac{\phi^{1/3} g^{\mu\sigma}}{2}\left[\partial_\nu(\phi^{-1/3} g_{\sigma\rho}) + \partial_\rho(\phi^{-1/3} g_{\sigma\nu}) - \partial_\sigma(\phi^{-1/3} g_{\nu\rho})\right] \\ &= \Gamma^{\mu}{}_{\nu\rho} - \frac{1}{6\phi}\left(\delta^{\mu}{}_{\rho}\partial_\nu\phi + \delta^{\mu}{}_{\nu}\partial_\rho\phi - g_{\nu\rho} g^{\mu\sigma}\partial_\sigma\phi\right)\end{aligned} \tag{D.11}$$

Now, one has to calculate the Ricci scalar:

$$\begin{aligned}\hat{R} =& \hat{g}^{AB}\left(\partial_C \hat{\Gamma}^{C}{}_{AB} - \partial_B \hat{\Gamma}^{C}{}_{AC} + \hat{\Gamma}^{D}{}_{AB}\hat{\Gamma}^{C}{}_{CD} - \hat{\Gamma}^{D}{}_{AC}\hat{\Gamma}^{C}{}_{BD}\right) \\ =& \phi^{-2/3}\left(-\frac{1}{3}\partial_\mu g^{\mu\nu}\partial_\nu\phi - \frac{1}{3}\Box\phi + \hat{\Gamma}^{\mu}{}_{44}\hat{\Gamma}^{\nu}{}_{\nu\mu} + \hat{\Gamma}^{\mu}{}_{44}\hat{\Gamma}^{4}{}_{4\mu} - \hat{\Gamma}^{4}{}_{4\mu}\hat{\Gamma}^{\mu}{}_{44} - \hat{\Gamma}^{\mu}{}_{44}\hat{\Gamma}^{4}{}_{4\mu}\right) + \\ &+ \phi^{1/3} g^{\mu\nu}\left(\partial_\rho \hat{\Gamma}^{\rho}{}_{\mu\nu} - \partial_\nu \hat{\Gamma}^{\rho}{}_{\mu\rho} - \partial_\nu \hat{\Gamma}^{4}{}_{\mu 4} + \hat{\Gamma}^{\rho}{}_{\mu\nu}\hat{\Gamma}^{4}{}_{4\rho} + \hat{\Gamma}^{\rho}{}_{\mu\nu}\hat{\Gamma}^{\sigma}{}_{\sigma\rho} - \hat{\Gamma}^{4}{}_{\mu 4}\hat{\Gamma}^{4}{}_{\nu 4} - \hat{\Gamma}^{\rho}{}_{\mu\sigma}\hat{\Gamma}^{\sigma}{}_{\nu\rho}\right) \\ =& \phi^{1/3} R - \frac{1}{3}\phi^{-2/3}\partial_\mu g^{\mu\nu}\partial_\nu\phi - \frac{1}{3}\phi^{-2/3}\Box\phi + \phi^{-2/3}\Bigg\{\left(-\frac{1}{3} g^{\mu\rho}\partial_\rho\phi\right)\Bigg[\Gamma^{\nu}{}_{\nu\mu} - \frac{1}{6\phi}(\delta^{\nu}{}_{\mu}\partial_\nu\phi + \\ &+ \delta^{\nu}{}_{\nu}\partial_\mu\phi - g_{\mu\nu} g^{\nu\sigma}\partial_\sigma\phi)\Bigg] - \left(\frac{1}{3\phi}\partial_\mu\phi\right)\left(-\frac{1}{3} g^{\mu\nu}\partial_\nu\phi\right)\Bigg\} + \\ &+ \phi^{1/3} g^{\mu\nu}\Bigg\{\partial_\rho\left[-\frac{1}{6\phi}(\delta^{\rho}{}_{\nu}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\nu\phi - g_{\mu\nu} g^{\rho\sigma}\partial_\sigma\phi)\right] - \partial_\nu\Bigg[-\frac{1}{6\phi}(\delta^{\rho}{}_{\rho}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\rho\phi - \\ &- g_{\mu\rho} g^{\rho\sigma}\partial_\sigma\phi)\Bigg] - \partial_\nu\left(\frac{1}{3\phi}\partial_\mu\phi\right) + \left[\Gamma^{\rho}{}_{\mu\nu} - \frac{1}{6\phi}(\delta^{\rho}{}_{\nu}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\nu\phi - g_{\mu\nu} g^{\rho\sigma}\partial_\sigma\phi)\right] \times \\ &\times \left(\frac{1}{3\phi}\partial_\rho\phi\right) + \Gamma^{\rho}{}_{\mu\nu}\left[-\frac{1}{6\phi}(\delta^{\sigma}{}_{\rho}\partial_\sigma\phi + \delta^{\sigma}{}_{\sigma}\partial_\rho\phi - g_{\sigma\rho} g^{\sigma\lambda}\partial_\lambda\phi)\right] + \Bigg[-\frac{1}{6\phi}(\delta^{\rho}{}_{\nu}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\nu\phi - \\ &- g_{\mu\nu} g^{\rho\lambda}\partial_\lambda\phi)\Bigg]\Gamma^{\sigma}{}_{\sigma\rho} + \left[-\frac{1}{6\phi}(\delta^{\rho}{}_{\nu}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\nu\phi - g_{\mu\nu} g^{\rho\lambda}\partial_\lambda\phi)\right]\left[-\frac{1}{6\phi}(4\partial_\rho\phi)\right] - \\ &- \left(\frac{1}{3\phi}\partial_\mu\phi\right)\left(\frac{1}{3\phi}\partial_\nu\phi\right) - \Gamma^{\rho}{}_{\mu\sigma}\left[-\frac{1}{6\phi}(\delta^{\sigma}{}_{\rho}\partial_\nu\phi + \delta^{\sigma}{}_{\nu}\partial_\rho\phi - g_{\nu\rho} g^{\sigma\lambda}\partial_\lambda\phi)\right] - \\ &- \left[-\frac{1}{6\phi}(\delta^{\rho}{}_{\sigma}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\sigma\phi - g_{\mu\sigma} g^{\rho\lambda}\partial_\lambda\phi)\right]\Gamma^{\sigma}{}_{\nu\rho} - \Bigg[-\frac{1}{6\phi}(\delta^{\rho}{}_{\sigma}\partial_\mu\phi + \delta^{\rho}{}_{\mu}\partial_\sigma\phi -\end{aligned}$$

