

The Cosmological Constant Problem, an Inspiration for New Physics

ISBN-10: 90-393-4269-5
ISBN-13: 978-90-393-4269-5

Cover Illustration: Stellar “eggs” emerge from molecular cloud. Star-Burst Clouds in M16. PRC95-44b, ST Scl OPO, November 2, 1995. J. Hester and P. Scowen (AZ. State University), NASA. HST-WFPC2, from <http://Hubblesite.org>. Symmetrized.

The Cosmological Constant Problem, an Inspiration for New Physics

Het Vraagstuk van de Kosmologische Constante,
een Inspiratie voor Nieuwe Fysica

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht, op gezag van de rector mag-
nificus, prof. dr. W.H. Gispen, ingevolge het
besluit van het college voor promoties in het
openbaar te verdedigen op donderdag 15 juni
2006 des middags te 2.30 uur

door

Stefanus Johannes Bernardus Nobbenhuis

geboren op 21 januari 1977 te Almelo

Promotor: Prof. Dr. G. 't Hooft

Preface

Standing outside, especially in the countryside, and looking up at the night sky it is hard not to be overwhelmed by the incredible extensiveness of our universe. Uncountable numbers of different stars in many galaxies, some over billions of lightyears away, seem to turn our spot in that vast world truly insignificant. For many centuries people have looked up to the sky wondering about its origin, its evolution and its contents.

Amazingly, through work of just the past few decades, we are now in a position to also truly scientifically discuss these cosmological questions. Much of the credits that have made this journey possible go to the astronomers and other experimental groups that have contributed to this dramatic progress. We will encounter their results at various instants throughout this work.

On the theoretical side, the two main building blocks are Einstein's General Relativity describing the force of gravity, and the Standard Model of elementary particle physics, describing the other three interactions: electromagnetism, and the weak and strong nuclear forces. Both general relativity and the standard model have been tested thoroughly and work excellently in their ranges of validity. General relativity is typically a theory describing physics over large distances, ranging from a tenth of a millimeter to sizes as large as the entire universe, while the Standard Model especially describes interactions and processes that take place at extremely short distances, shorter than the radius of the nucleus of an atom. Although we are all very familiar with the force of gravity from daily life, it is actually a very weak force, that only becomes important when large clumps of matter combine to make heavy objects like the earth. Moreover, these large clumps generally do not carry net electric or magnetic charges, so the electromagnetic interaction is negligible. That is why we can often separate the force of gravity from the other interactions: most of the time these two do not overlap.

In cosmology however, there are several situations where the strength of the gravitational interactions becomes of the same order of magnitude as the strength of the particle interactions described by the Standard Model, and the two theories cannot be treated separately anymore. This is the case for example during the extreme conditions at the beginning of the universe and in the treatment of black holes, or classical spacetime singularities in general. Under these extreme conditions, the two theories should be somehow combined in order to describe properly what is going on. It seems that a quantized theory of gravity is needed. This combination however of quantum mechanics and gravity, leads to severe clashes. General relativity is, in the language of field theory, a non-renormalizable theory. Infinities occur that cannot be removed by the standard renormalization procedure that works so well for ordinary quantum field theory. Instead, they have to be cancelled separately with new counterterms, introducing new free parameters in the theory. Obviously a very unsatisfactory situation, since

much predictability is lost. Moreover, knowledge from the past has taught us that this signals that we do not understand the short distance behavior of the theory, so clearly something better is needed.

Fortunately, due to the fact that gravity is such a weak force, the length scale at which the classical description of spacetime breaks down and some quantum theory of gravity seems to be needed is extremely small, about 10^{-33} cm, called the Planck length. This corresponds to an energy scale, the Planck energy, of 10^{19} GeV, some 16 orders of magnitude higher than what can be experimentally tested at the most advanced particle accelerators on earth¹. Nevertheless, it is an unsatisfactory situation that the two pillars of modern theoretical physics seem so incompatible, and even contradicting when one tries to combine them. Moreover, many observable phenomena related to much larger, even macroscopic distance scales, such as the apparent 4-dimensionality of our world, perhaps could be explained from such a theory.

One of the most peculiar and pressing clashes of a combination of gravity and quantum theory, where many of the subtleties involved in this merger arise, can be found in another large distance observation. This concerns the subject of this thesis: The Cosmological Constant problem.

The cosmological constant is a measure of the energy density of the vacuum and, through Einstein's equations of general relativity, causes spacetime to expand. This expansion can be measured, and an upper bound on the value of the cosmological constant can actually easily be obtained from everyday life. In an expanding universe the wavelengths of radiation, and thus light in particular, becomes redshifted. The dimension of a cosmological constant is one over distance squared, and if the cosmological constant would be as large as $1/\text{meter}^2$, we would see blue objects at a distance of a meter redder than they really are. We do not see such effects, not even at distances of many millions of lightyears, which already implies that the cosmological constant must be extremely small. However, from the physics we know, we have no understanding why this parameter should be so small. On the contrary, quantum field theory seems to predict a huge vacuum energy! Only by very carefully adjusting the different parameters in our theories to incredible sensitivity can we recover the world of our daily experience. This seems very artificial and an indication that we miss an important point here. In this thesis, we are trying to figure out what this point might be. We compare different approaches that have been tried over the years to solve this problem, to find out whether they can give us some insights about which directions we have to look at for a solution.

"I have again perpetrated something relating to the theory of gravitation that might endanger me of being committed to a madhouse. (Ich habe wieder etwas verbochen in der Gravitationstheorie, was mich ein wenig in Gefahr bringt, in ein Tollhaus interniert zu werden.)".

Einstein in a letter to P. Ehrenfest on February 4th, 1917 (CPAE 8, Doc. 294).

¹These orders of magnitude may change considerably in the presence of extra dimensions, but so far no experimental indication of their existence has been obtained.

Contents

Preface	v
1 The Cosmological Constant Problem	1
1.1 The ‘Old’ Problem, Vacuum Energy as Observable Effect	1
1.1.1 Reality of Zero-Point Energies	5
1.1.2 Repulsive Gravity	5
1.2 Two Additional Problems	6
1.3 Renormalization	7
1.4 Interpretations of a Cosmological Constant	10
1.4.1 Cosmological Constant as Lagrange Multiplier	10
1.4.2 Cosmological Constant as Constant of Integration	11
1.5 Where to Look for a Solution?	11
1.5.1 Weinberg’s No-Go Theorem	14
1.5.2 Some Optimistic Numerology	15
1.6 Outline of this thesis	16
2 Cosmological Consequences of Non-Zero Λ	19
2.1 The Expanding Universe	19
2.1.1 Some Historic Objections	22
2.2 Some Characteristics of FRW Models	22
2.2.1 Redshift	25
2.3 The Early Universe	28
2.3.1 Deriving Geometric Information from CMBR Anisotropies	30
2.3.2 Energy Density in the Universe	33
2.4 Evolution of the Universe	35
2.5 Precision cosmology in recent and upcoming experiments	37
2.5.1 Supernovae Type Ia	37
2.5.2 Pioneer Anomaly	39

3	Type I: Symmetry Mechanism	43
3.1	Supersymmetry	44
3.1.1	No-Scale SUGRA	45
3.1.2	Unbroken SUSY	47
3.1.3	Non-SUSY String Models	48
3.2	Scale Invariance, e.g. Conformal Symmetry	48
3.2.1	Λ as Integration Constant	50
3.3	Holography	51
3.3.1	Conceptual Issues in de Sitter Space	53
3.4	“Symmetry” between Sub- and Super-Planckian Degrees of Freedom	55
3.5	Interacting Universes, Antipodal Symmetry	56
3.6	Duality Transformations	57
3.6.1	S-Duality	57
3.6.2	Hodge Duality	58
3.7	Summary	58
4	Invariance Under Complex Transformations	61
4.1	Introduction	61
4.2	Classical Scalar Field	62
4.3	Gravity	63
4.4	Non-relativistic Particle	64
4.5	Harmonic Oscillator	66
4.6	Second Quantization	66
4.7	Pure Maxwell Fields	70
4.8	Relation with Boundary Conditions	72
4.9	Related Symmetries	73
4.9.1	Energy \rightarrow $-$ Energy	76
4.10	Summary	76
5	Type II: Back-reaction	79
5.1	Scalar Field, Instabilities <i>in</i> dS-Space	79
5.1.1	‘Cosmon’ Screening	82
5.1.2	Radiative Stability in Scalar Field Feedback Mechanism	84
5.1.3	Dilaton	84
5.2	Instabilities <i>of</i> dS-Space	85
5.2.1	Scalar-type Perturbations	85

5.3	Running Λ from Renormalization Group	89
5.3.1	Triviality as in $\lambda\phi^4$ Theory	94
5.4	Screening as a Consequence of the Trace Anomaly	95
5.5	Summary	96
6	Do Infrared Gravitons Screen the Cosmological Constant?	97
6.1	Introduction	97
6.2	A Review of the Scenario	98
6.2.1	A Newtonian Picture	99
6.3	Evaluation	101
6.4	The Full Quantum Gravitational Calculation	102
6.5	Renormalization in Perturbative QG	107
6.6	Summary	108
7	Type III: Violating the Equivalence Principle	109
7.1	Extra Dimensions, Braneworld Models	109
7.2	Randall-Sundrum Models, Warped Extra Dimensions	110
7.2.1	Self-Tuning Solutions	112
7.2.2	Extra Time-like Dimensions	114
7.3	Infinite Volume Extra Dimensions	114
7.3.1	Massive Gravitons	121
7.3.2	Non-Zero Brane Width	124
7.3.3	Non-local Gravity	124
7.3.4	A 5D DGP Brane-World Model	126
7.4	Ghost Condensation or Gravitational Higgs Mechanism	127
7.5	Fat Gravitons	129
7.6	Composite Graviton as Goldstone boson	131
7.6.1	Summary	132
8	Type IV: Statistical Approaches	135
8.1	Hawking Statistics	137
8.1.1	Wormholes	139
8.2	Anthropic Principle	142
8.2.1	Anthropic Prediction for the Cosmological Constant	143
8.2.2	Discrete vs. Continuous Anthropic Principle	146
8.3	Summary	146

9	Conclusions	149
9.1	Future Evolution of the Universe	150
A	Conventions and Definitions	151
B	Measured Values of Different Cosmological Parameters	153
B.1	Hubble Constant	153
B.2	Total Energy Density	153
B.3	Matter in the Universe	154
B.4	Dark Energy Equation of State	154
B.5	Summary	155
	Nederlandse Samenvatting	157
	Dankwoord	165
	Curriculum Vitae	167
	Bibliography	187

1 | The Cosmological Constant Problem

In this chapter we carefully discuss the different aspects of the cosmological constant problem. How it occurs, what the assumptions are, and where the difficulties lie in renormalizing vacuum energy. Different interpretations of the cosmological constant have been proposed and we briefly discuss these. We subsequently, as a warm-up, point at some directions to look at for a solution, before ending this chapter by giving an outline of the rest of the thesis.

1.1 The ‘Old’ Problem, Vacuum Energy as Observable Effect

In quantum mechanics, the energy spectrum, E_n , of a simple harmonic oscillator¹ is given by $E_n = (n + 1/2)\omega$. The energy of the ground state, the state with lowest possible energy, is non-zero, contrary to classical mechanics. This so-called zero-point energy is usually interpreted by referring to the uncertainty principle: a particle can never completely come to a halt.

Free quantum field theory is formulated as an infinite series of simple harmonic oscillators. The energy density of the vacuum in for example free scalar field theory therefore receives an infinite positive contribution from the zero-point energies of the various modes of oscillation:

$$\langle \rho \rangle = \int_0^\infty \frac{d^3k}{(2\pi)^3} \frac{1}{2} \sqrt{k^2 + m^2}, \quad (1.1)$$

and we set $\omega(k) = \sqrt{k^2 + m^2}$. With a UV-cutoff k_{\max} , regularizing the a priori infinite value for the vacuum energy density, this integral diverges² as k_{\max}^4 .

The filling of the ‘Dirac sea’ in the quantization of the free fermion theory leads to a downward shift in the vacuum energy with a similar ultraviolet divergence. The Hamiltonian is:

$$H = \int \frac{d^3k}{(2\pi)^3} \omega \sum_s (a^{s\dagger} a^s - b^{s\dagger} b^s), \quad (1.2)$$

where the $b_k^{s\dagger}$ creates negative energy. Therefore we have to require that these operators satisfy anti-commutation relations:

$$\{b_k^r, b_{k'}^{s\dagger}\} = (2\pi)^3 \delta^{(3)}(k - k') \delta^{rs}, \quad (1.3)$$

since this is symmetric between b^r and $b^{r\dagger}$, we can redefine the operators as follows:

$$\tilde{b}^s \equiv b^{s\dagger}; \quad \tilde{b}^{s\dagger} \equiv b^s, \quad (1.4)$$

¹See Appendix A for conventions and some definitions used throughout this thesis, e.g. $\hbar = 1$.

²This is just the leading divergence, see section (1.3) for a more precise discussion.

which obey the same anti-commutation relations. Now the second term in the Hamiltonian becomes:

$$-\omega b^{s\dagger} b^s = +\omega \tilde{b}^{s\dagger} \tilde{b}^s - (\text{const.}); \quad (\text{const}) = \int^{k_{max}} \frac{d^3 k}{(2\pi)^3} \omega(k). \quad (1.5)$$

Now we choose $|0\rangle$ to be the vacuum state that is annihilated by the a^s and \tilde{b}^s and all excitations have positive energy. All the negative energy states are filled; this is the Dirac sea. A hole in the sea corresponds to an excitation of the ground state with positive energy, compared to the ground state. The infinite constant contribution to the vacuum energy density has the same form as in the bosonic case, but enters with opposite sign.

Spontaneous symmetry breaking gives a finite but still possibly large shift in the vacuum energy density. In this case,

$$\mathcal{L} = \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \quad (1.6)$$

where the potential $V(\phi)$ is given by:

$$V(\phi) = -\mu^2(\phi^\dagger \phi) + \lambda(\phi^\dagger \phi)^2 + \epsilon_0, \quad (1.7)$$

where μ and λ both are positive constants and ϵ_0 is a constant with arbitrary sign. The potential is at its minimum value for $|\phi| = \sqrt{\mu^2/2\lambda}$, leading to a shift in the energy density of the ground state:

$$\langle \rho \rangle = \epsilon_0 - \frac{\mu^4}{4\lambda}. \quad (1.8)$$

The spontaneous breaking of the weak interaction $SU(2) \times U(1)$ symmetry and of the strong interaction chiral symmetry both would be expected to shift the vacuum energy density in this way. Through the additive constant ϵ_0 the minimum of the potential either after, or before symmetry breaking, can be tuned to any value one likes, but within field theory, this value is completely arbitrary.

More generally, in elementary particle physics experiments, the absolute value of the vacuum energy is unobservable. Experimentally measured particle masses, for example, are energy differences between the vacuum and certain excited states of the Hamiltonian, and the absolute vacuum energy cancels out in the calculation of these differences.

In GR however, each form of energy contributes to the energy-momentum tensor $T_{\mu\nu}$, hence gravitates and therefore reacts back on the spacetime geometry, as can be seen from Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} - \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}. \quad (1.9)$$

As it stands, the cosmological constant is a free parameter that can be interpreted as the curvature of empty spacetime. However, Lorentz invariance tells us that the energy-momentum tensor of the vacuum looks like:

$$T_{vac}^{\mu\nu} = -\langle \rho \rangle g^{\mu\nu}, \quad (1.10)$$

since the absence of a preferred frame for the vacuum means that $T_{vac}^{\mu\nu}$ must be the same for all observers connected by Lorentz transformations. Apart from zero, there is

only one isotropic tensor of rank two, which is the metric tensor. If we move the CC term to the right-hand-side of the equation (1.9), it looks exactly the same:

$$T_{vac}^{\mu\nu} = -\frac{\Lambda}{8\pi G}g^{\mu\nu}. \quad (1.11)$$

This interpretation of a cosmological constant measuring the energy density of the vacuum, was first explicitly given by Zel’dovich in 1968 [1, 2].

We can therefore define an effective cosmological constant³:

$$\Lambda_{eff} = \Lambda + 8\pi G\langle\rho\rangle, \quad \text{or} \quad \rho_{vac} = \langle\rho\rangle + \Lambda/8\pi G. \quad (1.12)$$

The vacuum energy density calculated from normal quantum field theory thus has a potentially significant observable effect through the coupling with gravity. It contributes to the effective cosmological constant and an infinite value would possibly generate an infinitely large spacetime curvature through the semiclassical Einstein equations:

$$G_{00} = -8\pi G\langle T_{00}\rangle. \quad (1.13)$$

Note at this point that only the effective cosmological constant, Λ_{eff} , is observable, not Λ , so the latter quantity may be referred to as a ‘bare’, or perhaps classical quantity that has to be ‘dressed’ by quantum corrections, analogously to all other physical parameters in ordinary quantum field theory. If one believes quantum field theory to be correct all the way up to the Planck scale, at 10^{19} GeV, then this scale provides a natural ultraviolet cutoff to all field theory processes like eqn. (1.1). Such a cutoff however, would lead to a vacuum energy density of $(10^{19} \text{ GeV})^4 = 10^{76} \text{ GeV}^4$, which is roughly 123 orders of magnitude larger than the currently observed value:

$$\rho_{vac} \simeq 10^{-47} \text{ GeV}^4 \sim 10^{-29} \text{ g/cm}^3 \sim (0, 1 \text{ mm})^{-4} \sim (10^{-12} \text{ s})^{-4}. \quad (1.14)$$

This means that the bare cosmological constant Λ has to be fine-tuned to a stunning 123 decimal places, in order to yield the correct physical result⁴. As is well known, even if we take a TeV scale cut-off the difference between experimental and theoretical results still requires a fine-tuning of about 50 orders of magnitude. Even a cutoff at for example the QCD scale (at ~ 200 MeV), worrying only about zero-point energies in quantum chromodynamics, would not help much; such a cutoff would still lead to a discrepancy of about 40 orders of magnitude. The answer clearly has to lie somewhere else. Even non-perturbative effects, like ordinary QCD instantons, would give far too large a contribution if not cancelled by some fundamental mechanism.

Physicists really started to worry about this in the mid-seventies, after the success of spontaneous symmetry breaking. Veltman [3] in 1975 “. . . concluded that this undermines the credibility of the Higgs mechanism.”

We have no understanding of why the effective cosmological constant, Λ_{eff} is so much smaller than the vacuum energy shifts generated in the known phase transitions of particle physics, and so much smaller again than the underlying field zero-point energies. No symmetry is known that can protect the cosmological constant to such a small

³Note that using this definition we use units in which the cosmological constant has dimension GeV^2 throughout.

⁴This often mentioned factor ~ 120 , depends on the dimension used for energy density. In Planckian units, the factor 120 is the correct one, relating dimensionless numbers.

value⁵. The magnificent fine-tuning needed to obtain the correct physical result seems to suggest that we miss an important point. In this thesis we give an overview of the main ideas that have appeared in trying to figure out what this point might be. For instance, one might suspect that it was naive to believe that vacuum energy, like any other form of energy contributes to the energy-momentum tensor and gravitates. It is however an assumption that is difficult to avoid.

In conclusion, the question is why is the effective cosmological constant so close to zero? Or, in other words, why is the vacuum state of our universe (at present) so close to the classical vacuum state of zero energy, or perhaps better, why is the resulting four-dimensional curvature so small, or why does Nature prefer a flat spacetime? Apparently spacetime is such, that it takes a lot of energy to curve it, while stretching it is (almost) for free, since the cosmological constant is (almost) zero. This is quite contrary to properties of objects from every day experience, where bending requires much less energy than stretching.

As a nice example, Pauli in the early 1920's was way ahead of his time when he wondered about the gravitational effect of the zero-point energy of the radiation field. He used a cutoff for the radiation field at the classical electron radius and realized that the entire universe "could not even reach to the moon" [4]. The calculation is straightforward and also restated in [5]. The vacuum energy-density of the radiation field is:

$$\langle \rho \rangle = \frac{8\pi}{(2\pi)^3} \int_0^{\omega_{max}} \frac{1}{2} \omega^3 d\omega = \frac{1}{8\pi^2} \omega_{max}^4 \quad (1.15)$$

With the cutoff $\omega_{max} = 2\pi/\lambda_{max} = \frac{2\pi m_e}{\alpha}$. This implies for the radius of the universe:

$$R \sim \frac{\alpha^2}{\pi} \frac{M_{Pl}}{m_e^2} \sim 31 \text{ km}, \quad (1.16)$$

indeed far less than the distance to the moon.

Nobody else seems to have been bothered by this, until the late 1960's when Zel'dovich [1, 2] realized that, even if one assumes that the zero-point contributions to the vacuum energy density are exactly cancelled by a bare term, there still remain very problematic higher order effects. On dimensional grounds the gravitational interaction between particles in vacuum fluctuations would be of order $G\mu^6$, with μ some cutoff scale. This corresponds to two-loop Feynman diagrams. Even for μ as low as 1 GeV, this is about 9 orders of magnitude larger than the observational bound.

This illustrates that all 'naive' predictions for the vacuum energy density of our universe are greatly in conflict with experimental facts, see table (1.1) for a list of order-of-magnitude contributions from different sources. All we know for certain is that the unification of quantum field theory and gravity cannot be straightforward, that there is some important concept still missing from our understanding. Note that the divergences in the cosmological constant problem are even more severe than in the case of the Higgs mass: The main divergences here are quartic, instead of quadratic. It is clear that the cosmological constant problem is one of the major obstacles for quantum gravity and cosmology to further progress.

⁵See however chapter (4) for an attempt.

Table 1.1: Energy density of the Quantum Vacuum

Observed	$\rho \simeq 10^{-30}$	gcm^{-3}
Quantum Field Theory	$\rho = \infty$	gcm^{-3}
Quantum Gravity	$\rho \simeq 10^{+90}$	gcm^{-3}
Supersymmetry	$\rho \approx 10^{+30}$	gcm^{-3}
Higgs Potential	$\rho \approx -10^{+25}$	gcm^{-3}
Other Sources	$\rho \approx \pm 10^{+20}$	gcm^{-3}

1.1.1 Reality of Zero-Point Energies

The reality of zero-point energies, i.e. their observable effects, has been a source of discussion for a long time and so far without a definite conclusion about it. Besides the gravitational effect in terms of a cosmological constant, there are two other observable effects often ascribed to the existence of zero-point energies: the Lamb shift and the Casimir effect.

A recent investigation by Jaffe [6] however, concluded that neither the experimental confirmation of the Casimir effect, nor of the Lamb shift, established the reality of zero-point fluctuations. However, no completely satisfactory description of, for example QED, is known in a formulation without zero-point fluctuations.

Note that in the light-cone formulation, a consistent description of e.g. QED can be given, with the vacuum energy density automatically set to zero. This is accomplished by selecting two preferred directions, and thus breaks Lorentz invariance. However, there is no physical reason to select these coordinates, and moreover, a discussion of tadpole diagrams and cosmological terms in de Sitter spacetime becomes problematic in these coordinates.

1.1.2 Repulsive Gravity

A positive cosmological constant gives a repulsive gravitational effect. This can easily be seen by considering the static gravitational field created by a source mass M at the origin, with density $\rho_M = M\delta^3(r)$. For weak fields $g_{\mu\nu} \simeq \eta_{\mu\nu}$ is the usual Lorentz metric. If we further assume a non-relativistic regime with $T_{00} \simeq \rho_M$ the (0, 0)-component of the Einstein equation reads:

$$G_{00} + \Lambda = -8\pi G_N \rho_M, \quad \text{with} \quad G_{00} = R_{00} + \frac{1}{2}R. \quad (1.17)$$

In the non-relativistic case we also expect that $T_{ij} \ll T_{00}$ which is equivalent to saying that we neglect pressure and stress compared to matter density. Therefore we can set $G_{ij} - \Lambda g_{ij} \simeq 0$, implying $R_{ij} \simeq (1/2R + \Lambda)g_{ij}$ and thus the curvature scalar becomes: $R = g^{\mu\nu} R_{\mu\nu} \simeq -R_{00} + 3(1/2R + \Lambda)$ or $R \simeq 2R_{00} - 6\Lambda$. Substituting this back in the

equation for the $(0, 0)$ component we find:

$$R_{00} - \Lambda = -4\pi G_N \rho_M. \quad (1.18)$$

Furthermore we can derive that within our approximation $R_{00} \simeq (-1/2)\nabla^2 g_{00}$. Finally recalling that Newton's potential ϕ is related to the deviation of the $(0, 0)$ -component of the metric tensor from $-\eta_{00} = 1$ (through $-g_{00} \simeq 1 + 2\phi$) we are led to the equation:

$$\nabla^2 \phi = 4\pi G_N \left(\rho_M - \frac{\Lambda}{4\pi G_N} \right). \quad (1.19)$$

This is of course nothing but Poisson's equation for the Newton potential with an additional term from the CC whose sign depends on that of Λ . For $\Lambda > 0$ the original gravitational field is diminished as though there were an additional repulsive interaction. These features are confirmed by explicitly solving the above equation:

$$\phi = -\frac{G_N M}{r} - \frac{1}{6} \Lambda r^2. \quad (1.20)$$

We are thus led to the expected gravitational potential plus a new contribution which is like a "harmonic oscillator" potential, repulsive for $\Lambda > 0$.

1.2 Two Additional Problems

Actually, after the remarkable discoveries and subsequent confirmations starting in 1997 (SN)[7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17] (WMAP) [18, 19] (Boomerang) [20, 21] (SdSS) [22, 23] (Hubble) [24] that the universe really is accelerating its expansion (more about this in the next chapter), there appear to be at present at least three cosmological constant problems.

The first, or sometimes called "old" cosmological constant problem is why is the effective cosmological constant so incredibly small, as described in the previous section.

The second problem is, if it is so small, then why is it not exactly equal to zero? Often in physics it is a lot easier to understand why a parameter is identically zero, than why it is a very small number.

And a third question may be posed, based on the measured value of the effective cosmological constant. The energy density of the vacuum that it represents, appears to be of the same order of magnitude as the present matter energy density in the universe. This is quite peculiar, since, as we will see in chapter (2), vacuum energy density, denoted Ω_Λ , remains constant during the evolution of the universe, whereas the matter energy density, Ω_m obviously decreases as the universe grows larger and larger. If the two energy densities are of the same order of magnitude nowadays, this means that their ratio, Ω_Λ/Ω_m had to be infinitesimal in the early universe, but fine-tuned to become equal now. Therefore, one obviously starts to wonder whether we are living in some special epoch, that causes these two forms of energy density to be roughly equal in magnitude. This has become known as the "cosmic coincidence problem" and is also sometimes phrased as the "Why now?" problem. In this thesis we mainly concentrate on the first problem.

1.3 Renormalization

In quantum field theory the value of the vacuum energy density has no observational consequences. We can simply rescale the zero point of energy, which amounts to adding or subtracting a constant to the action, without changing the equations of motion. It can also be done more elegantly, by a book-keeping method called “normal ordering”, denoted by placing a quantity between semicolons, i.e. $:T_{\mu\nu}:$. In this prescription, one demands that wherever a product of creation and annihilation operators occurs, it is understood that all creation operators stand to the left of all annihilation operators. This differs from the original notation by commutator terms which renormalize vacuum energy and mass, since:

$$\langle\psi| :T_{\mu\nu} : |\psi\rangle = \langle\psi|T_{\mu\nu}|\psi\rangle - \langle 0|T_{\mu\nu}|0\rangle. \quad (1.21)$$

However, it is not a very satisfactory way of dealing with the divergences, considering that we have seen that vacuum energy is observable, when we include gravity. By way of comparison, we note that in solid state physics, through X-ray diffraction, zero-point energy is measurable. Here, a natural UV-cutoff to infinite integrals like (1.1) exists, because the system is defined on a lattice. And indeed, this zero-point energy turns out to be $\omega/2$ per mode.

So let us return to eqn. (1.1). In fact, there are not just quartic divergences in the expression for the vacuum energy, but also quadratic and logarithmic ones. More precisely, for a field with spin $j > 0$:

$$\begin{aligned} \langle\rho\rangle &= \frac{1}{2}(-1)^{2j}(2j+1) \int_0^{\Lambda_{UV}} \frac{d^3k}{(2\pi)^3} \sqrt{\mathbf{k}^2 + m^2} \\ &= \frac{(-1)^{2j}(2j+1)}{16\pi^2} \left(\Lambda_{UV}^4 + m^2\Lambda_{UV}^2 - \frac{1}{4}m^4 \left[\log(\Lambda_{UV}^2/m^2) + \frac{1}{8} - \frac{1}{2}\log 2 \right] \right) \\ &+ \mathcal{O}(\Lambda_{UV}^{-1}). \end{aligned} \quad (1.22)$$

where we have imposed an ultraviolet cutoff Λ_{UV} to the divergent integral. This shows that quartic divergences to the vacuum energy come from any field, massive or massless⁶.

In combination with gravity, these divergences must be subtracted, as usual, by a counterterm, a ‘bare’ cosmological constant, as we will see in the following. Then there are quadratic and logarithmic divergences, but these only arise for massive fields. The quadratic divergences have to be treated in the same way as the quartic ones. The logarithmic divergences are more problematic, because, even after having been cancelled by counterterms, their effect is still spread through the renormalization group. Since the theory is not renormalizable, the infinities at each loop order have to be cancelled separately with new counterterms, introducing new free parameters in the theory. Because of dimensional reasons, the effects of these new terms are small, and are only felt in the UV. However, this is an unsatisfactory situation, since much predictability

⁶It should be noted however, that since this cutoff is only on spatial momenta, it violates Lorentz invariance. It was shown in [25, 26] that the quartic divergence for the energy density and pressure of a free scalar field in fact do not describe vacuum energy, but a homogeneous sea of background radiation. However, when one does the calculation correctly, the quartic divergences are again recovered [27]. When evaluating the integral using dimensional regularization, one finds: $\rho = -m^4/(64\pi^2) [\log \Lambda_{UV}^2/m^2 + 3/2]$; no quartic or quadratic term.

is lost. Moreover, knowledge from the past has taught us that this signals that we do not understand the short distance behavior of the theory, so clearly something better is needed.

The subtraction procedure between bare terms and quantum corrections depends on the energy scale μ in the divergent integral. The renormalization of the cosmological constant therefore must be performed such that at energy scales where it is measured today, we get the observed value. The renormalization condition, in other words, has to be chosen at some fixed energy scale μ_c . For the cosmological constant the renormalization condition becomes:

$$\Lambda_{eff}(\mu_c) = \Lambda_{vac}(\mu_c) + \Lambda_{ind}(\mu_c), \quad (1.23)$$

where Λ_{eff} is the observable, physical cosmological constant, Λ_{vac} is the bare cosmological constant in Einstein's equations, and Λ_{ind} are the quantum corrections. The coupling constants of the theory become a function of μ , where μ is called the renormalization scale. Physical observables must be independent of μ and this is expressed by the renormalization group equations, or more precisely, the one-dimensional subgroup of scale transformations, sending $\mu \rightarrow \alpha\mu$, with α some constant.

Despite the fact that there exists no renormalizable theory combining gravity and particle physics, the hope is that under certain restrictions, these considerations still make sense. It has been tried, in vain, to argue that this type of running with energy scale μ under the renormalization group, could explain the very small value Λ_{eff} nowadays. We will discuss this scenario in section (5.3).

The restrictions mentioned above are quite severe. General relativity is a non-linear theory, which implies that gravitons not only transmit the force of gravity, but also set up a gravitational field themselves, similar, mutatis mutandis, to gluons in QCD. Gravitons 'feel' a gravitational field just as much as for example photons do. The approach usually taken, also in most parts of this thesis, is to consider these graviton contributions as a metric perturbation on a background spacetime, $g_{\mu\nu} = g_{\mu\nu}^{bg} + h_{\mu\nu}$, where $h_{\mu\nu}$ represents the gravitons, or gravitational waves. The gravitational waves are then treated as a null fluid, and considered to contribute to the energy momentum tensor, i.e. to the right-hand-side of Einstein's equation, instead of the left-hand-side. The combined action for gravity and matter fields can be expanded in both g^{bg} and h , and Feynman rules can be derived, see e.g. the lectures by Veltman [28].

As discussed above, at higher orders the gravitational corrections are out of control. However, if one truncates the expansion at a finite number of loops, then the finite number of divergent quantities that appear can be removed by renormalizing a finite number of physical quantities, see e.g. [29, 30, 31] for a more detailed description of this setup.

This procedure provides a way to use the semi-classical Einstein equations:

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = -8\pi G \langle T_{\mu\nu} \rangle, \quad (1.24)$$

where $\langle T_{\mu\nu} \rangle$ is the quantum expectation value of the energy momentum tensor. For this to make any sense at all, usually the expansion is truncated already at one-loop level. In this semi-classical treatment, the gravitational field is treated classically, while the matter fields, including the graviton to one-loop order, are treated quantum mechanically.

To say something meaningful about $\langle T_{\mu\nu} \rangle$ one now has to renormalize this parameter, since the expectation value in principle diverges terribly. To do this in a curved background is not trivial. First one has to introduce a formal regularization scheme, which renders the expectation value finite, but dependent on an arbitrary regulator parameter. Several choices for regularization are available, see e.g. [30].