$$
\begin{aligned}
& - g_{\mu\sigma} g^{\rho\lambda} \partial_\lambda \phi) \Big] \Big[-\frac{1}{6\phi} (\delta^\sigma{}_\rho \partial_\nu \phi + \delta^\sigma{}_\nu \partial_\rho \phi - g_{\nu\rho} g^{\sigma\lambda} \partial_\lambda \phi) \Big] \Big\} \\
= & \phi^{1/3} R - \frac{1}{3}\phi^{-2/3} \partial_\mu g^{\mu\nu} \partial_\nu \phi - \frac{1}{3}\phi^{-2/3} \Box \phi - \frac{1}{3}\phi^{-2/3} g^{\mu\rho} \partial_\rho \phi \Gamma^\nu{}_{\nu\mu} + \frac{1}{3}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi - \\
& - \frac{1}{3}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi + \frac{1}{3}\phi^{-2/3} \Box \phi + \frac{1}{6}\phi^{-2/3} g^{\mu\nu} \partial_\rho g_{\mu\nu} g^{\rho\sigma} \partial_\sigma \phi + \frac{2}{3}\phi^{-2/3} \partial_\mu g^{\mu\nu} \partial_\nu \phi - \\
& - \frac{1}{3}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi + \frac{1}{3}\phi^{-2/3} \Box \phi + \frac{1}{3}\phi^{-2/3} g^{\mu\nu} \Gamma^\rho{}_{\mu\nu} \partial_\rho \phi + \frac{1}{9}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi - \\
& - \frac{2}{3}\phi^{-2/3} g^{\mu\nu} \Gamma^\rho{}_{\mu\nu} \partial_\rho \phi + \frac{1}{3}\phi^{-2/3} g^{\mu\nu} \Gamma^\rho{}_{\rho\mu} \partial_\nu \phi - \frac{2}{9}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi - \frac{1}{9}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi + \\
& + \frac{2}{6}\phi^{-2/3} g^{\mu\nu} \Gamma^\rho{}_{\mu\nu} \partial_\rho \phi + \frac{1}{18}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi \\
= & \phi^{1/3} R - \frac{1}{2}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi + \frac{1}{3}\phi^{-2/3} \Box \phi + \frac{1}{3}\phi^{-2/3} \partial_\mu g^{\mu\nu} \partial_\nu \phi + \frac{1}{6}\phi^{-2/3} g^{\mu\nu} \partial_\rho g_{\mu\nu} g^{\rho\sigma} \partial_\sigma \phi \\
= & \phi^{1/3} R - \frac{1}{2}\phi^{-5/3} \partial_\mu \phi \partial^\mu \phi + \frac{1}{3}\phi^{-2/3} \left(\Box \phi + \partial_\mu g^{\mu\nu} \partial_\nu \phi + \frac{1}{2} g^{\mu\nu} \partial_\rho g_{\mu\nu} g^{\rho\sigma} \partial_\sigma \phi \right) ; \qquad \text{(D.12)}
\end{aligned}
$$

where $\Box\phi \equiv g^{\mu\nu}\partial_\mu\partial_\nu\phi$, as usual.

Realizing that $\Gamma^\mu{}_{\mu\rho} = (1/2)g^{\mu\nu}(\partial_\mu g_{\nu\rho} + \partial_\rho g_{\nu\mu} - \partial_\nu g_{\mu\rho}) = (1/2)g^{\mu\nu}\partial_\rho g_{\mu\nu}$, one can rewrite the expression inside the brackets in the last line of equation (D.12) as $\nabla_\mu(g^{\mu\nu}\partial_\nu\phi)$.

$$
\nabla_\mu(g^{\mu\nu}\partial_\nu\phi) = \partial_\mu(g^{\mu\nu}\partial_\nu\phi) + \Gamma^\mu{}_{\mu\rho} g^{\rho\nu}\partial_\nu\phi = \partial_\mu g^{\mu\nu}\partial_\nu\phi + \Box\phi + \frac{1}{2} g^{\mu\nu}\partial_\rho g_{\mu\nu} g^{\rho\sigma}\partial_\sigma\phi \qquad \text{(D.13)}
$$

Thus:

$$
\phi^{-1/3}\hat{R} = R - \frac{1}{2}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} + \frac{1}{3\phi}\nabla_\mu(g^{\mu\nu}\partial_\nu\phi) \qquad \text{(D.14)}
$$

It is desirable to rewrite the last term of equation (D.14) as a total covariant derivative in order to transform this piece of the action into a boundary term which is assumed to be kept fixed as one varies the action and therefore does not contribute to the equations of motion. Thus, note that:

$$
\begin{aligned}
\nabla_\mu X^\mu \equiv \nabla_\mu \left(\frac{g^{\mu\nu}\partial_\nu\phi}{\phi} \right) &= \frac{1}{\phi}\nabla_\mu(g^{\mu\nu}\partial_\nu\phi) - \left(\frac{\partial_\mu\phi}{\phi^2} \right) g^{\mu\nu}\partial_\nu\phi \\
\Rightarrow \frac{1}{\phi}\nabla_\mu(g^{\mu\nu}\partial_\nu\phi) &= \nabla_\mu X^\mu + \frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} \qquad \text{(D.15)}
\end{aligned}
$$

Plugging the result of equation (D.15) into equation (D.14) one finds out:

$$\phi^{-1/3}\hat{R} = R - \frac{1}{6}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} + \frac{1}{3}\nabla_\mu X^\mu \tag{D.16}$$

Finally, we must substitute the results obtained in equations (D.4) and (D.16) in equation (D.2) to find the dimensionally reduced action:

$$\begin{aligned}\hat{S}_{E-H} &= \left(\frac{1}{16\pi\hat{G}}\int dy\right)\int d^4x\sqrt{-g}\phi^{-1/3}\hat{R} \\ &= \frac{1}{16\pi G}\left[\int d^4x\sqrt{-g}\left(R - \frac{1}{6}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2}\right) + \frac{1}{3}\int_\Omega d^4x\sqrt{-g}\nabla_\mu X^\mu\right]\end{aligned} \tag{D.17}$$

As anticipated, using the trick of equation (A.11), $\sqrt{-g}\nabla_\mu X^\mu = \partial_\mu(\sqrt{-g}X^\mu)$, and applying the Stokes's theorem one transforms the last integral in the above equation into a boundary integral which doesn't contribute to the equations of motion. One could bring the action in equation (D.17) to an even more familiar form by redefining the scalar field as $\phi \to \varphi \equiv -\frac{\log\phi}{4\sqrt{3\pi G}}$ leading to a canonical kinetic term for the scalar field in the action, $\frac{1}{6(16\pi G)}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi$.

Now, we are going to generalize the above calculation for the case where $A_\mu \neq 0$ in equation (D.1). The first thing to note is that it is not that straight forward to invert the metric as it was in the previous block-diagonal case. So, let us guess the generic form of the inverse metric as:

$$\hat{g}^{MN} = \phi^{1/3}\begin{pmatrix} g^{\mu\nu} + \delta g^{\mu\nu} & B^\mu \\ B^\nu & \varphi \end{pmatrix}; \tag{D.18}$$

and try to write $\delta g^{\mu\nu}$, B^μ and φ in terms of the metric (or its inverse), A_μ, ϕ and κ by imposing $\hat{g}_{MN}\hat{g}^{NP} = \delta_M{}^P$.

$$\hat{g}_{MN}\hat{g}^{NP} = \begin{pmatrix} g_{\mu\nu} + \kappa^2\phi A_\mu A_\nu & \kappa\phi A_\mu \\ \kappa\phi A_\nu & \phi \end{pmatrix} \begin{pmatrix} g^{\nu\rho} + \delta g^{\nu\rho} & B^\nu \\ B^\rho & \varphi \end{pmatrix}$$

$$\Rightarrow \begin{cases} \hat{g}_{\mu N}\hat{g}^{N\rho} = & \delta_\mu{}^\rho + g_{\mu\nu}\delta g^{\nu\rho} + \kappa^2\phi A_\mu A_\nu g^{\nu\rho} + \kappa^2\phi A_\mu A_\nu \delta g^{\nu\rho} + \kappa\phi A_\mu B^\rho = \delta_\mu{}^\rho \\ \hat{g}_{\mu N}\hat{g}^{N4} = & g_{\mu\nu}B^\nu + \kappa^2\phi A_\mu A_\nu B^\nu + \kappa\phi\varphi A_\mu = 0 \\ \hat{g}_{4N}\hat{g}^{N\rho} = & \kappa\phi A_\nu g^{\nu\rho} + \kappa\phi A_\nu \delta g^{\nu\rho} + \phi B^\rho = 0 \\ \hat{g}_{4N}\hat{g}^{N4} = & \kappa\phi A_\nu B^\nu + \phi\varphi = 1 \end{cases} \tag{D.19}$$