One option is to separate the spacetime points at which the fields in $T_{\mu\nu}$ are evaluated and then average over the direction of separation. This is a covariant regularization scheme which leaves $\langle T_{\mu\nu} \rangle$ dependent on an invariant measure of the distance between the two points. The price to pay for any regularization scheme is the breaking of conformal invariance, so massless fields no longer have traceless stress-tensors. Often this distance is chosen to be one-half times the square of the geodesic distance between them, denoted by σ . Asymptotically, the regularized expression then becomes:

$$\langle T_{\mu\nu} \rangle \sim A \frac{g_{\mu\nu}}{\sigma^2} + B \frac{G_{\mu\nu}}{\sigma} + (C_1 H_{\mu\nu}^{(1)} + C_2 H_{\mu\nu}^{(2)}) \ln \sigma, \quad (1.25)$$

where A, B, C_1 and C_2 are constants, $G_{\mu\nu}$ is the Einstein tensor and the $H_{\mu\nu}^{(1),(2)}$ are covariantly conserved tensors, quadratic in the Riemann tensor. So this gives a correction to the bare cosmological constant present in the Einstein-Hilbert action (term linear in $g_{\mu\nu}$), a correction to Newton's constant (term $\propto G_{\mu\nu}$) plus higher order corrections to $T_{\mu\nu}$. The tensors $H_{\mu\nu}^{(1),(2)}$ are the functional derivatives with respect to the metric tensor of the square of the scalar curvature and of the Ricci tensor, respectively. With $\delta/\delta g_{\mu\nu}$ the Euler-derivative, one arrives at (see e.g. [30]):

$$\begin{aligned} H_{\mu\nu}^{(1)} &\equiv \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} [\sqrt{-g} R^2] \\ &= 2\nabla_\nu \nabla_\mu R - 2g_{\mu\nu} \nabla_\rho \nabla^\rho R - \frac{1}{2} g_{\mu\nu} R^2 + 2R R_{\mu\nu}, \end{aligned} \quad (1.26)$$

and

$$\begin{aligned} H_{\mu\nu}^{(2)} &\equiv \frac{1}{\sqrt{-g}} \frac{\delta}{\delta g^{\mu\nu}} [\sqrt{-g} R_{\alpha\beta} R^{\alpha\beta}] = 2\nabla_\alpha \nabla_\nu R_\mu^\alpha - \nabla_\rho \nabla^\rho R_{\mu\nu} \\ &\quad - \frac{1}{2} g_{\mu\nu} \nabla_\rho \nabla^\rho R - \frac{1}{2} g_{\mu\nu} R_{\alpha\beta} R^{\alpha\beta} + 2R_\mu^\rho R_{\rho\nu}. \end{aligned} \quad (1.27)$$

The divergent parts of $\langle T_{\mu\nu} \rangle$ can then be taken into account by adding counterterms to the standard Einstein-Hilbert action:

$$S_G = \frac{1}{16\pi G_0} \int d^4x \sqrt{-g} (R - 2\Lambda_0 + \alpha_0 R^2 + \beta_0 R_{\alpha\beta} R^{\alpha\beta}). \quad (1.28)$$

We did not include a term $R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta}$, since it can be absorbed in the other counterterms, using that the combination:

$$\int d^4x \sqrt{-g} (R^2 - 4R_{\mu\nu}^2 + R_{\alpha\beta\mu\nu}^2), \quad (1.29)$$

is a topological invariant and does not affect the field equations. Furthermore, in pure gravity, with no matter fields, the Einstein equations tell us that $R = 0$ and $R_{\mu\nu} = 0$. In this case the counterterms are unphysical, which means, together with the observation

that (1.29) is a topological invariant, that pure gravity has no infinities at one-loop level and thus is one-loop renormalizable.

If we now again include the matter action and vary both with respect to the metric in order to obtain the field equations, and replace the classical $T_{\mu\nu}$ with the quantum mechanical expectation value $\langle T_{\mu\nu} \rangle$ we obtain:

$$G_{\mu\nu} - \Lambda_0 g_{\mu\nu} = -8\pi G_0 \langle T_{\mu\nu} \rangle - \alpha_0 H_{\mu\nu}^{(1)} - \beta_0 H_{\mu\nu}^{(2)}. \quad (1.30)$$

The divergent parts in $\langle T_{\mu\nu} \rangle$ may be removed by renormalizing the bare coupling constants, G_0 , Λ_0 , α_0 and β_0 after which they become the physical parameters of the theory. Of course, in order for this semiclassical treatment to be a good approximation, the terms $H_{\mu\nu}^{(1),(2)}$ are assumed to be very small.

This explicitly clarifies some of the statements we made in this section. At the one-loop level, renormalization of G , Λ , and the coupling constants of two new geometrical tensors, suffices to render the theory finite. At one-loop order, we had to introduce already two new free parameters to absorb the infinities and this only becomes worse at higher orders.

1.4 Interpretations of a Cosmological Constant

Since the 1970's the standard interpretation of a CC is as vacuum energy density. It is an interesting philosophical question, closely connected with the reality of spacetime, whether this interpretation really differs from the original version where the CC was simply a constant, one of the free parameters of the universe. Perhaps the CC is determined purely by the fabric structure of spacetime. However, we won't go into those discussions in this thesis. In this section we consider different interpretations that have sometimes been put forward.

1.4.1 Cosmological Constant as Lagrange Multiplier

The action principle for gravity in the presence of a CC can be written:

$$\begin{aligned} A &= \frac{1}{16\pi G} \int d^4x \sqrt{-g} (R + 2\Lambda) \\ &= \frac{1}{16\pi G} \int d^4x \sqrt{-g} R + \frac{1}{8\pi G} \int d^4x \sqrt{-g} \Lambda \end{aligned} \quad (1.31)$$

This can be viewed as a variational principle where the integral over R is extremized, subject to the condition that the 4-volume of the universe remains constant. The second term has the right structure, to mathematically think of the CC as a Lagrange multiplier ensuring the constancy of the 4-volume of the universe when the metric is varied.

It does not help at all in solving the cosmological constant problem, but it might be useful to have a different perspective.

1.4.2 Cosmological Constant as Constant of Integration

If one assumes that the determinant g of $g_{\mu\nu}$ is not dynamical we could admit only those variations which obey the condition $g^{\mu\nu}\delta g_{\mu\nu} = 0$ in the action principle. The trace part of Einstein's equation then is eliminated. Instead of the standard result, after varying the standard action, without a cosmological constant, we now obtain:

$$R^{\mu\nu} - \frac{1}{4}g^{\mu\nu}R = -8\pi G \left(T^{\mu\nu} - \frac{1}{4}g^{\mu\nu}T^\lambda{}_\lambda \right) \quad (1.32)$$

just the traceless part of Einstein's equation. The general covariance of the action still implies that $T^\mu{}_{;\mu} = 0$ and the Bianchi identities $(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R)_{;\mu} = 0$ continue to hold. These two conditions imply:

$$\partial_\mu R = 8\pi G \partial_\mu T \quad \Rightarrow \quad R - 8\pi GT = \text{constant} \equiv -4\Lambda \quad (1.33)$$

Defining the constant term this way, we arrive at:

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R - \Lambda g^{\mu\nu} = -8\pi GT^{\mu\nu} \quad (1.34)$$

which is precisely Einstein's equation in the presence of a CC, only in this approach is has not so much to do with any term in the action or vacuum fluctuations. It is merely a constant of integration to be fixed by invoking "suitable" boundary conditions for the solutions.

There are two main difficulties in this approach, which has become known as the 'unimodular' approach. The first is how to interpret the assumption that g must remain constant when the variation is performed. See [32] for some attempts in this direction and section (3.2.1) for more details on this. A priori, the constraint keeping the volume-element constant, is just a gauge restriction in the coordinate frame chosen. The second problem is that it still does not give any control over the value of the CC which is even more worrying in case of a non-zero value for the CC.

1.5 Where to Look for a Solution?

The solution to the cosmological constant problem may come from one of many directions. As we will argue in chapter (3) a symmetry would be the most natural candidate. Let us therefore first investigate what the symmetries of the gravity sector are, with and without a cosmological constant, since these are quite different.

The unique vacuum solution to Einstein's equations in the presence of a (negative) cosmological constant is (anti-) de Sitter spacetime, (A)dS for short. Many lectures on physics of de Sitter space are available, for example [33, 34, 35, 36, 37, 38]. The cosmological constant is a function of α , the radius of curvature of de Sitter space, and in D-dimensions given by:

$$\Lambda = \frac{(D-1)(D-2)}{2\alpha^2} \stackrel{(D=4)}{=} \frac{3}{\alpha^2}. \quad (1.35)$$

The explicit form of the metric is most easily obtained by thinking of de Sitter space as a hypersurface embedded in $D+1$ -dimensional Minkowski spacetime. The embedding equation is:

$$-z_0^2 + z_1^2 + \dots + z_D^2 = \alpha^2. \quad (1.36)$$

This makes it manifest that the symmetry group of dS-space is the ten parameter group $SO(1, D)$ of homogeneous ‘Lorentz transformations’ in the D -dimensional embedding space, and the metric is the induced metric from the flat Minkowski metric on the embedding space. In the literature, one encounters several different coordinate systems, which, after quantization, all lead to what appear to be different natural choices for a vacuum state. The metric often used in cosmology is described by coordinates (t, \vec{x}) defined as:

$$\begin{aligned} z_0 &= \alpha \sinh(t/\alpha) + \frac{1}{2}\alpha^{-1}e^{t/\alpha}|\vec{x}|^2 \\ z_4 &= \alpha \cosh(t/\alpha) - \frac{1}{2}\alpha^{-1}e^{t/\alpha}|\vec{x}|^2 \\ z_i &= e^{t/\alpha}x_i, \quad i = 1, 2, 3, \quad -\infty < t, x_i < \infty \end{aligned} \quad (1.37)$$

in which the metric becomes:

$$ds^2 = -dt^2 + e^{\pm 2t/\alpha}d\vec{x}^2. \quad (1.38)$$

covering only half the space, with $z_0 + z_4 > 0$, describing either a universe originating with a big bang, or one ending with a big crunch, depending on the signs. This portion of dS-space is conformally flat. The apparent time-dependence of the metric is just a coordinate artifact. In the absence of any source other than a cosmological constant, there is no preferred notion of time. The translation along the time direction merely slides the point on the surface of the hyperboloid. This time-independence can be made explicit, by choosing so-called static coordinates, defined by:

$$\begin{aligned} z_0 &= (\alpha^2 - r^2)^{\frac{1}{2}} \sinh(t/\alpha) \\ z_1 &= (\alpha^2 - r^2)^{\frac{1}{2}} \cosh(t/\alpha) \\ z_2 &= r \sin \theta \cos \phi \\ z_3 &= r \sin \theta \sin \phi \\ z_4 &= r \cos \theta, \quad 0 \leq r < \infty. \end{aligned} \quad (1.39)$$

Then the metric takes the form:

$$ds^2 = - [1 - (r^2/\alpha^2)] dt^2 + \frac{dr^2}{[1 - (r^2/\alpha^2)]} + r^2(d\theta^2 + \sin^2 \theta d\phi^2). \quad (1.40)$$

These coordinates also cover only half of the de Sitter manifold, with $z_0 + z_1 > 0$. The key feature of the manifold, is that it possesses a coordinate singularity at $r = \alpha = \sqrt{3/\Lambda} = H^{-1}$, with H the Hubble parameter. This represents the event horizon for an observer situated at $r = 0$, following the trajectory of the Killing vector ∂_t (obviously not a global Killing vector, since t is a timelike coordinate only in the region $r < H^{-1}$). In these coordinates, the metric looks very similar to the Schwarzschild metric:

$$ds^2 = - \left(1 - \frac{2M}{r}\right) dt^2 + \frac{dr^2}{\left(1 - \frac{2M}{r}\right)} + r^2(d\theta^2 + \sin^2 \theta d\phi^2), \quad (1.41)$$

describing empty spacetime around a static spherically symmetric source, with a singularity at $r = 2M$, which turns out to be a coordinate artefact, and a ‘real’ singularity at $r = 0$.

In the coordinate system (1.40), the notion of a vacuum state is especially troublesome. The energy-momentum tensor diverges at the event horizon $r = \alpha$, but this horizon is an observer dependent quantity, i.e. it depends on the origin of the radial coordinates. The vacuum state is not even translationally invariant, each observer must be associated with a different choice of vacuum state and comoving observers appear to perceive a bath of thermal radiation [30].

Moreover, like the Bekenstein-Hawking entropy of a black hole, one can assign an entropy S to de Sitter spacetime:

$$S = \frac{A}{4G} \quad (1.42)$$

with A the area of the horizon. Its size, generalizing to D -dimensions again, is given by [34, 35, 36, 37, 38]:

$$A = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{(D-1)(D-2)}{2\Lambda} \right)^{\frac{D-2}{2}} \stackrel{(D=4)}{=} \frac{12\pi}{\Lambda}. \quad (1.43)$$

As noted, however, the horizon in this case is observer-dependent and it is not immediately clear which concepts about black holes carry over to de Sitter space.

One can also derive the Hawking temperature of de Sitter space, by demanding that the metric be regular across the horizon. This yields:

$$T_H = \frac{1}{4\pi} \sqrt{\frac{8\Lambda}{(D-1)(D-2)}} \stackrel{(D=4)}{=} \frac{1}{\pi} \sqrt{\frac{\Lambda}{12}}. \quad (1.44)$$

There are a number of difficulties when one tries to formulate quantum theories in de Sitter space. One is that since there is no globally timelike Killing vector in dS-space, one cannot define the Hamiltonian in the usual way. There is no positive conserved energy in de Sitter space. Consequently, there cannot be unbroken supersymmetry, since, if there were, there would be a non-zero supercharge Q that we can take to be Hermitian⁷ and Q^2 would have to be non-zero, a non-negative bosonic conserved quantity and we arrive at a contradiction. For a long time, this was a serious difficulty for string theory.

Based on the finite entropy, it has been argued that the Hilbert space in de Sitter has finite dimension [39]. This could imply that the standard Einstein-Hilbert action with cosmological constant cannot be quantized for general values of G and Λ , but that it must be derived from a more fundamental theory, which determines these values [34]. Another issue, especially encountered in string theory, is that since the de Sitter symmetry group is non-compact, it cannot act on a finite dimensional Hilbert space [34].

We will return to these interesting issues in much more detail in section (3.3), where we will also discuss arguments from holography. In chapter (5) we will encounter arguments that intend to show that de Sitter space is unstable, leading to a decaying effective cosmological constant.

⁷Possibly after replacing Q by $Q + Q^\dagger$ or $i(Q - Q^\dagger)$.

1.5.1 Weinberg's No-Go Theorem

Another route, often tried to argue that the effective cosmological constant would gradually decay over time, involves a screening effect using the potential of a scalar field ϕ .

Note that in order to make the effective cosmological constant time dependent, this necessarily involves either introducing extra degrees of freedom, or an energy-momentum tensor that is not covariantly constant. This follows directly by taking the covariant derivative on both sides of the Einstein equation. One arrives at:

$$\partial_\mu \Lambda - 8\pi G T_{\mu;\nu}^\nu = 0 \quad (1.45)$$

since the covariant derivative on the Einstein tensor vanishes because of the contracted Bianchi-identity. The energy-momentum tensor is obtained from varying the matter action with respect to $g_{\mu\nu}$, and if general covariance is unbroken, $T_{\mu;\nu}^\nu$ is zero, as a result of the equations of motion. Hence, Λ must be a constant. Therefore, to make the cosmological constant time-dependent, the best one can do is introduce a new dynamical field.

Consider the source of this field to be proportional to the trace of the energy-momentum tensor, or the curvature scalar. Suppose furthermore, that T_μ^μ depends on ϕ , and vanishes at some field value ϕ_0 . Then ϕ will evolve until it reaches its equilibrium value ϕ_0 , where T_μ^μ is zero, and the Einstein equations have a flat space solution. We will consider these approaches in chapter 5. However, on rather general terms Weinberg has derived a no-go theorem, first given in his 1989 review [40], stating that many of these approaches are fatally flawed. We follow the 'derivation' of this no-go theorem as given by Weinberg in [41], see also [42].

He assumes that there will be an equilibrium solution to the field equations in which $g_{\mu\nu}$ and all matter fields ϕ_n are *constant* in spacetime. The field equations are:

$$\partial\mathcal{L}/\partial g_{\mu\nu} = 0 \quad \text{and} \quad \partial\mathcal{L}/\partial\phi_n = 0 \quad (1.46)$$

With N ϕ 's, there are $N+6$ equations for $N+6$ unknowns, since the Bianchi identities remove four of the ten metric coefficients $g_{\mu\nu}$. So one might expect to find a solution without fine-tuning.

The problem Weinberg argues, is in satisfying the trace of the gravitational field equation, which receives a contribution from ρ_{vac} which for $\rho_{vac} \neq 0$ prevents a solution. The trace of the left-hand-side of the Einstein equation is $-R + 4\Lambda$. This contribution of the cosmological constant to the trace of the gravitational field equations, should be cancelled in the screening. Therefore, one tries to make the trace a linear combination of the ϕ_n field equations, as follows:

$$g_{\lambda\nu} \frac{\partial\mathcal{L}(g, \phi)}{\partial g_{\lambda\nu}} = \sum_n \frac{\partial\mathcal{L}(g, \phi)}{\partial\phi_n} f_n(\phi) \quad (1.47)$$

for all constant $g_{\mu\nu}$ and ϕ_n , and $f(\phi)$ arbitrary, except for being finite.

Now, if there is a solution of the the field equation $\partial\mathcal{L}/\partial\phi = 0$ for constant ϕ , then the trace $g_{\mu\nu}\partial\mathcal{L}/\partial g_{\mu\nu} = 0$ of the Einstein field equation for a spacetime-independent metric is also satisfied, despite the fact there is a bare cosmological constant term in the Einstein equation.

However, Weinberg points out that under these assumptions, the Lagrangian has such a simple dependence on ϕ that it is not possible to find a solution of the field equation for ϕ . With the action stationary with respect to variations of all other fields, general covariance and (1.47) imply that the following transformations are a symmetry:

$$\delta g_{\lambda\nu} = 2\epsilon g_{\lambda\nu}, \quad \delta\phi = -\epsilon f(\phi). \quad (1.48)$$

The Lagrangian density for spacetime-independent fields $g_{\mu\nu}$ and ϕ can therefore be written as⁸:

$$\mathcal{L} = c\sqrt{-g} \exp\left(2 \int^{\phi} \frac{d\phi'}{f(\phi')}\right), \quad (1.49)$$

where c is a constant whose value depends on the lower limit chosen for the integral. Only when $c = 0$ is this stationary with respect to ϕ .