Let us now suppose $\delta g^{\nu\rho} = 0$. Of course, this is an ansatz and one needs to make sure this is valid at the end. It turns out this ansatz indeed allows us to find the inverse metric. An example of a bad ansatz would be $\varphi = \phi^{-1}$ since this would automatically require $A_\mu A_\nu g^{\mu\nu} = 0$, which is not the case of our original guess. Then, from the second and last equations of (D.19):

$$B^\mu = -\kappa A_\nu g^{\nu\mu}\,; \tag{D.20}$$

$$\varphi = \phi^{-1} - \kappa A_\nu B^\nu = \phi^{-1} + \kappa^2 A_\mu A_\nu g^{\mu\nu} \tag{D.21}$$

The first and third equations of (D.19) are supposed to be respected too and they indeed are by the expressions for B^μ and φ above. Thus, we end up with:

$$\hat{g}^{MN} = \phi^{1/3} \begin{pmatrix} g^{\mu\nu} & -\kappa A_\rho g^{\rho\mu} \\ -\kappa A_\rho g^{\rho\nu} & \phi^{-1} + \kappa^2 A_\mu A_\nu g^{\mu\nu} \end{pmatrix} \tag{D.22}$$

From equation (D.1), if all the fields in the 4-dimensional theory have the same average amplitude over space and time, since $\kappa \equiv 4\sqrt{\pi G} = \sqrt{2}M_{Pl}^{-1}$, one can approximate:

$$\det(\hat{g}_{MN}) \equiv \hat{g} = \phi^{-5/3} \det \left[\begin{pmatrix} g_{\mu\nu} & 0 \\ 0 & \phi \end{pmatrix} + \kappa \begin{pmatrix} \kappa\phi A_\mu A_\nu & \phi A_\mu \\ \phi A_\nu & 0 \end{pmatrix} \right]$$
$$= \phi^{-5/3} \det \begin{pmatrix} g_{\mu\nu} & 0 \\ 0 & \phi \end{pmatrix} \det \left(I + \mathcal{O}(M_{Pl}^{-1}) \right) \simeq \phi^{-2/3} \det(g_{\mu\nu}) \, ; \qquad \text{(D.23)}$$

where I is the identity matrix in five dimensions. The above equation leads to:

$$\sqrt{-\hat{g}} \simeq \phi^{-1/3} \sqrt{-g} \qquad \text{(D.24)}$$

Now, we are supposed to calculate the Christoffel symbols and then find an expression for the Ricci scalar in the same manner it was done in our warm up example. Since the metric isn't block-diagonal anymore, the $\hat{\Gamma}$ expressions get messier. The connection coefficient functions are:

(i). First, the $\hat{\Gamma}$ whose indices are all that of the compactified direction:

$$\hat{\Gamma}^4{}_{44} = \frac{\kappa A_\rho g^{\rho\mu}}{2} \phi^{1/3} \left[\partial_\mu \left(\phi^{2/3} \right) \right] = \frac{\kappa}{3} A_\rho g^{\rho\mu} \partial_\mu \phi \qquad \text{(D.25)}$$

(ii). For one μ index down:

$$\hat{\Gamma}^4{}_{\mu 4} = \hat{\Gamma}^4{}_{4\mu} = \frac{\phi^{1/3}}{2} \left(\phi^{-1} + \kappa^2 A_\rho A_\nu g^{\rho\nu} \right) \left[\partial_\mu \left(\phi^{2/3} \right) \right] - \frac{\phi^{1/3}}{2} \left(\kappa A_\rho g^{\rho\nu} \right) \times$$
$$\times \left[\partial_\mu (\kappa \phi^{2/3} A_\nu) - \partial_\nu (\kappa \phi^{2/3} A_\mu) \right]$$
$$= \left(\frac{\phi^{-1} + \kappa^2 A_\rho A_\nu g^{\rho\nu}}{3} \right) \partial_\mu \phi - \left(\frac{\kappa^2 \phi A_\rho g^{\rho\nu}}{2} \right) F_{\mu\nu}$$
$$- \frac{\kappa^2 A_\rho g^{\rho\nu}}{3} \left(A_\nu \partial_\mu \phi - A_\mu \partial_\nu \phi \right) \, ; \qquad \text{(D.26)}$$

where $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual field strength tensor in 4D.

(iii). For one upstairs μ index:

$$\hat{\Gamma}^{\mu}{}_{44} = -\frac{\phi^{1/3} g^{\mu\nu}}{2}\left[\partial_\nu\left(\phi^{2/3}\right)\right] = -\frac{g^{\mu\nu}}{3}\partial_\nu\phi \tag{D.27}$$

(iv). For one μ index up and another one down:

$$\begin{aligned}\hat{\Gamma}^{\mu}{}_{4\nu} = \hat{\Gamma}^{\mu}{}_{\nu 4} = & -\frac{\kappa A_\rho g^{\rho\mu}}{2}\phi^{1/3}\left[\partial_\nu\left(\phi^{2/3}\right)\right] + \frac{\phi^{1/3}}{2}g^{\mu\rho}\left[\partial_\nu(\kappa\phi^{2/3}A_\rho) - \partial_\rho(\kappa\phi^{2/3}A_\nu)\right] \\ = & -\left(\frac{\kappa A_\rho g^{\rho\mu}}{3}\right)\partial_\nu\phi + \frac{\kappa\phi}{2}g^{\mu\rho}F_{\nu\rho} + \frac{\kappa}{3}g^{\mu\rho}(A_\rho\partial_\nu\phi - A_\nu\partial_\rho\phi)\end{aligned} \tag{D.28}$$

(v). Now, considering two μ indices down:

$$\begin{aligned}\hat{\Gamma}^{4}{}_{\mu\nu} = \ldots = & -\frac{\kappa}{2}A_\sigma\Gamma^{\sigma}{}_{\mu\nu} + \frac{\kappa}{2}\phi^{-1}(A_\nu\partial_\mu\phi + A_\mu\partial_\nu\phi) + \frac{\kappa}{2}(\partial_\mu A_\nu + \partial_\nu A_\mu) + \frac{\kappa^3}{2}\phi g^{\sigma\rho}A_\sigma\times \\ & \times(A_\mu F_{\rho\nu} + A_\nu F_{\rho\mu}) - \frac{\kappa}{6}\phi^{-1}g_{\mu\nu}g^{\sigma\rho}A_\sigma\partial_\rho\phi + \frac{\kappa^3}{3}g^{\sigma\rho}A_\sigma A_\mu A_\nu\partial_\rho\phi\end{aligned} \tag{D.29}$$

(vi). Finally, for all indices running from 0 to 3:

$$\begin{aligned}\hat{\Gamma}^{\mu}{}_{\nu\rho} = \ldots = \Gamma^{\mu}{}_{\nu\rho} & - \frac{\kappa^2}{3}A_\nu A_\rho g^{\mu\sigma}\partial_\sigma\phi + \frac{\kappa^2}{2}\phi g^{\mu\sigma}(A_\rho F_{\nu\sigma} + A_\nu F_{\rho\sigma}) - \frac{1}{6\phi}(\delta^{\mu}{}_{\rho}\partial_\nu\phi + \\ & + \delta^{\mu}{}_{\nu}\partial_\rho\phi - g^{\mu\sigma}g_{\nu\rho}\partial_\sigma\phi)\end{aligned} \tag{D.30}$$

Note that, in equations (D.29) and (D.30), we have omitted the algebra done and have jumped right to the final expression to make it shorter. No physics insights are needed to get to the final expressions though. One can check that all the Christoffel symbols in 5D were expressed above. One must calculate now the Ricci scalar as in equation (D.12). This would demand a bunch of algebra, i.e. it is a much longer calculation than in the warm up case, with no physical relevance so we will jump to the end and claim that the final expression one gets is, see [46] to check it:

$$\phi^{-1/3}\hat{R} = R - \frac{\kappa^2}{4}\phi F_{\mu\nu}F^{\mu\nu} - \frac{1}{6}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2} + \nabla_\mu X^\mu\,; \tag{D.31}$$

where X^μ is a vector field whose exact form doesn't concern us since it doesn't contribute to the equations of motion as we have already seen. Therefore, the dimensionally reduced Einstein-Hilbert action becomes:

$$\begin{aligned}\hat{S}_{E-H} &= \left(\frac{1}{16\pi\hat{G}}\int dy\right)\int d^4x\sqrt{-g}\phi^{-1/3}\hat{R} \\ &= \int d^4x\sqrt{-g}\left[\frac{R}{16\pi G} - \frac{1}{4}\phi F_{\mu\nu}F^{\mu\nu} - \frac{1}{6(16\pi G)}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2}\right] + \int_\Omega d^4x\partial_\mu(\sqrt{-g}X^\mu) \\ &= \int d^4x\sqrt{-g}\left[\frac{R}{16\pi G} - \frac{1}{4}\phi F_{\mu\nu}F^{\mu\nu} - \frac{1}{6(16\pi G)}\frac{\partial_\mu\phi\partial^\mu\phi}{\phi^2}\right] + \oint_{\partial\Omega} d\Sigma_\mu\sqrt{-g}X^\mu \end{aligned} \tag{D.32}$$

Again, one can redefine the scalar field in order to get the usual kinetic term, $\phi \to \varphi \equiv -\frac{\log\phi}{4\sqrt{3\pi G}} \Rightarrow \phi = e^{-4\sqrt{3\pi G}\varphi} = e^{-\sqrt{6}\varphi/M_{Pl}}$. Thus:

$$\hat{S}_{E-H} = \int d^4x\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{1}{4}e^{-\sqrt{6}\frac{\varphi}{M_{Pl}}}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi\right) + \oint_{\partial\Omega} d\Sigma_\mu\sqrt{-g}X^\mu \tag{D.33}$$

One can see that there is no linear term in A_μ in the above action, that is why it was said to be consistent to put $A_\mu = 0$ in the warm up example. The definition of consistence is discussed in chapter 3. The warm up example is equivalent to a specific type of Brans-Dicke theory, but since it will not be our main focus here, we only refer to [41].