Another way to see how this scenario fails, originally due to Polchinski [40], is that the above symmetry $\delta g_{\lambda\nu} = 2\epsilon g_{\lambda\nu}$ ensures that for constant fields, the Lagrangian can depend on $g_{\lambda\nu}$ and ϕ only in the combination $e^{2\phi}g_{\lambda\nu}$, which can be considered as just a coordinate rescaling of the metric $g_{\lambda\nu}$ and therefore cannot have any physical effects. In terms of the new metric $\hat{g}_{\lambda\nu} = e^{2\phi}g_{\lambda\nu}$, ϕ is just a scalar field with only derivative couplings.

Many proposals have been put forward based on such spontaneous adjustment mechanism using one or more scalar fields. However, on closer inspection, they either do not satisfy (1.47), in which case a solution for ϕ does not imply a vanishing vacuum energy, or they do satisfy (1.47), but no solution for ϕ exists. We will return to these types of ‘solutions’ in chapter 5 and see that in many proposals not only the cosmological constant is screened to zero value, but also Newton’s constant. A flat space solution of course always can be obtained if there is no gravity.

This argument can be cast in the form of a no-go theorem. This theorem states that the vacuum energy density cannot be cancelled without fine-tuning in any effective four-dimensional theory that satisfies the following conditions [43]:

1. General Covariance;
2. Conventional four-dimensional gravity is mediated by a *massless* graviton;
3. Theory contains a finite number of fields below the cutoff scale;
4. Theory contains no negative norm states,
5. The fields are assumed to be spacetime independent at late times.

In section (5.4) quantum anomalies to the energy momentum tensor are discussed. These cannot circumvent this no-go theorem.

1.5.2 Some Optimistic Numerology

One often encounters some optimistic numerology trying to relate the value of the effective cosmological constant to other ‘fundamental’ mass scales. For example [44]:

$$\rho_{vac} \sim \left(\frac{M_P}{L}\right)^2 \quad \rightarrow \quad M_\Lambda \sim (M_P M_U)^{\frac{1}{2}}, \quad (1.50)$$

⁸Note the different pre-factor in the exponential.

with $L = M_U^{-1} = 10^{-33}$ eV the size of the universe, and its ‘Compton mass’. The mass scale of the cosmological constant is given by the geometric mean of the UV cutoff M_P and an IR cutoff M_U .

Another relation arises if one includes the scale of supersymmetry breaking M_{SUSY} , and the Planck mass M_P , to the value of the effective cosmological constant [45]. Experiment indicates:

$$M_{susy} \sim M_P \left(\frac{\Lambda}{M_P^2} \right)^\alpha, \quad \text{with } \alpha = \frac{1}{8} \quad (1.51)$$

The standard theoretical result however indicates $M_{susy} \sim \Lambda^{1/2}$.

Another relation often mentioned, see for example [46] is:

$$\Lambda^{1/2} \sim \left(\frac{M_{susy}}{M_P} \right) M_{susy}, \quad (1.52)$$

where it is guessed that the supersymmetry breaking scale is the geometric mean of the vacuum energy and the Planck energy. In both relations, the experimental input is used that $\Lambda^{1/2} \sim 10^{-15} M_{susy}$.

Another non-SUSY numerological match can be given [47], related to the fine-structure constant $\alpha = 1/137$:

$$\rho_\Lambda = \frac{M_P^4}{(2\pi^2)^3} e^{-2/\alpha} \sim 10^{-123} M_P^4. \quad (1.53)$$

Note that this looks very similar to ‘t Hooft’s educated guess originating from gravitational instantons, for the electron mass [48]:

$$Gm_e^2 \cong \left(\alpha\sqrt{2} \right)^{-1} e^{-\pi/4\alpha} \quad (1.54)$$

Whether any of these speculative relations holds, is by no means certain but they may be helpful guides in looking for a solution.

1.6 Outline of this thesis

In this thesis we will consider different scenario’s that have been put forward as possible solutions of the cosmological constant problem. The main objective is to critically compare them and see what their perspectives are, in the hope to get some better idea where to look for a solution. We have identified as many different, credible mechanisms as possible and provided them with a code for future reference. They can roughly be divided in five categories: Fine-tuning, symmetry, back-reaction, violating the equivalence principle and statistical approaches.

In the next chapter we describe the cosmological consequences of the latest experimental results, since they have dramatically changed our view of the universe. This is standard cosmological theory and is intended to demonstrate the large scale implications of a non-zero cosmological constant.

The remaining chapters are subsequently devoted to the above mentioned five different categories of proposals to solve the cosmological constant problem and is largely based

on my paper [49]. We will discuss each of these proposals separately, indicating where the difficulties lie and what the various prospects are. See table (1.2) for a list.

Three approaches are studied in great detail. The first can be found in chapter 4 and is based on my paper written together with Gerard 't Hooft [50], in which we explore a new symmetry based on a transformation to imaginary space. The idea is that the laws of nature have a much wider symmetry than previously expected, and that quantum field theory can be analytically continued to the full complex plane. This generally leads to negative energy states. Positivity of energy arises only after imposing hermiticity and boundary conditions, which opens the way for a vacuum state invariant under these transformations to have zero energy, leading to zero cosmological constant.

The second one is the proposal by Tsamis and Woodard, which we carefully analyze in chapter 6. This scenario is based on a purely quantum gravitational screening of the cosmological constant. However, in our opinion there are several fatal flaws in their arguments.

Thirdly, starting in section 7.3, we carefully study the so-called DGP-gravity model. This string-theory inspired model requires at least three extra infinite volume spatial dimensions in order to solve the cosmological constant problem. As a result, general relativity is modified at both very short and very long distances. At first sight this is a very interesting prospect, however, modifying GR without destroying its benefits at distance scales where the theory is tested, is a very difficult task and we will see that also the DGP-model faces serious obstacles. It shows just how difficult it is to modify general relativity.

The conclusions of this work, as well as an outlook on further research that still needs to be done in this field, are given in chapter (9). Finally, there are two appendices, Appendix A gives the definitions and conventions, used throughout this thesis. Appendix B provides a full list of best fit values for the different cosmological parameters.

For reviews on the history of the cosmological constant (problem) and many phenomenological considerations, see [51, 40, 52, 53, 42, 38, 54, 55, 56, 57].

Table 1.2: Classification of different approaches. Each of them can also be thought of as occurring 1) Beyond 4D, or 2) Beyond Quantum Mechanics, or both.

Type 0: Just Finetuning	
Type I: Symmetry; A: Continuous	a) Supersymmetry
	b) Scale invariance
	c) Conformal Symmetry
B: Discrete	d) Imaginary Space
	e) Energy \rightarrow -Energy
	f) Holography
	g) Sub-super-Planckian
	h) Antipodal Symmetry
	i) Duality Transformations
Type II: Back-reaction Mechanism	a) Scalar
	b) Gravitons
	c) Running CC from Renormalization Group
	d) Screening Caused by Trace Anomaly
Type III: Violating Equiv. Principle	a) Non-local Gravity, Massive Gravitons
	b) Ghost Condensation
	c) Fat Gravitons
	d) Composite graviton as Goldst. boson
Type IV: Statistical Approaches	a) Hawking Statistics
	b) Wormholes
	c) Anthropic Principle, Cont.
	d) Anthropic Principle, Discrete

2 | Cosmological Consequences of Non-Zero Λ

Recent experiments may have shown that the cosmological constant is non-zero. In this chapter we will study the cosmological effects of a non-zero cosmological constant, from the introduction of this term by Einstein, to the evolution of the universe as unravelled by these recent results. They indicate that the universe started a phase of accelerated expansion, about 5 billion years ago. This expansion could very well be driven by a non-zero cosmological constant. We give a brief review of supernovae measurements since these are most important in tracking down the expansion history of the universe.

In this chapter we are concerned with the cosmological effects of a non-zero cosmological constant and assume that general relativity is the correct theory of gravity, also at the largest distance scales. Experimental results have led to the introduction, not only of dark energy, which may be due to a non-zero cosmological constant, but also of dark matter. Together these two spurious forms of energy seem to make up roughly 96% of the total energy density of the universe. In chapter (7) we discuss modifications of GR, intended to explain the observed phenomena without introducing any new form of matter or energy density.

2.1 The Expanding Universe

The cosmological constant was introduced by Einstein in 1917 when he first applied his equations of GR to cosmology, assuming that the universe even at the largest scales can be described by these equations. Without a cosmological constant, they read:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = -8\pi GT_{\mu\nu} \quad (2.1)$$

Furthermore, at large scales, larger than a few hundred Megaparsecs, the universe appears to be very homogeneous and isotropic; our position in the universe seems in no way exceptional. This observation is known as the ‘cosmological principle’ and can be formulated more precisely as follows [58]:

1. The hypersurfaces with constant comoving time coordinate are maximally symmetric subspaces of the whole of spacetime.
2. Not only the metric $g_{\mu\nu}$, but all cosmic tensors, such as the energy-momentum tensor $T_{\mu\nu}$, are form-invariant with respect to the isometries of these subspaces.

To see that this formal definition says the same, consider a general coordinate transformation:

$$x^\mu \rightarrow \tilde{x}^\mu = x^\mu + \xi^\mu(x), \quad (2.2)$$

under which the metric transforms as:

$$\delta g_{\mu\nu} = -(D_\mu \xi_\nu + D_\nu \xi_\mu), \quad (2.3)$$

where D_μ is the covariant derivative. Therefore, if the vector ξ^μ satisfies:

$$D_\mu \xi_\nu + D_\nu \xi_\mu = 0, \quad (2.4)$$

the metric is unchanged by the coordinate transformation, which is then called an isometry and the associated vector ξ^μ is called a Killing vector. A space that admits the maximum number of Killing vectors, given by $d(d+1)/2$ in d dimensions, is called a maximally symmetric space.

A tensor is called form-invariant if the transformed tensor is the same function of \tilde{x}^μ as the original tensor was of x^μ . Specifically, the metric tensor is form-invariant under an isometry and a tensor is called maximally form-invariant if it is form-invariant under all isometries of a maximally symmetric space.

The above mathematical definition of the cosmological principle therefore first states that the universe is spatially homogeneous and isotropic and secondly that our space-time position is in no way special since cosmic observables are invariant under isometries like translations.

Note that both homogeneity and isotropy are symmetries only of space, not of space-time. Homogeneous, isotropic spacetimes have a family of preferred three-dimensional spatial slices on which the three dimensional geometry is homogeneous and isotropic. In particular, these solutions are not Lorentz invariant.

Assuming the cosmological principle to be correct, and ignoring local fluctuations, the metric takes the Robertson-Walker form:

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (2.5)$$

where $a(t)$, called the scale-factor, characterizes the relative size of the spatial sections, and k is the curvature parameter. Coordinates can always be chosen such, that k takes on the value -1, 0 or +1, indicating respectively negatively curved, flat or positively curved spatial sections. Robertson-Walker metrics are the most general homogeneous, isotropic metrics one can write down. They are called Friedman-Robertson-Walker (FRW) metrics if the scale factor obeys the Einstein equation. If the scale factor increases in time this line element describes an expanding universe.

Matter and energy in FRW-model is modelled as a perfect cosmological fluid, with an energy-density ρ and a pressure p . An individual galaxy behaves as a particle in this fluid with zero velocity, since otherwise it would establish a preferred direction, in contradiction with the assumption of isotropy. The coordinates are comoving, an individual galaxy has the same coordinates at all times.

The energy-momentum tensor for a perfect fluid is:

$$T_{\mu\nu} = (\rho + p)U_\mu U_\nu + pg_{\mu\nu}, \quad (2.6)$$

where U_μ is the fluid four-velocity. The rest-frame of the fluid must coincide with a comoving observer in the FRW-metric and in that case, the Einstein equations (2.1)

reduce to the two Friedmann equations. The (00)-component gives:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \quad (2.7)$$

where H is called the Hubble parameter. Using conservation of energy-momentum, we derive:

$$T_{;\nu}^{\mu\nu} = 0 \quad \Rightarrow \quad \frac{d}{dt}(\rho a^3) + p \frac{d}{dt}(a^3) = 0 \quad \Rightarrow \quad \dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0, \quad (2.8)$$

where the first equation is direct consequence of the Bianchi-identity, we find:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + p) \quad (2.9)$$

Einstein was interested in finding static solutions, with $\dot{a} = 0$, since the universe was assumed to be static at the time. One reason for this belief was that the relative velocities of the stars were known to be very small. Furthermore, he assumed that space is both finite and globally closed, as he believed this was the only way to incorporate Mach's principle stating that the metric field should be completely determined by the energy-momentum tensor. A static universe with $\dot{a} = 0$ and positive energy density is compatible with (2.7) if the spatial curvature is negative and the density is appropriately tuned. However, (2.9) implies that \ddot{a} will never vanish in such a spacetime if the pressure p is also non-negative, which is indeed true for most forms of matter such as stars and gas. However, Einstein realized that mathematically his equations allowed an extra term, that could become important at very large distances. It is invisible locally, and therefore also not noticed earlier, by for example Newton. So he proposed a modification to:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = -8\pi GT_{\mu\nu} \quad (2.10)$$

where Λ is a new free parameter, the cosmological constant, and is interpreted as the curvature of empty spacetime. The left-hand-side of (2.10) now indeed is the most general local, coordinate-invariant, divergenceless, symmetric two-index tensor one can construct solely from the metric and its first and second derivatives. In other words, mathematically there was no reason not to put it there right from the start. With this modification, the two Friedmann equations become:

$$\begin{aligned} H^2 &= \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + p) + \frac{\Lambda}{3}. \end{aligned} \quad (2.11)$$

These equations now do admit a static solution with both \dot{a} and \ddot{a} equal to zero, and all parameters ρ , p and Λ non-negative. This solution is called the 'Einstein static universe'.

However, it is not a stable solution; any slight departure of any of the terms from their balanced equilibrium value, leads to a rapid runaway solution. Therefore, even with a cosmological constant a genuinely stationary universe cannot be a solution of the Einstein equations. From the first moment on, the acceptance of the cosmological constant and its physical implications have been the topic of many discussions, which continue till this day.

2.1.1 Some Historic Objections

Already in the same year 1917, De Sitter found that with a cosmological constant, an expanding cosmological model as a solution of Einstein's equations could be obtained, which is 'anti-Machian'. This model universe contains no matter at all.

At about the same time Slipher had observed that most galaxies show redshifts of up to 6%, whereas only a few show blueshifts [59]. However, the idea of an expanding universe was accepted much later, only after about 1930, despite the breakthrough papers of Friedmann in 1922 and 1924 [60, 61] and Lemaitre in 1927. Einstein also found this hard to swallow, according to Lemaitre, Einstein was telling him at the Solvay conference in 1927: "Vos calculs sont correct, mais votre physique est abominable" [5]. Moreover, in order to model an expanding universe, one does not need a cosmological constant.

Even Hubble's stunning discovery in 1924 at first did not really change this picture, for he also did not interpret his data as evidence for an expanding universe. It was Lemaitre's interpretation of Hubble's results that finally changed the paradigm and at this point Einstein rejected the cosmological constant as superfluous and no longer justified [62]. Rumors go that Einstein rejected the CC term in his equations calling it the biggest blunder in his life¹ where he might have been referring to the missed opportunity of predicting the expansion of the universe. There are however also indications that Einstein already had doubts at an earlier stage. A postcard has been found from 1923 where Einstein writes to Weyl: "If there is no quasi-static world, then away with the cosmological term" [5].

It should be noted, that there was a problem interpreting Hubble's data as evidence for an expanding universe, the so-called age problem. The age of the universe derived from Hubble's distance-redshift relation² was a mere two billion years, which clearly cannot be correct, since already the Earth itself is older. For some, for example Eddington, this was reason to keep the cosmological constant alive. For a detailed history of the cosmological constant problem, see [5].

This history of acceptance and refusal, of struggling to understand the constituents and the evolution of the universe, goes on to this very day and especially concerns the role and interpretation of the cosmological constant. There were some major turning points on the road, some of them we will encounter in this thesis.

2.2 Some Characteristics of FRW Models

For many details on cosmology, one can check for example [64, 65, 66, 67]; we especially use the lectures by Garcia-Bellido. In these sections we will review those issues that are most important to us.

To find explicit solutions to the Friedmann equations (2.11), we need to know the matter and energy content of the universe and how they evolve with time. Recall eqn. (2.8):

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0, \quad (2.12)$$

¹Actually there is only one reference about this, by Gamow [63] referring to a private conversation.

²See section (2.2.1) for more details on this.

an expression for the total energy density and pressure. For the individual components i , one has:

$$\dot{\rho}_i + 3\frac{\dot{a}}{a}(\rho_i + p_i) = X_i, \quad \sum_i X_i = 0, \quad (2.13)$$

with X_i a measure of the interactions between the different components. For most purposes in cosmology, the interaction can be set to zero and we can explicitly solve (2.13). To do so, we need a relation between ρ_i and p_i , known as the equation of state. The most relevant fluids in cosmology are barotropic, i.e. fluids whose pressure is linearly proportional to the density: $p_i = w_i\rho_i$, with w_i a constant, the equation of state parameter. In these fluids the speed of sound $c_s^2 = dp/d\rho$ is constant. The solution to (2.13) now becomes:

$$\rho_i = \rho_0 a^{-3(1+w_i)} \quad (2.14)$$

where ρ_0 is an integration constant, set equal to ρ_i , when $a^{-3(1+w_i)} = 1$. Furthermore, it should be noted that we assume here that there is no interaction between the different components ρ_i .