Equation (D.33) is the final version of the Kaluza-Klein original action in terms of 4D fields we were looking for in chapter 3. The calculation developed here is also important in chapter 6.

Bibliography

[1] R. H. Brandenberger and C. Vafa, "Superstrings in the Early Universe," *Nucl. Phys. B* **316** (1989) 391–410.

[2] S. Alexander, R. H. Brandenberger, and D. A. Easson, "Brane gases in the early universe," *Phys. Rev. D* **62** (2000) 103509, `arXiv:hep-th/0005212`.

[3] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," *JHEP* **05** (2003) 013, `arXiv:astro-ph/0210603`.

[4] P. McFadden and K. Skenderis, "The Holographic Universe," *J. Phys. Conf. Ser.* **222** (2010) 012007, `arXiv:1001.2007 [hep-th]`.

[5] P. McFadden and K. Skenderis, "Holography for Cosmology," *Phys. Rev. D* **81** (2010) 021301, `arXiv:0907.5542 [hep-th]`.

[6] H. Nastase, "Solution of the cosmological constant problem within holographic cosmology," *Phys. Lett. B* **801** (2020) 135168, `arXiv:1812.07597 [hep-th]`.

[7] J. Khoury and A. Weltman, "Chameleon fields: Awaiting surprises for tests of gravity in space," *Phys. Rev. Lett.* **93** (2004) 171104, `arXiv:astro-ph/0309300`.

[8] J. Khoury and A. Weltman, "Chameleon cosmology," *Phys. Rev. D* **69** (2004) 044026, `arXiv:astro-ph/0309411`.

[9] A. Einstein, "Zur elektrodynamik bewegter körper," *Annalen der Physik* **17** (1905) no. 10, 891–921.

[10] A. Einstein, "Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?," *Annalen der Physik* **18** (1905) 639–641.

[11] E. Hubble, "A Relation Between Distance and Radial Velocity among Extra-Galactic Nebulae," *Proceedings of the national academy of sciences* **15** (1929) no. 3, 168–173.

[12] H. Năstase, *Cosmology and String Theory*, vol. 197 of *Fundamental Theories of Physics*. Springer, 2019.

[13] R. M. Wald, *General Relativity*. Chicago Univ. Pr., Chicago, USA, 1984.

[14] S. M. Carroll, *Spacetime and Geometry*. Cambridge University Press, 7, 2019.

[15] A. Einstein, "Die Feldgleichungen der Gravitation," *Sitzung der physikalische-mathematischen Klasse* **25** (1915) 844–847.

[16] D. Hilbert, "Die Grundlagen der Physik.(Erste Mitteilung.)," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* **1915** (1915) 395–408.

[17] E. W. Kolb and M. S. Turner, *The Early Universe*, vol. 69. CRC Press, 2018.

[18] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2018 Results. VI. Cosmological Parameters," *Astron. Astrophys.* **641** (2020) A6, `arXiv:1807.06209 [astro-ph.CO]`.

[19] S. Birrer *et al.*, "H0LiCOW - IX. Cosmographic Analysis of the Doubly Imaged Quasar SDSS 1206+4332 and a New Measurement of the Hubble Constant," *Mon. Not. Roy. Astron. Soc.* **484** (2019) 4726, `arXiv:1809.01274 [astro-ph.CO]`.

[20] A. G. Riess *et al.*, "Milky Way Cepheid Standards for Measuring Cosmic Distances and Application to Gaia DR2: Implications for the Hubble Constant," *Astrophys. J.* **861** (2018) no. 2, 126, `arXiv:1804.10655 [astro-ph.CO]`.

[21] K. C. Wong *et al.*, "H0LiCOW – XIII. A 2.4 per Cent Measurement of H0 from Lensed Quasars: 5.3σ Tension Between Early- and Late-Universe Probes," *Mon. Not. Roy. Astron. Soc.* **498** (2020) no. 1, 1420–1439, `arXiv:1907.04869 [astro-ph.CO]`.

[22] V. Poulin, T. L. Smith, T. Karwal, and M. Kamionkowski, "Early Dark Energy Can Resolve The Hubble Tension," *Phys. Rev. Lett.* **122** (2019) no. 22, 221301, `arXiv:1811.04083 [astro-ph.CO]`.

[23] F. Zwicky, "Die Rotverschiebung von Extragalaktischen Nebeln," *Helv. Phys. Acta* **6** (1933) 110–127.

[24] A. A. Starobinsky, "Spectrum of Relict Gravitational Radiation and the Early State of the Universe," *JETP Lett.* **30** (1979) 682–685.

[25] A. A. Starobinsky, "A New Type of Isotropic Cosmological Models Without Singularity," *Phys. Lett. B* **91** (1980) no. 1, 99 – 102. [Adv. Ser. Astrophys. Cosmol.3,130(1987)].

[26] A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," *Phys. Rev. D* **23** (1981) 347–356. [Adv. Ser. Astrophys. Cosmol.3,139(1987)].

[27] A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," *Phys. Lett. B* **108** (1982) no. 6, 389–393. [Adv. Ser. Astrophys. Cosmol.3,149(1987)].

[28] A. Albrecht and P. J. Steinhardt, "Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking," *Phys. Rev. Lett.* **48** (1982) 1220–1223. [Adv. Ser. Astrophys. Cosmol.3,158(1987)].

[29] T. Kibble, "Topology of Cosmic Domains and Strings," *J. Phys. A* **9** (1976) 1387–1398.

[30] T. Kibble, "Some Implications of a Cosmological Phase Transition," *Phys. Rept.* **67** (1980) 183.

[31] S. Weinberg, *Cosmology*. Oxford University Press, 9, 2008.

[32] D. H. Lyth and A. R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure*. Cambridge University Press, 2009.

[33] D. Baumann, "Inflation," in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 523–686. 2011. arXiv:0907.5424 [hep-th].

[34] J. Martin, "Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)," *Comptes Rendus Physique* **13** (2012) 566–665, arXiv:1205.3365 [astro-ph.CO].

[35] H. Năstase, *Introduction to Quantum Field Theory*. Cambridge University Press, 2019.

[36] R. Caldwell, R. Dave, and P. J. Steinhardt, "Cosmological imprint of an energy component with general equation of state," *Phys. Rev. Lett.* **80** (1998) 1582–1585, arXiv:astro-ph/9708069.

[37] K. Hinterbichler, J. Khoury, H. Nastase, and R. Rosenfeld, "Chameleonic inflation," *JHEP* **08** (2013) 053, arXiv:1301.6756 [hep-th].

[38] D. Tong, "String Theory," arXiv:0908.0333 [hep-th].

[39] T. Kaluza, "Zum Unitätsproblem der Physik," *Sitzungsber. Preuss. Akad. Wiss. Berlin.(Math. Phys.)* (1921) .

[40] O. Klein, "Quantentheorie und fünfdimensionale relativitätstheorie," *Zeitschrift für Physik* **37** (1926) .

[41] J. Overduin and P. Wesson, "Kaluza-Klein gravity," *Phys. Rept.* **283** (1997) 303–378, arXiv:gr-qc/9805018.

[42] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "The Hierarchy problem and new dimensions at a millimeter," *Phys. Lett. B* **429** (1998) 263–272, arXiv:hep-ph/9803315.

[43] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "New dimensions at a millimeter to a Fermi and superstrings at a TeV," *Phys. Lett. B* **436** (1998) 257–263, arXiv:hep-ph/9804398.

[44] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373, arXiv:hep-ph/9905221.

[45] L. Randall and R. Sundrum, "An Alternative to compactification," *Phys. Rev. Lett.* **83** (1999) 4690–4693, arXiv:hep-th/9906064.

[46] M. Duff, B. Nilsson, and C. Pope, "Kaluza-Klein Supergravity," *Phys. Rept.* **130** (1986) 1–142.

[47] T. Appelquist and A. Chodos, "Quantum Dynamics of Kaluza-Klein Theories," *Phys. Rev. D* **28** (1983) 772.

[48] H. Năstase, *Introduction to the AdS/CFT Correspondence*. Cambridge University Press, 2015.

[49] H. Năstase, *Classical Field Theory*. Cambridge University Press, 2019.

[50] B. Zwiebach, *A First Course in String Theory*. Cambridge University Press, 2009.

[51] S. Weinberg, "Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations," *Phys. Rev.* **138** (1965) B988–B1002.

[52] J. Dai, R. Leigh, and J. Polchinski, "New Connections Between String Theories," *Mod. Phys. Lett. A* **4** (1989) 2073–2083.

[53] J. Polchinski, *String theory. Vol. 1: An introduction to the bosonic string*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 12, 2007.

[54] E. Witten, "String theory dynamics in various dimensions," *Nucl. Phys. B* **443** (1995) 85–126, arXiv:hep-th/9503124.

[55] R. Bousso, "The Holographic principle," *Rev. Mod. Phys.* **74** (2002) 825–874, arXiv:hep-th/0203101.

[56] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Int. J. Theor. Phys.* **38** (1999) 1113–1133, arXiv:hep-th/9711200.

[57] A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy," *Phys. Lett. B* **379** (1996) 99–104, arXiv:hep-th/9601029.

[58] C. G. Callan and J. M. Maldacena, "D-brane approach to black hole quantum mechanics," *Nucl. Phys. B* **472** (1996) 591–608, arXiv:hep-th/9602043.

[59] J. M. Maldacena and A. Strominger, "Statistical entropy of four-dimensional extremal black holes," *Phys. Rev. Lett.* **77** (1996) 428–429, arXiv:hep-th/9603060.

[60] N. Berkovits and C. Vafa, "Towards a Worldsheet Derivation of the Maldacena Conjecture," *AIP Conf. Proc.* **1031** (2008) 21–42, arXiv:0711.1799 [hep-th].

[61] H. Nastase, "Towards deriving the AdS/CFT correspondence," arXiv:1812.10347 [hep-th].

[62] N. Berkovits, "Sketching a Proof of the Maldacena Conjecture at Small Radius," *JHEP* **06** (2019) 111, arXiv:1903.08264 [hep-th].