In cosmology the number of barotropic fluids is often restricted to only three:

- Radiation: $w = 1/3$, associated with relativistic degrees of freedom, kinetic energy much greater than the mass energy. Radiation energy density decays as $\rho_R \sim a^{-4}$ with the expansion of the universe.
- Matter: $w = 0$, associated with non-relativistic degrees of freedom, energy density is the matter energy density. It decays as $\rho_M \sim a^{-3}$. Also called 'dust'.
- Vacuum energy: $w = -1$, associated with energy density represented by a cosmological constant. Due to this peculiar equation of state, vacuum energy remains constant throughout the expansion of the universe.

From the Friedmann equations it can be seen that an accelerating universe $\ddot{a} > 0$ is possible, not only for non-zero cosmological constant $w = -1$, but more generally for 'fluids' with $w < -1/3$. Fluids satisfying $\rho + 3p \geq 0$, or $w \geq -1/3$ are said to satisfy the 'strong energy condition' (SEC). Dark energy thus violates this SEC. The 'weak energy condition' (WEC), is satisfied if $\rho + p \geq 0$, or $w \geq -1$. This condition is usually assumed to hold at all times, but recently been called into question in so-called 'phantom dark energy' models, see [68, 69, 70, 71]. The effective speed of sound in such a medium $v = \sqrt{|dp/d\rho|}$ can become larger than the speed of light. A universe dominated by phantom energy has some bizarre properties. For example, it culminates in a future curvature singularity ('Big Rip'). Models constructed simply with a wrong sign kinetic term, are plagued with instabilities at the quantum level, but it has been argued that braneworld models can be constructed devoid of these troubles [72, 73].

From the Friedmann equations (2.11), we can define a critical energy density ρ_c that corresponds to a flat universe:

$$\rho_c \equiv \frac{3H_0^2}{8\pi G}, \quad (2.15)$$

where the subscript 0 denotes parameters measured at the present time. In terms of this critical density, we can rewrite the first Friedmann equation of (2.11) in terms

of density parameters $\Omega_i \equiv \rho_i/\rho_c$ where the subscript i runs over all possible energy sources. For matter, radiation, cosmological constant and curvature, these are:

$$\begin{aligned}\Omega_M &= \frac{8\pi G}{3H_0^2}\rho_M & \Omega_R &= \frac{8\pi G}{3H_0^2}\rho_R \\ \Omega_\Lambda &= \frac{\Lambda}{3H_0^2} & \Omega_k &= -\frac{k}{a_0^2 H_0^2}.\end{aligned}\quad (2.16)$$

With these definitions the Friedmann equation can be written as:

$$H^2(a) = H_0^2 \left(\Omega_R \frac{a_0^4}{a^4} + \Omega_M \frac{a_0^3}{a^3} + \Omega_\Lambda + \Omega_k \frac{a_0^2}{a^2} \right), \quad (2.17)$$

and therefore, the Friedmann equation today ($a = a_0$) becomes:

$$1 = \Omega_M + \Omega_R + \Omega_\Lambda + \Omega_k \quad (2.18)$$

known as the ‘‘cosmic sum rule’’. Sometimes, a dimensionless scalefactor $R(t)$ is defined: $R(t) \equiv a(t)/a_0$, such that $R(t) = 1$ at present to make manipulations with the above expression a little easier.

The energy density in radiation nowadays is mainly contained in the density of photons from the cosmic microwave background radiation (CMBR):

$$\begin{aligned}\rho_{CMBR} &= \pi^2 k^4 T_{CMBR}^4 / (15\hbar^3 c^3) = 4.5 \times 10^{-34} \text{ g/cm}^3, \\ \Omega_{R,CMBR} &= 2.4 \times 10^{-5} h^{-2}\end{aligned}\quad (2.19)$$

where $h \sim 0.72$ is the dimensionless Hubble parameter, defined in Appendix A. Three approximately massless neutrinos would contribute a similar amount, whereas the contribution of possible gravitational radiation would be much less. Therefore, we can safely neglect the contribution of Ω_R to the total energy density of the universe today. Moreover, CMBR measurements indicate that the universe is spatially flat to a high degree of precision, which means that Ω_k is also negligibly small.

From the second Friedmann equation, we can define another important quantity, the deceleration parameter q , defined as follows:

$$q \equiv -\frac{a\ddot{a}}{\dot{a}^2} = \frac{4\pi G}{3H^2}(\rho + 3p) - \frac{\Lambda}{3H^2}. \quad (2.20)$$

This shows that when vacuum energy is the dominant energy contribution in the universe, the deceleration parameter is negative, indicating an accelerated expansion, whereas it is positive for matter dominance. Uniform expansion corresponds to the case $q = 0$. In terms of the Ω_i , the deceleration parameter today can be written as:

$$q_0 = \Omega_R + \frac{1}{2}\Omega_M - \Omega_\Lambda + \frac{1}{2}\sum_x (1 + 3w_x)\Omega_x, \quad (2.21)$$

where we have included the option of possible other fluids, with deviating equations of state parameters w_x . Recent measurements indicate that q_0 is about -0.6 , indicating that the expansion of the universe is accelerating. Astronomers actually measure this quantity by making use of a different relation, where as a time variable the redshift, denoted by z , is used.

2.2.1 Redshift

Since FRW models are time dependent, the energy of a particle will change as it moves through this geometry, similarly to moving in a time dependent potential. The trajectory of a particle moving in a gravitational field obeys the following geodesic equation:

$$\frac{du^\mu}{d\lambda} + \Gamma_{\nu\alpha}^\mu u^\nu u^\alpha = 0, \quad (2.22)$$

where $u^\mu \equiv dx^\mu/ds$ and λ is some affine parameter, that we can choose to be the proper length $g_{\mu\nu}dx^\mu dx^\nu$. The $\mu = 0$ component of the geodesic equation then is very simple, since the only non-vanishing $\mu = 0$ Christoffel is $\Gamma_{ij}^0 = (\dot{a}/a)g_{ij}$. Also using that $g_{ij}u^i u^j = |\vec{u}|^2$ we have:

$$\frac{du^0}{ds} + \frac{\dot{a}}{a}|\vec{u}|^2 = 0 \quad \Rightarrow \quad \frac{1}{u^0} \frac{d|\vec{u}|}{ds} + \frac{\dot{a}}{a}|\vec{u}| = 0, \quad (2.23)$$

and since $u^0 = dt/ds$ this reduces to:

$$\frac{|\dot{\vec{u}}|}{|\vec{u}|} = -\frac{\dot{a}}{a} \quad \Rightarrow \quad |\vec{u}| \propto \frac{1}{a}. \quad (2.24)$$

Moreover, $p^\mu = mu^\mu$, therefore the 3-momentum of a freely propagating particle redshifts as $1/a$. The factors ds in (2.23) cancel, so this also applies to massless particles for which ds is zero. This can be derived in a similar way, choosing a different affine parameter.

This momentum-redshift can also be derived by specifying a particular metric and writing down the Hamilton-Jacobi equations, see [65].

In quantum mechanics, the momentum of a photon is inversely proportional to the wavelength of the radiation, thus a similar shift occurs. For a photon this results in a redshift of its wavelength, hence the name. Photons travel on null geodesics of zero proper time and thus travel between emission time t_e and observation time t_o a distance R , given by:

$$ds^2 = 0 = -dt^2 + a^2(t)dr^2 \quad \Rightarrow \quad R = \int_{t_e}^{t_o} \frac{dt}{a(t)}. \quad (2.25)$$

Furthermore, since R is a comoving quantity, changing the upper and lower limits to account for photons emitted, and observed, at later times, does not affect the result. In other words, $dt_e/dt_o = a(t_e)/a(t_o)$. The redshift is then simply defined as:

$$1 + z \equiv \frac{\lambda_{obs}}{\lambda_{emit}} = \frac{\nu_{emit}}{\nu_{obs}}. \quad (2.26)$$

In practice, astronomers observe light emitted from objects at large distances and compare their spectra with similar ones known in their restframe. This can be done since our galaxy is a gravitationally bound object that has decoupled from the expansion of the universe. The distance between galaxies changes with time, not the sizes of galaxies, or measuring rods within them. Stars move with respect to their local environments at typical velocities of about 10^{-3} times the speed of light [74], leading to a special relativistic Doppler redshift of about $\Delta z \sim 10^{-3}$. If spacetime were not expanding, this

would be the only source of non-zero redshift, and averaging over many stars at the same luminosity distance would give zero redshift. This is precisely what happens for stars within our galaxy, but changes for stars further away.

In a similar manner we can calculate the distance to some far away object. The assumptions of homogeneity and isotropy give us the freedom to choose our position at the origin of our spatial section and to ignore the angular coordinates. In a general FRW-geometry we thus have:

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = \int_0^{r_1} \frac{dr}{\sqrt{1 - kr^2}} \equiv f(r_1) = \begin{cases} \arcsin r_1 & k = 1 \\ r_1 & k = 0 \\ \operatorname{arcsinh} r_1 & k = -1 \end{cases} \quad (2.27)$$

If we now Taylor expand the scale factor to third order [65, 64, 66, 67],

$$\frac{a(t)}{a_0} = 1 + H_0(t - t_0) - \frac{q_0}{2!}H_0^2(t - t_0)^2 + \frac{j_0}{3!}H_0^3(t - t_0)^3 + \mathcal{O}(t - t_0)^4, \quad (2.28)$$

where, using the second Friedmann equation:

$$\begin{aligned} q_0 &= -\frac{\ddot{a}}{aH^2}(t_0) = \frac{1}{2} \sum_i (1 + 3w_i)\Omega_i = \frac{1}{2}\Omega_M - \Omega_\Lambda \\ j_0 &= +\frac{\dddot{a}}{aH^3}(t_0) = \frac{1}{2} \sum_i (1 + 3w_i)(2 + 3w_i)\Omega_i = \Omega_M + \Omega_\Lambda, \end{aligned} \quad (2.29)$$

where we have set Ω_R and Ω_k to zero. We find, to first approximation,

$$r_1 \approx f(r_1) = \frac{1}{a_0}(t_0 - t_1) + \dots = \frac{z}{a_0H_0} + \dots \quad (2.30)$$

This yields the famous Hubble law:

$$H_0 d = a_0H_0r_1 = z \simeq vc, \quad (2.31)$$

Astronomers track the expansion history of the universe by plotting redshift z versus a quantity d_L , the luminosity distance. This luminosity distance is defined as the distance at which a source of absolute luminosity \mathcal{L} gives a flux $\mathcal{F} = \mathcal{L}/4\pi d_L^2$. The expression for d_L as a function of z for a FRW-metric, returning to general Ω_k , but keeping $\Omega_R = 0$, is [65, 64, 66, 67]:

$$H_0 d_L(z) = \frac{1 + z}{|\Omega_k|^{1/2}} \operatorname{sinn} \left[\int_0^z \frac{|\Omega_k|^{1/2} dz'}{\sqrt{(1 + z')^2(1 + z'\Omega_M) - z'(2 + z')\Omega_\Lambda}} \right], \quad (2.32)$$

where $\operatorname{sinn}(x) = x$ if $k = 0$; $\sin(x)$ if $k = +1$ and $\sinh(x)$ if $k = -1$, and we have used the cosmic sum rule (2.18). Substituting eqn's (2.29) into Eqn. (2.32) we find:

$$H_0 d_L(z) = z + \frac{1}{2}(1 - q_0)z^2 - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^3 + \mathcal{O}(z^4). \quad (2.33)$$

The leading term yields Hubble's law, which is just a kinematical law, whereas the higher order terms are sensitive to the cosmological parameters Ω_Λ and Ω_M .

An interesting point in the evolution of the universe was the transition between deceleration and acceleration. This transition must have occurred since the early universe had to be matter dominated in order to form structure. Only recently clear evidence of the existence of this transition, or *coasting point* has been obtained from supernovae data [17]. The coasting point is defined as the time, or redshift, at which the deceleration parameter vanishes,

$$q(z) = -1 + (1+z) \frac{d}{dz} \ln H(z) = 0, \quad (2.34)$$

using $q = -1 - d/dt H(a)$, where

$$H(z) = H_0 \left[\Omega_M (1+z)^3 + \Omega_x e^{3 \int_0^z (1+w_x(z')) \frac{dz'}{1+z'}} + \Omega_k (1+z)^2 \right]^{1/2}, \quad (2.35)$$

from eqn. (2.17) and using that $a_0/a = 1+z$, and we have assumed that the dark energy is parameterized by a density Ω_x today, with a redshift-dependent equation of state, $w_x(z)$, not necessarily equal to -1 .

Assuming that w is constant, the coasting redshift can be determined from

$$q(z) = \frac{1}{2} \left[\frac{\Omega_M + (1+3w) \Omega_x (1+z)^{3w}}{\Omega_M + \Omega_x (1+z)^{3w} + \Omega_k (1+z)^{-1}} \right] = 0, \quad (2.36)$$

$$\Rightarrow z_c = \left(\frac{(3|w| - 1) \Omega_x}{\Omega_M} \right)^{\frac{1}{3|w|}} - 1, \quad (2.37)$$

which, in the case of a true cosmological constant, reduces to

$$z_c = \left(\frac{2\Omega_\Lambda}{\Omega_M} \right)^{1/3} - 1. \quad (2.38)$$

When substituting $\Omega_\Lambda \simeq 0.7$ and $\Omega_M \simeq 0.3$, one obtains $z_c \simeq 0.6$, in excellent agreement with recent observations [17]. The plane $(\Omega_M, \Omega_\Lambda)$ can be seen in Fig. (2.2.1), which shows a significant improvement with respect to previous data. The best determination of the Hubble parameter H_0 was made by the Hubble Space Telescope Key Project, [24] to be $H_0 = 72 \pm 8$ km/s/Mpc, based on objects at distances up to 500 Mpc, corresponding to redshifts $z \leq 0.1$. In Appendix B, we provide a full list of the best fit values for the different cosmological parameters.

As a nice example, we can calculate the effect of a cosmological constant term, on our motion in the Milky Way [54]. Using the non-relativistic limit, we have:

$$\frac{d^2 \vec{r}}{dt^2} = \vec{g} + \Omega_\Lambda H_0^2 \vec{r}, \quad (2.39)$$

where \vec{g} is the relative gravitational acceleration produced by the distribution of ordinary matter. Our solar system is moving in the Milky Way galaxy at speed roughly $v_c = 220$ km/s at radius $r = 8$ kpc. The ratio of the acceleration g_Λ produced by the cosmological constant, to the total gravitational acceleration $g = v_c^2/r$ is:

$$g_\Lambda/g = \Omega_\Lambda H_0^2 r^2 / v_c^2 \sim 10^{-5}, \quad (2.40)$$

a small number. The precision of celestial dynamics in the solar system is much better, but of course, the effect of a cosmological constant is much smaller, $g_\Lambda/g \sim 10^{-22}$.

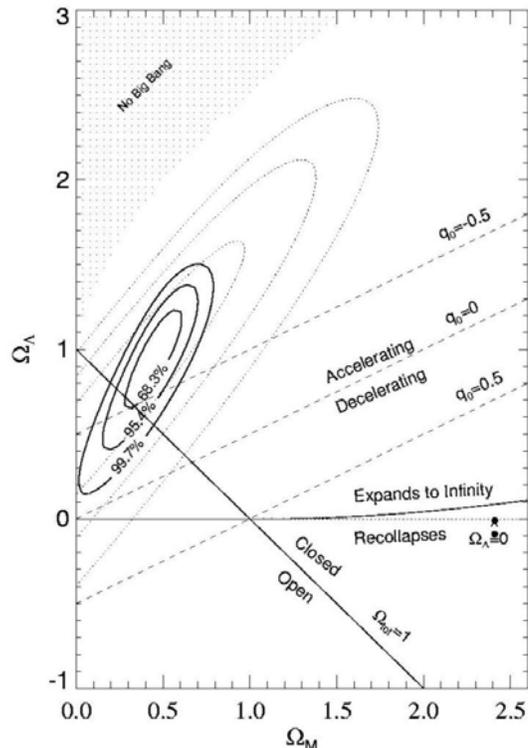


Figure 2.1: The recent supernovae data on the $(\Omega_M, \Omega_\Lambda)$ plane. Shown are the 1-, 2- and 3- σ contours, as well as the data from 1998, for comparison. It is clear that the old EdS cosmological model at $(\Omega_M = 1, \Omega_\Lambda = 0)$ is many standard deviations away from the data. From Ref. [12].

2.3 The Early Universe

We have already used the value of several cosmological parameters in the above calculations. An important source of information for these values comes from the Cosmic Microwave Background Radiation, or CMBR.

In table (2.1) we list the major events that took place in the history of the universe, based on the inflationary big bang model. Important for us are the end of inflation at $\sim 10^{12}$ GeV, after which the effective cosmological constant was very small, and the origin of the CMBR, since many cosmological observations that are the backbone of our theoretical models, originate from it.

When the CMBR was discovered in 1965, its temperature was found to be 2.725 Kelvins, no matter which direction of the sky one looks at; it is the best blackbody spectrum ever measured. The CMBR appeared isotropic, indicating that the universe was very uniform at those early times.

It can be considered as the “afterglow” of the Big Bang. After the recombination of protons and electrons into neutral hydrogen, about 380,000 years after the Big Bang, the mean free path of the photons became larger than the horizon size, and the universe became transparent for the photons produced in the earlier phases of the evolution of the universe. This radiation therefore provides a snapshot of the universe at that time.