[63] G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys. B* **72** (1974) 461–473.

[64] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys. Lett. B* **428** (1998) 105–114, `arXiv:hep-th/9802109`.

[65] E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2** (1998) 253–291, `arXiv:hep-th/9802150`.

[66] A. Borde and A. Vilenkin, "Eternal inflation and the initial singularity," *Phys. Rev. Lett.* **72** (1994) 3305–3309, `arXiv:gr-qc/9312022`.

[67] D. A. Easson, "Towards a stringy resolution of the cosmological singularity," *Phys. Rev. D* **68** (2003) 043514, `arXiv:hep-th/0304168`.

[68] R. H. Brandenberger, "String Gas Cosmology," pp. 193–230. 8, 2008. `arXiv:0808.0746 [hep-th]`.

[69] A. A. Tseytlin and C. Vafa, "Elements of string cosmology," *Nucl. Phys. B* **372** (1992) 443–466, `arXiv:hep-th/9109048`.

[70] T. Battefeld and S. Watson, "String gas cosmology," *Rev. Mod. Phys.* **78** (2006) 435–454, `arXiv:hep-th/0510022`.

[71] D. Mitchell and N. Turok, "Statistical Properties of Cosmic Strings," *Nucl. Phys. B* **294** (1987) 1138–1163.

[72] R. H. Brandenberger, A. Nayeri, S. P. Patil, and C. Vafa, "String gas cosmology and structure formation," *Int. J. Mod. Phys. A* **22** (2007) 3621–3642, `arXiv:hep-th/0608121`.

[73] R. H. Brandenberger, "Moduli stabilization in string gas cosmology," *Prog. Theor. Phys. Suppl.* **163** (2006) 358–372, `arXiv:hep-th/0509159`.

[74] R. H. Brandenberger, "String Gas Cosmology after Planck," *Class. Quant. Grav.* **32** (2015) no. 23, 234002, arXiv:1505.02381 [hep-th].

[75] A. Strominger, "The dS/CFT correspondence," *JHEP* **10** (2001) 034, arXiv:hep-th/0106113.

[76] E. Witten, "Quantum gravity in de Sitter space," in *Strings 2001: International Conference*. 6, 2001. arXiv:hep-th/0106109.

[77] N. Afshordi, E. Gould, and K. Skenderis, "Constraining holographic cosmology using Planck data," *Phys. Rev. D* **95** (2017) no. 12, 123505, arXiv:1703.05385 [astro-ph.CO].

[78] N. Afshordi, C. Coriano, L. Delle Rose, E. Gould, and K. Skenderis, "From Planck data to Planck era: Observational tests of Holographic Cosmology," *Phys. Rev. Lett.* **118** (2017) no. 4, 041301, arXiv:1607.04878 [astro-ph.CO].

[79] H. Nastase and K. Skenderis, "Holography for the very early Universe and the classic puzzles of Hot Big Bang cosmology," *Phys. Rev. D* **101** (2020) no. 2, 021901, arXiv:1904.05821 [hep-th].

[80] H. Nastase, "Holographic cosmology solutions of problems with pre-inflationary cosmology," *JHEP* **12** (2020) 026, arXiv:2008.05630 [hep-th].

[81] H. Nastase and U. M. Portugal, "The monopole problem in holographic cosmology," *Phys. Lett. B* **815** (2021) 136141, arXiv:2012.14829 [hep-th].

[82] H. Nastase, "Reheating in holographic cosmology and connecting to Λ-MSSM constructions for particle physics," arXiv:2008.06144 [hep-th].

[83] K. Hinterbichler, J. Khoury, and H. Nastase, "Towards a UV Completion for Chameleon Scalar Theories," *JHEP* **03** (2011) 061, arXiv:1012.4462 [hep-th]. [Erratum: JHEP 06, 072 (2011)].

[84] P. Brax, C. van de Bruck, A. C. Davis, J. Khoury, and A. Weltman, "Chameleon dark energy," *AIP Conf. Proc.* **736** (2004) no. 1, 105–110, arXiv:astro-ph/0410103.

[85] S. Kachru, R. Kallosh, A. D. Linde, and S. P. Trivedi, "De Sitter vacua in string theory," *Phys. Rev. D* **68** (2003) 046005, arXiv:hep-th/0301240.

[86] G. Obied, H. Ooguri, L. Spodyneiko, and C. Vafa, "De Sitter Space and the Swampland," arXiv:1806.08362 [hep-th].

[87] J. K. Hoskins, R. D. Newman, R. Spero, and J. Schultz, "Experimental tests of the gravitational inverse square law for mass separations from 2-cm to 105-cm," *Phys. Rev. D* **32** (1985) 3084–3095.

[88] H. Nastase and A. Weltman, "A natural cosmological constant from chameleons," *Phys. Lett. B* **747** (2015) 200–204, arXiv:1302.1748 [hep-th].

[89] H. Nastase and A. Weltman, "Chameleons on the Racetrack," *JHEP* **08** (2013) 059, arXiv:1301.7120 [hep-th].