The collection of points where the photons of the CMBR that are now arriving on earth had their last scattering before the universe became transparent, is called the last-scattering surface.

The universe just before recombination is a tightly coupled fluid, where photons scatter off charged particles and since they carry energy, they feel perturbations imprinted in the metric during inflation. These small perturbations propagate very similar to sound waves, a train of slight compressions and rarefactions. The compressions heated the gas, while the rarefactions cooled it, leading to a shifting pattern of hot and cold spots, the temperature anisotropies. A distinction is made between primary and secondary anisotropies, the first arise due to effects at the time of recombination, whereas the latter are generated by scattering along the line of sight. There are three basic primary perturbations, important on respectively large, intermediate, and small angular scales (see [64, 75, 76] for many details on the CMBR and its anisotropies):

1. Gravitational Sachs-Wolfe. Photons from high density regions at last scattering have to climb out of potential wells, and are redshifted: $\delta T/T = \delta\Phi/c^2$, with $\delta\Phi$ the (perturbations in the) gravitational potential. These perturbations also cause a time dilation at the surface of last scattering, so these photons appear to come from a younger, hotter universe: $\delta T/T = -2\delta\Phi/3c^2$, so the combined effect is $\delta T/T = \delta\Phi/3c^2$.
2. Intrinsic, adiabatic. Recombination occurs later in regions of higher density, causing photons coming from overly dense regions to have smaller redshift from the universal expansion, and so appear hotter: $\delta T/T = -\delta z/(1+z) = \delta\rho/\rho$.
3. Velocity, Doppler. The plasma has a certain velocity at recombination, leading to Doppler shifts in frequency and hence temperature: $\delta T/T = \delta\vec{v} \cdot \hat{r}/c$, with \hat{r} , the direction along the line of sight, and \vec{v} the characteristic velocity of the photons in the plasma.

Through the Cosmic Background Explorer (COBE) and, more recently, the Wilkinson Microwave Anisotropy Probe (WMAP) [18, 19] satellites, the small variations in the radiation's temperature were detected. They are perturbations of about one part in 100,000. These tiny anisotropies are images of temperature fluctuations on the last-scattering surface and contain a wealth of cosmological information. Their angular sizes depend on their physical size at this time of last scattering, but they also depend on the geometry of the universe, through which the light has been travelling before reaching us. Maps of the temperature fluctuations are a picture of this last-scattering surface processed through the geometry and evolution of a FRW model.

During their journey through the universe, a small fraction of the CMBR photons is scattered by hot electrons in gas in clusters of galaxies. These CMBR photons gain energy as a result of this Compton scattering, which is known as the Sunyaev-Zeldovich effect, see for example [77, 78]. It is observed as a deficit of about 0.05 % of CMBR photons, as they have shifted to higher energy, with about 2 % increase. This effect can be seen as a verification of the cosmological origin of the CMBR.

A widely used code to calculate the anisotropies using linear perturbation theory is CMBFAST [79].

2.3.1 Deriving Geometric Information from CMBR Anisotropies

When looking at the CMBR, we are observing a projection of soundwaves onto the sky. A particular mode with wavelength λ , subtends an angle θ on the sky. The observed spectrum of CMB anisotropies is mapped as the magnitude of the temperature variations, versus the sizes of the hot and cold spots, and this is usually plotted through a multipole expansion in Legendre polynomials $P_l(\cos \theta)$, of a correlation function $C(\theta)$. The order l of the polynomial, related to the multipole moments, plays a similar role in the angular decomposition as the wavenumber $k \propto 1/\lambda$ does for a Fourier decomposition. Thus the value of l is inversely proportional to the characteristic angular size of the wavemode it describes.

There are many good lectures on CMB physics, e.g. [64, 66, 67, 75, 80, 76] and especially the website by Hu [81].

The correlation function $C(\theta)$, is defined as follows: Let $\Delta T(\vec{n})/T$ be the fractional deviation of the CMBR temperature from its mean value in the direction of a unit vector \vec{n} . Take two vectors \vec{n} and \vec{n}' that make a fixed angle θ with each other: $\vec{n} \cdot \vec{n}' = \cos \theta$. The correlation function $C(\theta)$ is then defined by averaging the product of the two $\Delta T/T$'s over the sky:

$$C(\theta) \equiv \left\langle \frac{\Delta T(\vec{n})}{T} \frac{\Delta T(\vec{n}')}{T} \right\rangle, \quad (2.41)$$

where the angle brackets denote the all-sky average over \vec{n} and \vec{n}' and it is assumed that the fluctuations are Gaussian³. In terms of Legendre polynomials:

$$C(\theta) = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} C_l P_l(\cos \theta). \quad (2.42)$$

Modes caught at extrema of their oscillations become the peaks in the CMB power spectrum. They form a harmonic series based on the distance sound can travel by recombination, called the sound horizon. The first peak represents the fundamental wave of the universe, and represents the mode that compressed once inside potential wells before recombination, the second the mode that compressed and then rarefied, the third the mode that compressed then rarefied then compressed, etc. These subsequent peaks in the power spectrum represent the temperature variations caused by the overtones. All peaks have nearly the same amplitude, as predicted by inflation, except for a sharp drop-off after the third peak. The physical scale of these fluctuations is so small that they are comparable to the distance photons travel during recombination. Recombination does not occur instantaneously, the surface of last scattering has a certain thickness. In that short period during which the universe recombines, the photons bounce around the baryons and execute a random walk. If the random walk takes the photons across a wavelength of the perturbation, then the hot and cold photons mix and average out. The acoustic oscillations are exponentially damped on scales smaller than the distance photons randomly walk during recombination. See figure (2.3).

From the position of the peaks we can infer information about the geometry of the universe. In the same way as the angle subtended by say the planet Jupiter depends on

³If the n -point distribution is Gaussian, it is defined by its mean vector $\langle \delta(\vec{x}) \rangle$, which is identically zero, and its covariance matrix $C_{mn} \equiv \langle \delta(\vec{x}_m) \delta(\vec{x}_n) \rangle = \xi(|\vec{x}_m - \vec{x}_n|)$. $\xi(x)$ is a correlation function: $\langle \delta(\vec{x}_2) \delta(\vec{x}_1) \rangle$, that because of homogeneity and isotropy only depends on $x \equiv |\vec{x}_2 - \vec{x}_1|$ [76].

both its size and its distance from us, so depend the angular sizes of the anisotropies on our distance to the surface of last scattering d_{sls} , from which the CMBR originated, and on what is called the “sound horizon” r_s . The sound horizon is given by the distance sound waves could have travelled in the time before recombination, it is a fixed scale at the surface of last scattering. The angular size θ_s of the sound horizon thus becomes:

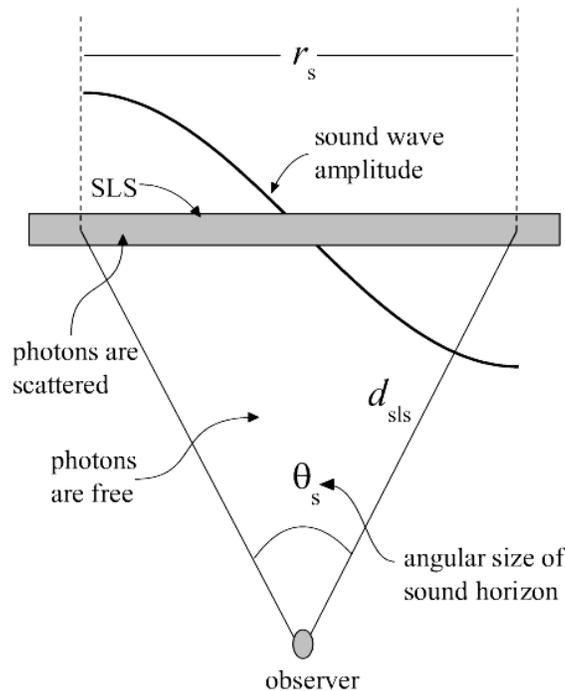
$$\theta_s \approx \frac{r_s}{d_{sls}}. \quad (2.43)$$

The sound horizon r_s and the distance to the surface of last scattering, both depend on the cosmological parameters, Ω_i . The distance sound can travel, from the big bang to the time of recombination is:

$$r_s(z_*, \Omega_i) \approx \int_0^{t_*} c_s dt, \quad (2.44)$$

where z_*, t_* are the redshift and time at recombination. See figure (2.2).

Figure 2.2: Surface of last scattering (SLS), fundamental acoustic mode, and the sound horizon. From [80].



The speed of sound in the baryon-photon plasma is given by [75]:

$$c_s \approx c [3 (1 + 3\Omega_b/4\Omega_R)]^{-1/2}, \quad (2.45)$$

where Ω_b is the density of baryons and c the speed of light.

The distance to the surface of last scattering, corresponding to its angular size, is given by what is called the angular diameter distance. It is proportional to the luminosity

distance d , and explicitly given by:

$$d_{sls} = \frac{d(z_*; \Omega_M; \Omega_\Lambda)}{(1 + z_*)^2}. \quad (2.46)$$

The location of the first peak is given by $l \approx d_{sls}/r_s$ and is most sensitive to the curvature of the universe.

In [80], a very instructive calculation is presented, considering just the first peak of the spectrum, that gives a clear idea how all of this works in practise. We will review that calculation here. Taking the speed of sound (2.45) to leading order equal to $c/\sqrt{3}$ and assuming that the early universe was matter dominated, we obtain for the sound horizon r_s :

$$r_s = \frac{c/\sqrt{3}}{H_0\sqrt{\Omega_M}} \int_{z_*}^{\infty} (1+z)^{-5/2} dz = \frac{2c/\sqrt{3}}{3H_0\sqrt{\Omega_M}} (1+z_*)^{-3/2}. \quad (2.47)$$

The distance to the surface of last scattering d_{sls} , assuming a flat universe i.e. $\Omega_k = 0$, depends only on Ω_M and Ω_Λ . It can be derived, using that $d_{sls} = r_{sls}/(1+z_*)$, where r_{sls} is the radial coordinate of the surface of last scattering. It is given by [80]:

$$r_{sls} = \frac{c}{H_0} \int_0^{\infty} [\Omega_M(1+z)^3 + \Omega_\Lambda]^{-1/2} dz. \quad (2.48)$$

This integral can be solved making use of a binomial expansion approximating the integrand to:

$$r_{sls} = \frac{c}{H_0} \int_0^{\infty} \left(\Omega_M^{-1/2} (1+z)^{-3/2} - \frac{\Omega_\Lambda}{2\Omega_M^{3/2}} (1+z)^{-9/2} \right) dz. \quad (2.49)$$

We obtain:

$$d_{sls} = \frac{2c}{7H_0(1+z_*)} \left(7\Omega_M^{-1/2} - 2\Omega_\Lambda\Omega_M^{-3/2} + \mathcal{O}[(1+z_*)^{-1/2}] \right). \quad (2.50)$$

With our assumption of a flat universe the cosmic sum rule simply becomes $\Omega_\Lambda = 1 - \Omega_M$. Using this, and ignoring higher order terms, we arrive at:

$$d_{sls} \approx \frac{2c\Omega_M^{-1/2}}{7H_0(1+z_*)} (9 - 2\Omega_M^3). \quad (2.51)$$

Together with eqn. (2.47), this gives the prediction for the first acoustic peak in a flat universe, ignoring density in radiation:

$$l \approx \frac{d_{sls}}{r_s} \approx 0.74\sqrt{(1+z_*)} (9 - 2\Omega_M^3) \approx 221. \quad (2.52)$$

Consistent with the more accurate result, for example [82], from the MAXIMA-1 collaboration:

$$l \approx 200/\sqrt{1 - \Omega_k}. \quad (2.53)$$

To summarize, the expansion in multipoles (2.3), shows that for some feature at angular size $\Delta\theta$ in radians, the C_l 's will be enhanced for a value l inversely related to $\Delta\theta$. For a

flat universe, it shows up at lower l , than it would in an open universe, see figure. The dependence of the position of the first peak on the spatial curvature can approximately be given by $l_{peak} \simeq 220 \Omega^{-1/2}$, with $\Omega = \Omega_M + \Omega_\Lambda = 1 - \Omega_k$. With the high precision WMAP data, this lead to $\Omega = 1.02 \pm 0.02$ at 95% confidence level.

Much more information can be gained from the power-spectrum. A very nice place to get a feeling for the dynamics of the power spectrum is Wayne Hu's website [81], whose discussion we also closely follow in the remainder of this section.

The effect of baryons on the CMB power spectrum is threefold:

- 1: The more baryons there are, the more the second peak will be suppressed compared to the first. This results from the fact that the odd numbered peaks are related to how much the plasma is compressed in gravitational potential wells, whereas the even numbered peaks originate from the subsequent rarefaction of the plasma.
- 2: With more baryons, the peaks are pushed to slightly higher multipoles l , since the oscillations in the plasma will decrease.
- 3 Also at higher multipole moments, smaller angular scales, an effect can be seen, due to their effect on how sound waves are damped.

The effects of dark matter are best identified by the higher acoustic peaks, since they are sensitive to the energy density ratio of dark matter to radiation in the universe and the energy density in radiation is fairly well known. Especially the third peak is interesting. If it is higher than the second peak, this indicates dark matter dominance in the plasma, before recombination. The third peak gives the best picture for this, since in the first two peaks the self-gravity of photons and baryons is still important.

Dark matter also changes the location of the peaks, especially the first one. This is because the ratio of matter to radiation determines the age of the universe at recombination,. This in turn limits how far sound could have travelled before recombination relative to how far light travels after recombination. The spatial curvature has a similar effect, so to disentangle the two, at least three peaks have to be measured.

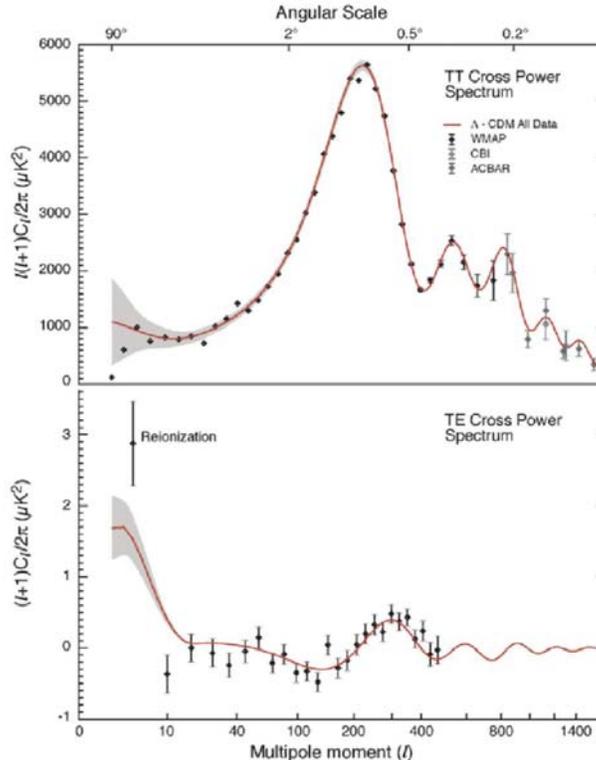
The effect of a positive cosmological constant is a small change in how far light can travel since recombination, and hence produces a slight shift to lower multipoles.

2.3.2 Energy Density in the Universe

There is considerable evidence that most of the mass density of the universe is neither in luminous matter, nor in radiation. Since every mass gravitates, it can be detected by its gravitational influence. Most direct evidence for unseen 'dark matter' comes from weighing spiral galaxies. By measuring the Doppler shifts in the 21-cm line of neutral hydrogen the velocity of clouds of this gas in the disk can be mapped as a function of the distance r from the center of the galaxy. One should expect this velocity to fall off as $r^{-1/2}$, but rather it remains constant in almost all cases, see figure (2.4).

This seems to imply that, even in the outskirts of the galaxy, the amount of mass is still growing with distance. Almost every galaxy seems to contain a 'halo' of dark, unseen matter, the amount of which depends on the kind of galaxy. Ranging from roughly ten

Figure 2.3: The Angular Power Spectrum of CMB temperature anisotropies, from the WMAP satellite. Also shown is the correlation between the temperature anisotropies and the (E-mode) polarization. From Ref. [83].



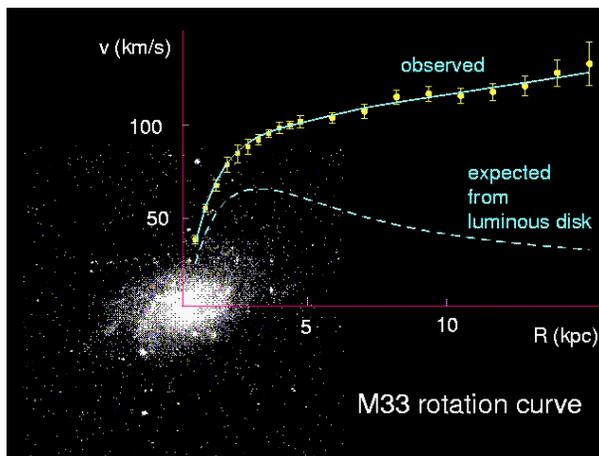
times the mass seen in visible matter in spiral galaxies, to about 100 times as much in low surface brightness galaxies and dwarfs, like the Draco dwarf spheroidal galaxy, which is nearby at only 79 kpc from the Milky Way [84, 85].

This “missing mass” is a central problem in cosmology and speculations about its origin range from black holes to new species of particles, see for example [86] for a recent review. One of the most important dark matter candidates is the lightest supersymmetric particle. From the WMAP data and under certain model restrictions, an upper limit of the mass of the LSP can be derived: $m_{LSP} \leq 500$ GeV [87].

However, there could also be a different explanation as propagated by proponents of the Modified Newtonian Dynamics (MOND) approach [88, 89, 90, 91]. MOND can perfectly describe the rotation curves of individual galaxies, but it is not without problems. Its predictions do not agree with observations of clusters of galaxies [92].

Using measurements of CMB anisotropies, statistical analysis of galaxy redshift surveys and measurements of the number density of massive galaxy clusters, it has been shown that the universe is flat to great precision and that matter, including dark matter, makes up about a quarter of the critical energy density. Combining these results with those of distant supernovae, it is inferred that about 75% of the universe is made up of dark energy. These combined experiments moreover show that the equation of state of this dark energy component is very close to -1. For more details on exactly how the different parameters are measured, see any cosmology book, like [64, 65], or more

Figure 2.4: The observed rotation curve, of the dwarf spiral galaxy M33, extends considerably beyond its optical image (shown superimposed); from [84, 85].



recent reviews like [66, 67].

In the past two decades or so, the picture of the contents of the universe we now have is often abbreviated as Λ CDM. A universe, with dominant cosmological constant, or dark energy and Cold Dark Matter.

2.4 Evolution of the Universe

We will now consider universes dominated by matter and vacuum energy, $\Omega = \Omega_M + \Omega_\Lambda = 1 - \Omega_k$ as given in the previous section. As discussed, a positive cosmological constant tends to accelerate the universal expansion, while ordinary matter and a negative cosmological constant tend to decelerate it. The relative contributions to the energy density of the universe scale like:

$$\Omega_\Lambda \propto a^2 \Omega_k \propto a^3 \Omega_M. \quad (2.54)$$

From the Friedmann equations:

$$\begin{aligned} H^2 &= \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + p) + \frac{\Lambda}{3}. \end{aligned} \quad (2.55)$$

we can immediately infer that, if $\Omega_\Lambda < 0$, the universe will always recollapse to a big crunch, either because the matter density is sufficiently high or because eventually the negative cosmological constant will dominate. For $\Omega_\Lambda > 0$ the universe will expand forever, unless there is sufficient matter to cause a recollapse before Ω_Λ becomes dynamically important. For $\Omega_\Lambda = 0$ on the hand the universe will expand forever if $\Omega_M \leq 1$ or will recollapse if $\Omega_M > 1$. To identify these boundaries properly, let us have look again at the equation for the time-dependent Hubble parameter:

$$H^2(a) = H_0^2 \left(\Omega_R \frac{a_0^4}{a^4} + \Omega_M \frac{a_0^3}{a^3} + \Omega_\Lambda + \Omega_k \frac{a_0^2}{a^2} \right), \quad (2.56)$$

with dimensionless scalefactor $R(t) \equiv a(t)/a_0$ and writing $\Omega_k = -\Omega_M - \Omega_\Lambda + 1$ from the cosmic sum rule, we have:

$$\frac{H^2(a)}{H_0^2} = \Omega_\Lambda(1 - R^{-2}) + \Omega_M(R^{-3} - R^{-2}) + R^{-2}. \quad (2.57)$$

Now we have to find solutions for which the left-hand-side vanishes, since these define the turning points in the expansion. There are four conditions, see for example [93, 64, 66, 67]:

- 1: As already stated before, negative Λ always implies recollapse.
- 2: If Λ is positive and $\Omega_M < 1$, the model always expands to infinity.
- 3: If $\Omega_M > 1$, recollapse is only avoided if Ω_Λ exceeds a critical value:

$$\Omega_\Lambda > 4\Omega_M \cos^3 \left[\frac{1}{3} \cos^{-1} (\Omega_M^{-1} - 1) + \frac{4\pi}{3} \right]. \quad (2.58)$$

- 4: Conversely, if the cosmological constant is sufficiently large compared to the matter density, there was, no initial singularity, no Big Bang. Its early history consisted of a period of gradually slowing contraction to a minimum radius before beginning its current expansion; it underwent a bounce. The criterion for there to have been no singularity in the past is

$$\Omega_\Lambda > 4\Omega_M \text{coss}^3 \left[\frac{1}{3} \text{coss}^{-1} (\Omega_M^{-1} - 1) \right], \quad (2.59)$$

where “coss” represents cosh when $\Omega_M < 1/2$, and cos when $\Omega_M \geq 1/2$.

If the universe lies exactly on the critical line, the bounce is at infinitely early times. Models that almost satisfy the critical relation $\Omega_\Lambda(\Omega_M)$ are called loitering models, since they spent a long time close to constant scale factor. However, these bounce models can be ruled out quite strongly, since the same cubic equations also give a relation for the maximum possible redshift [64]:

$$1 + z_{\text{bounce}} \leq 2 \text{coss} \left(\frac{1}{3} \text{coss}^{-1} (\Omega_M^{-1} - 1) \right). \quad (2.60)$$

With Ω_M as low as 0.1, a bounce is ruled out once objects are seen at redshift $z > 2$. Since galaxies have even been observed at $z = 10$, these universes are indeed ruled out.

This defines the boundaries between the different regions, summarized in the $(\Omega_M, \Omega_\Lambda)$ -plane (2.2.1).

The behavior of the universe can also be seen by referring to our previous expression for the deceleration parameter q :

$$q_0 = \Omega_R + \frac{1}{2}\Omega_M - \Omega_\Lambda + \frac{1}{2} \sum_x (1 + 3w_x)\Omega_x \simeq \frac{1}{2}\Omega_M - \Omega_\Lambda. \quad (2.61)$$

Uniform expansion ($q_0 = 0$) corresponds to the line $\Omega_\Lambda = \Omega_M/2$. Points above this line correspond to universes that are accelerating today, while those below correspond to

decelerating universes, in particular the old cosmological model of Einstein-de Sitter (EdS), with $\Omega_\Lambda = 0$, $\Omega_M = 1$. Since 1998, all the data from Supernovae of type Ia appear above this line, many standard deviations away from EdS universes. Nevertheless, sometimes it is argued, that the EdS universe is a sound alternative to the concordance model, and that it is in better agreement with the low quadrupole in the CMBR [94]. In such a universe, the Hubble constant would be significantly lower, approximately 46 km/s/Mpc, the SNIa results would have to be discarded, for some systematical error for example, and one would have to introduce another source of dark matter, to suppress power on small scales. This seems a very contrived conspiracy of effects, and has not many proponents.

The line $\Omega_\Lambda = 1 - \Omega_M$ corresponds with a flat universe, $\Omega_k = 0$. Points to the right of this line correspond to closed universes, while those to the left correspond to open ones. In the last few years we have mounting evidence that the universe is spatially flat (in fact Euclidean).

2.5 Precision cosmology in recent and upcoming experiments

In the past two decades or so cosmology has really grown as a very interesting field of study. Especially observational cosmology has made astonishing progress producing the first maps of the cosmos and of course showing the acceleration of the expansion of the universe. In this section we will review the supernovae experiments, since these have been so important for the discovery of an accelerating universe, and briefly discuss the Pioneer anomaly, as it has been used to argue for deviations of general relativity.

2.5.1 Supernovae Type Ia

From the end of the 1990's, two independent teams, the ‘‘Supernova Cosmology project’’, led by Perlmutter of Lawrence Berkeley Laboratory and ‘‘The High-Z Supernova Search Team’’, led by Schmidt of Mt. Stromlo and Siding Observatories, have used the apparent brightness of supernovae in order to study the speed of the expansion of the universe. Both teams found that these supernovae look fainter than expected. With the assumption, based on CMBR experiments that the universe is flat, this could be the result of a cosmological constant. This can be easily seen from the expressions for the luminosity distance d_L , which in a matter dominated universe ($\Omega_M = 1, \Omega_\Lambda = 0$) take the form:

$$d_L = 2H_0^{-1} [(1+z) - (1+z)^{1/2}], \quad (2.62)$$

whereas in a cosmological constant dominated universe ($\Omega_M = 0, \Omega_\Lambda = 1$):

$$d_L = H_0^{-1} z(1+z). \quad (2.63)$$

This shows that at given redshift z , the luminosity distance d_L is larger for a universe where the cosmological constant dominates, and hence, a given object, at a fixed redshift, will appear fainter.

For objects close by there is a linear relation between these two quantities and the fixed ratio of the two is known as Hubble's constant. For objects farther away though, deviations from this linear dependence can be expected, either because the speed of the

expansion has changed over time, or as a result of spacetime curvature. Therefore, to track down the history of the expansion one has to find the distance-redshift relation of objects located very far away. In Hubble's day, distances were determined by assuming that all galaxies have the same intrinsic brightness. So the fainter a galaxy really appeared, the further away from us it is located and vice versa. This is a very crude way of determining distances since different galaxies can have very different properties, and therefore different intrinsic brightness. Moreover, when looking at galaxies located far away, evolutionary effects play an important role since light takes so long to travel to us. We then view these galaxies as they were billions of years ago, in their youth, and their intrinsic brightness could be very different from the more mature ones, seen closer at home.

Distances are measured in terms of the "distance modulus" $m - M$, where m is the apparent magnitude of the source, and M its absolute magnitude. This distance modulus is related to the luminosity distance:

$$m - M = 5 \log_{10} [d_L(\text{Mpc})] + 25. \quad (2.64)$$

A very useful relation, if you know the absolute magnitude, which is notoriously hard to infer for a distant object. It is very difficult to disentangle evolutionary changes from the effects of the expansion, so astronomers have been looking for objects called "standard candles", whose intrinsic brightness, and therefore distance, can be determined more unambiguously. The best candidate for this task is a particular type of supernovae (SNe). Supernovae, extremely bright explosions of a dead star, about 10 billion times as luminous as our sun, are among the most violent phenomena in our universe. They come in two main classes [95] called Type II for supernovae whose optical spectra exhibit hydrogen lines, while hydrogen deficient SNe are designated Type I. These SNe Type I are further subdivided according to the detailed appearance of their early-time spectrum. SNe Ia are characterized by strong absorption near 6150 \AA , corresponding to SiII. SNe Ib lack this feature but instead show prominent HeI lines. SNe Ic finally have neither SiII nor HeI lines. In practise there is often little distinction between the latter two types and they are most commonly designated as Type Ib/c.

For the Type II SNe, four subclasses exist based on the shape of optical light curves, see [96] for the details. For our purposes the Type Ia are the only relevant ones.

There are two physical mechanisms that produce supernovae: thermonuclear (the SNe Ia are the only ones of this type) and the more common core collapse of a massive star. This last type occurs when a massive red supergiant becomes old and produces more and more metals in its core. Stars never fuse elements heavier than iron, since this would cost energy, rather than produce it. Thus iron piles up until this iron core reaches the Chandrasekhar mass, which is about 1.4 solar masses. This is the critical mass below which electron degeneracy pressure can stand up to the gravitational pressure, but above this mass, no such equilibrium is possible and the core collapses. At the extreme high pressure in the core protons and electrons are smashed together to form neutrons and neutrinos, while at the same time the outer layers crash into the core and rebound, sending shockwaves outward. These two effects together cause the entire star outside the core to be blown apart in a huge explosion, a Type II supernova! What is left behind of the core will be either a neutron star or, if it is heavy enough, a black hole.

SNe Ia on the other hand originate from a white dwarf star that accretes matter from

another star orbiting nearby, until the Chandrasekhar limit is reached. This happens for example in a binary system, when a white dwarf can start slurping up matter from a main sequence star that expands into a giant or supergiant. When the critical mass is reached, the star is no longer stable against gravitational collapse: the radius decreases, while the density and temperature increase. The fusion of carbon and oxygen into iron now occurs very rapidly, converting the star to a fusion bomb until a thermonuclear firestorm ignites. The dwarf star gets completely blown apart, spewing out material at about 10,000 kilometers per second. The amount of energy released in this cosmic showpiece is about 10^{44} joules, as much as the Sun has radiated away during its entire lifetime. The glow of this fireball takes about three weeks to reach its maximum brightness and then declines over a period of months. The spectrum contains no hydrogen or helium lines, since the dwarf that is blown apart consists of carbon and oxygen. Yet it does show silicon lines, since that is one of the products of fusing carbon and oxygen.

In a typical galaxy such a SNe 1a lights up every 300 years or so. In our own Milky Way therefore, it is a rare celestial event. The last supernova in our galaxy was seen in 1604 by Kepler. However, if you monitor a few thousand galaxies, you can expect to witness about one type 1a supernova every month. There are actually so many galaxies in the universe that a supernova bright enough to study erupts every few seconds! You only have to find them. Catching a supernova at its peak brightness however is not so easy as their occurrence cannot be predicted and observing time at the worlds largest telescopes therefore cannot be scheduled in advance. Moreover, these type 1a also vary in their brilliance, but brighter explosions last somewhat longer than fainter ones. Therefore, they must be very carefully observed multiple times during the first weeks, to monitor how long they last. This way their intrinsic brightness can be deduced to within 12 percent (absolute magnitude $M \sim -19.5$). In terms of redshift, supernovae with redshifts as high as $z \sim 1.5$ have been observed. This corresponds to a time when the universe was about one-third its current age⁴.

The supernova results show that although the universe is undergoing accelerated expansion now, this has not always been the case. Up until about 5 billion years ago, the universe was matter dominated and decelerating. The change between these two phases occurred at about $z \sim 0.5$. For a recent review of type Ia supernovae in cosmology, see [97].

2.5.2 Pioneer Anomaly

Another phenomenon that perhaps should be explained as a deviation of GR originates from within our solar system and has become known as the Pioneer anomaly. The two space probes Pioneer 10 and 11 were launched in the beginning of the seventies to do measurements on Jupiter and Saturn along the outskirts of our solar system. At distances r from the sun between 20 and 70 astronomical units (AU), the Doppler data have shown a deviation from calculations based on GR. Analysis indicated a small, constant Doppler blue shift drift of order 6×10^{-9} Hz/s. After accounting for systematics, this may be interpreted as an unexplained, constant acceleration directed towards the sun with a roughly constant amplitude: $a_P \simeq 8 \times 10^{-10} \text{ms}^{-2}$, or perhaps a time acceleration of $a_t = (2.92 \pm 0.44) \times 10^{-18} \text{s/s}^2$, see e.g. [98] for a recent review and

⁴That is, for a Hubble constant $H_0 = 72 \text{ km/s/Mpc}$ leading to a current age of the universe of about 14 billion years.

many references.

Various mechanisms have been considered to explain this acceleration, but so far no satisfactory scenario has been put forward see [99, 100, 101, 102] and references therein. The inability of conventional physics to explain the anomaly, has triggered a growing number of new theoretical explanations. One popular explanation was based on a Yukawa modification of Newton's law, but it quickly became clear that explaining the Pioneer anomaly in terms of a long-range Yukawa correction, would also lead to very large deviations from Kepler's law or of the precession of perihelions which obviously have not been observed [103, 104]. Other proposals suggest that we see here effects due to dark matter. If dark matter should be held accountable for the anomalous acceleration, it would have to be distributed in the form of a disk in the outer solar system, with a constant density of $\sim 4 \times 10^{-16} \text{ kg/m}^3$. The acceleration, due to any drag force from an interplanetary medium is:

$$a_d(r) = -K_d \rho(r) v_s^2(r) A/m, \quad (2.65)$$

with K_d the effective reflection/absorption/transmission coefficient of the spacecraft surface, $\rho(r)$ the density of the medium, v_s^2 the effective relative velocity of the craft to the medium, A the spacecraft's cross-section, and m its mass. The mass of the spacecraft is 241 kg, its cross-section about 5 m^2 and its speed nowadays roughly 12 km/s. A constant density of $\sim 4 \times 10^{-16} \text{ kg/m}^3$ would therefore explain the Pioneer Anomaly. This would be in correspondence with bounds on dark matter within the outer regions of our solar system of about $10^{-6} M_\odot$, [105]. A spherical halo of a degenerate gas of massive neutrinos with $m_\nu \leq 16 \text{ keV}$, around the sun [105] and mirror matter [106, 107] have been suggested.

Another proposal suggests that the modified inertia interpretation of gravity called 'Modified Newtonian Dynamics' or MOND for short, gives a correct explanation [108, 109]. Also certain braneworld scenarios [110] presumably can explain the observed effect, with a right adjustment of parameters for the potential of a scalar field.

Table 2.1: The “history” of the universe from the Planck scale. Some of the major events are also shown, together with the dominant type of physics.

Energy of back-ground radiation	Description	Time	Redshift
$\sim 10^{19}$ GeV	Planck Scale	10^{-44} sec	$z > 10^{46}$
$\sim 10^{16}$ GeV	Breaking of GUT	10^{-37} sec	
$\sim 10^{12}$ GeV	End of Inflation/Reheating	10^{-30} sec	$z \sim 10^{20}$
$\sim 10^3$ GeV	EW symm. breaking	10^{-10} sec	
1 GeV	quark-gluon plasma condens.	10^{-4} sec	
~ 100 MeV	pion decay/annihilation neutrino decoupling	10^{-4} sec	
~ 0.1 MeV	nucleosynthesis, start creation light elements	100 sec	$z \sim 10^4$
~ 30 keV	end nucleosynthesis	15 min	
~ 2 eV	matter-radiation equality	10.000 yr	3500
~ 0.35 eV	recombination, hydrogen-photon decoupling, origin CMBR	380.000 yr.	1100
	Galaxy formation	10^9 years	
10^{-4} eV	now	$1.4 \cdot 10^{10}$ yr.	$z=0$

3 | Type I: Symmetry Mechanism

One can set the cosmological constant to any value one likes, by simply adjusting the value of the bare cosmological constant. No further explanation then is needed. This fine-tuning has to be precise to better than at least 55 decimal places (assuming some TeV scale cut-off), but that is of course not a practical problem. Since we feel some important aspects of gravity are still lacking in our understanding and nothing can be learned from this ‘mechanism’, we do not consider this to be a physical solution. However, it is a possibility that we can not totally ignore and it is mentioned here just for sake of completeness.

A natural way to understand the smallness of a physical parameter is in terms of a symmetry that altogether forbids any such term to appear. This is also often referred to as ‘naturalness’: a theory obeys naturalness only if all of its small parameters would lead to an enhancement of its exact symmetry group when replaced by zero. Nature has provided us with several examples of this. Often mentioned in this respect is the example of the mass of the photon. The upper bound on the mass (squared) of the photon from terrestrial measurements of the magnetic field yields:

$$m_\gamma^2 \lesssim \mathcal{O}(10^{-50})\text{GeV}^2. \quad (3.1)$$

The most stringent estimates on Λ_{eff} nowadays give:

$$\Lambda_{eff} \lesssim \mathcal{O}(10^{-84})\text{GeV}^2 \quad (3.2)$$

We ‘know’ the mass of the photon to be in principle exactly equal to 0, because due to the $U(1)$ gauge symmetry of QED, the photon has only two physical degrees of freedom (helicities). In combination with Lorentz invariance this sets the mass equal to zero. A photon with only two transverse degrees of freedom can only get a mass if Lorentz invariance is broken. This suggests that there might also be a symmetry acting to keep the effective cosmological constant an extra 34 orders of magnitude smaller.

A perhaps better example to understand the smallness of a mass is chiral symmetry. If chiral symmetry were an exact invariance of Nature, quark masses and in particular masses for the pseudoscalar mesons (π, K, η) would be zero. The spontaneous breakdown of chiral symmetry would imply pseudoscalar Goldstone bosons, which would be massless in the limit of zero quark mass. The octet (π, K, η) would be the obvious candidate and indeed the pion is by far the lightest of the mesons. Making this identification of the pion being the pseudo-Goldstone boson associated with spontaneous breaking of chiral symmetry, we can understand why the pion-mass is so much smaller than for example the proton mass.

3.1 Supersymmetry

One symmetry with this desirable feature for the energy density of the ground state, is supersymmetry, or SUSY for short. Contributions to the energy density of the vacuum in field theory coming from fields with spin j are (eqn. (1.22)):

$$\begin{aligned} \langle \rho \rangle &= \frac{1}{2}(-1)^{2j}(2j+1) \int_0^{\Lambda_{UV}} \frac{d^3k}{(2\pi)^3} \sqrt{\mathbf{k}^2 + m^2} \\ &= \frac{(-1)^{2j}(2j+1)}{16\pi^2} \left(\Lambda_{UV}^4 + m^2 \Lambda_{UV}^2 - \frac{1}{4} m^4 \left[\log(\Lambda_{UV}^2/m^2) + \frac{1}{8} - \frac{1}{2} \log 2 \right] \right) \\ &+ \mathcal{O}(\Lambda_{UV}^{-1}). \end{aligned} \quad (3.3)$$

so if for each mass m there are an equal number of fermionic and bosonic degrees of freedom, the net contribution to $\langle \rho \rangle$ would be zero. Supersymmetry posits exactly such a symmetry, and hence realizes elegantly that the energy density of the ground state is zero.

The spin-1/2 generators of the supersymmetry transformations, called supercharges are denoted by Q and they satisfy anticommutation relations:

$$\{Q_\alpha, \bar{Q}_\beta\} = 2\sigma_{\alpha\beta}^\mu P_\mu, \quad (3.4)$$

where α and β are two-component spin indices, σ^μ are the Pauli-matrices, and P_μ is the energy-momentum operator. Summing over α and β , we find the Hamiltonian:

$$H = \frac{1}{4} \sum_\alpha \{Q_\alpha, \bar{Q}_\alpha\} = P^0, \quad (3.5)$$

where P^0 is the energy operator. Matrix elements of P^0 can be written as:

$$\langle \psi | P^0 | \psi \rangle = \frac{1}{4} \sum_\alpha \langle \psi_\alpha | \psi_\alpha \rangle, \quad (3.6)$$

with

$$\psi_\alpha = (Q_\alpha + \bar{Q}_\alpha) | \psi \rangle. \quad (3.7)$$

Thus in a supersymmetric theory, the energy of any non-vacuum state in a positive definite Hilbert space, is positive definite. Moreover, if supersymmetry is unbroken, the vacuum state satisfies $Q_\alpha | 0 \rangle = \bar{Q}_\alpha | 0 \rangle = 0$ for all α) and we see that this state has vanishing energy:

$$\langle 0 | P^0 | 0 \rangle = 0. \quad (3.8)$$

This includes all quantum corrections and would nicely explain a vanishing vacuum energy.

This same result can also be obtained by looking at the potential. The scalar field potential in supersymmetric theories takes on a special form. Scalar fields ϕ^i must be complex, to match the degrees of freedom coming from the fermions, and the potential is derived from a function, called the superpotential $W(\phi^i)$ which is necessarily holomorphic (written in terms of ϕ^i and not its complex conjugate $\bar{\phi}^i$). In the simple

Wess-Zumino models of just spin-0 and spin-1/2 fields, for example, the scalar potential is given by:

$$V(\phi^i, \bar{\phi}^i) = \sum_i |\partial_i W|^2, \quad (3.9)$$

where $\partial_i W = \partial W / \partial \phi^i$. Supersymmetry will be unbroken only for values of ϕ^i such that $\partial_i W = 0$, implying that V takes its minimum value $V(\phi^i, \bar{\phi}^i) = 0$. Quantum effects do not change this conclusion, since the boson-fermion symmetry ensures that boson-loops are cancelled against fermion-loops.

So if supersymmetry were an exact symmetry of nature, there would be bosons and superpartner-fermions with the same mass and the vacuum state of this theory would have zero energy. However, these supersymmetric partners of the Standard Model particles have not been found, so SUSY has to be broken at least at the TeV scale, which again induces a large vacuum energy.

Besides, we still have the freedom to add any constant to the Lagrangian, thereby changing the value of the effective cosmological constant. In order to discuss the cosmological constant problem properly, we need to bring gravity into the picture. This implies making the supersymmetry transformations local, leading to the theory of supergravity or SUGRA for short. In such a theory the value of the effective cosmological constant is given by the expectation value of the potential.

3.1.1 No-Scale SUGRA

In exact SUGRA the lowest energy state of the theory, generically has negative energy density¹. Or, in other words, the vacuum of supergravity is AdS. This has inspired many to consider so-called no-scale supergravity models in which supersymmetry breaking contributes precisely the amount of positive vacuum energy to make the net result equal to zero. See [40] or supersymmetry textbooks such as [111] for excellent reviews. In this section we closely follow [111].

The pure supergravity (SUGRA) Lagrangian, that is without interactions, is:

$$\mathcal{L}_{grav} = \mathcal{L}^{(2)} + \mathcal{L}^{(3/2)} - \frac{e}{3} (S^2 + P^2 - A_m^2), \quad (3.10)$$

where $\mathcal{L}^{(2)}$ is the Einstein-Hilbert term, $\mathcal{L}^{(3/2)}$ is the Rarita-Schwinger Lagrangian describing the free gravitino and P , a pseudoscalar, S a scalar and A_μ an axial vector, are auxiliary fields to match the on- and off-shell fermionic and bosonic degrees of freedom. The full Lagrangian is characterized by two arbitrary functions of the scalar fields (in the presence of gravity, renormalizability is lost anyway, so there is no reason to require only low-order polynomials in ϕ^i): a real function $G(\phi^i, \phi_j^*)$, called the Kähler potential and an analytic function $f_{ab}(\phi^i)$, where Latin indices are gauge indices. These functions determine the general forms allowed for the kinetic energy terms of the scalar fields ϕ^i and of the gauge fields A_μ^a , respectively. Some of the terms from the full Lagrangian that are important for now are:

$$\begin{aligned} \frac{1}{e} \mathcal{L}_{tot} = & - \frac{1}{2} R + G_i^j \partial_\mu \phi^i \partial^\mu \phi_j^* - e^G [G_i (G^{-1})_j^i G^j - 3] \\ & - \frac{1}{4} \text{Re}(f_{ab}) F_{\mu\nu}^a F^{\mu\nu b} + e^{G/2} \bar{\psi}_\mu \sigma^{\mu\nu} \psi_\nu + \dots \end{aligned} \quad (3.11)$$

¹This negative energy density can also be forbidden by postulating an unbroken R-symmetry.

in which the gravitational coupling κ has been set equal to one and:

$$G_i \equiv \frac{\partial G}{\partial \phi^i}, \quad G^j \equiv \frac{\partial G}{\partial \phi_j^*}, \quad G_i^j \equiv \frac{\partial^2 G}{\partial \phi^i \partial \phi_j^*}. \quad (3.12)$$

The scalar kinetic energy term shows that G_i^j plays the role of the metric in the space spanned by the scalar fields. A metric G_i^j of this form is referred to as a Kähler metric and G is called the Kähler potential. In the absence of gravity $G_i^j \rightarrow \delta_i^j$ and $f_{ab} \rightarrow \delta_{ab}$. The function G is invariant under transformations of the gauge group, whereas f_{ab} transforms as a symmetric product of two adjoint representations of that group.

The Lagrangian contains a scalar potential which generally consists of so-called D-terms and F-terms. The D-terms arise from removing the auxiliary scalar fields contributing to the gauge multiplets, and the F-terms from removing the auxiliary scalar fields of the chiral matter multiplets. With W denoting the superpotential, one finds:

$$|F|^2 \rightarrow \frac{\partial W}{\partial \phi^i} \frac{\partial W^*}{\partial \phi_j^*} (G^{-1})_j^i, \quad (3.13)$$

where we now must include the Kähler metric. This gives the first term of the scalar potential:

$$V(\phi, \phi^*) = e^G [G_i (G^{-1})_j^i G^j - 3], \quad (3.14)$$

In case of supergravity, there is an additional contribution from eliminating the auxiliary scalar field terms $-|S + iP|^2$ in (3.10) and this yields the second term of (3.14). The negative sign has important consequences. The e^G factor arises from the Weyl rescaling of the $e_{m\mu}$ fields required to bring the first term in (3.11) into the canonical Einstein form. This rescaling also implies a redefinition of the fermion fields and hence the factor $e^{G/2}$ in the last term as well. Owing to this term, when local SUSY is spontaneously broken the gravitino acquires a mass:

$$m_{3/2} = e^{G/2}, \quad (3.15)$$

where G then has to be evaluated at the minimum of the potential.

In general the Kähler potential has to satisfy certain conditions for the theory to be well defined, for example $G_i^j > 0$, in order for the kinetic terms of the scalar fields to have the correct sign. A special choice is:

$$G(\phi, \phi^*) = \phi^i \phi_i^* + \log (W(\phi^i)^2)^2, \quad (3.16)$$

with W the superpotential. This choice gives $G_i^j = \delta_i^j$ and hence the minimal kinetic terms as in global SUSY. The scalar potential becomes:

$$V = \exp(\phi^i \phi_i^*) \left[\left| \frac{\partial W}{\partial \phi^i} + \phi_i^* W \right|^2 - 3|W|^2 \right]. \quad (3.17)$$

Now we come to the important point. There is an elegant way of guaranteeing a flat potential, with $V = 0$ after supersymmetry breaking, by using a nontrivial form of the Kähler potential G . So far we have used the minimal form of G which lead to the above

scalar potential. For a single scalar field z :

$$\begin{aligned} V &= e^G \left[\frac{\partial_z G \partial_{z^*} G}{\partial_z \partial_{z^*} G} - 3 \right] \\ &= \frac{9e^{4G/3}}{\partial_z \partial_{z^*} G} (\partial_z \partial_{z^*} e^{-G/3}). \end{aligned} \quad (3.18)$$

A flat potential with $V = 0$ is obtained if the expression in brackets vanishes for all z , which happens if:

$$G = -3 \log(z + z^*), \quad (3.19)$$

and one obtains a gravitino mass:

$$m_{3/2} = \langle e^{G/2} \rangle = \langle (z + z^*)^{-3/2} \rangle, \quad (3.20)$$

which, as required, is not fixed by the minimization of V . Thus *provided one chooses a suitable, nontrivial form for the Kähler potential G , it is possible to obtain a zero CC and to leave the gravitino mass undetermined*, just fixed dynamically through non-gravitational radiative corrections. The minimum of the effective potential occurs at:

$$V_{eff} \approx -(m_{3/2})^4, \quad (3.21)$$

where in this case, after including the observable sector and soft symmetry-breaking terms, $m_{3/2} \approx M_W$. Such a mass is ruled out cosmologically, see e.g. [112], and so other models with the same ideas have been constructed that allow a very small mass for the gravitino, also by choosing a specific Kähler potential, see [113].

That these constructions are possible is quite interesting and in the past there has been some excitement when superstring theory seemed to implicate precisely the kinds of Kähler potential as needed here, see for example [114]. However, that is not enough, these simple structures are not expected to hold beyond zeroth order in perturbation theory.

$D = 11$ SUGRA seems to be a special case; its symmetries implicitly forbid a CC term, see [115]. However, also here, the vanishing of the vacuum energy is a purely classical phenomenon, which is spoiled by quantum corrections after supersymmetry breaking.

3.1.2 Unbroken SUSY

To paraphrase Witten [116]: “Within the known structure of physics, supergravity in four dimensions leads to a dichotomy: either the symmetry is unbroken and bosons and fermions are degenerate, or the symmetry is broken and the vanishing of the CC is difficult to understand”. However, as he also argues in the same article, in $2 + 1$ dimensions, this unsatisfactory dichotomy does not arise: SUSY can explain the vanishing of the CC without leading to equality of boson and fermion masses, see also [117].

The argument here is that in order to have equal masses for the bosons and fermions in the same supermultiplet one has to have unbroken global supercharges. These are determined by spinor fields which are covariantly constant at infinity. The supercurrents J^μ from which the supercharges are derived are generically not conserved in the usual sense, but covariantly conserved: $D_\mu J^\mu = 0$. However, in the presence of a covariantly

constant spinor ($D_\mu \epsilon = 0$), the conserved current $\bar{\epsilon} J^\mu$ can be constructed and therefore, a globally conserved supercharge:

$$Q = \int d^3x \bar{\epsilon} J^0. \quad (3.22)$$

But in a 2+1 dimensional spacetime any state of non-zero energy produces a geometry that is asymptotically conical at infinity (see also [118]). The spinor fields are then no longer covariantly constant at infinity [119] and so even when supersymmetry applies to the vacuum and ensures the vanishing of the vacuum energy, it does not apply to the excited states. Explicit examples have been constructed in [120, 121, 122, 123]. Two further ideas in this direction, one in $D < 4$ and one in $D > 4$ are [124, 125], however the latter later turned out to be internally inconsistent [126].

In any case, what is very important is to make the statement of “breaking of supersymmetry” more precise. As is clear, we do not observe mass degeneracies between fermions and bosons, therefore supersymmetry, even if it were a good symmetry at high energies between excited states, is broken at lower energies. However, and this is the point, as the example of Witten shows, the issue of whether we do or do not live in a supersymmetric vacuum state is another question. In some scenarios it is possible to have a supersymmetric vacuum state, without supersymmetric excited states. This really seems to be what we are looking for. The observations of a small or even zero CC could point in the direction of a (nearly) supersymmetric vacuum state.

Obviously the question remains how this scenario and the absence nevertheless of a supersymmetric spectrum can be incorporated in 4 dimensions, where generically spacetime is asymptotically flat around matter sources.

3.1.3 Non-SUSY String Models

It has been suggested that the cosmological constant might vanish in certain string models [127, 128] due to an equality between the number of boson and fermion mass states, despite the fact that supersymmetry is broken. The one-loop contribution to the cosmological constant therefore vanishes trivially, but it is claimed that even higher order corrections would not spoil this cancellation.

These ideas have been challenged in [129, 130], where it was argued that higher loop corrections would indeed spoil the tree level results. The most important drawbacks however for this proposed scenario, is that the non-Abelian gauge sector, was always supersymmetric [131, 132, 133]. These approaches thus do not lead to viable models.

3.2 Scale Invariance, e.g. Conformal Symmetry

Another interesting symmetry with respect to the cosmological constant problem is conformal symmetry, $g_{\mu\nu} \rightarrow f(x^\mu)g_{\mu\nu}$. Massless particles are symmetric under a bigger group than just the Lorentz group, namely, the conformal group. This group does not act as symmetries of Minkowski spacetime, but under a (mathematically useful) completion, the “conformal compactification of Minkowski space”. This group is 15-dimensional and corresponds to $SO(2, 4)$, or if fermions are present, the covering group

$SU(2, 2)$. Conformal symmetry forbids any term that sets a length scale, so a cosmological constant is not allowed, and indeed also particle masses necessarily have to vanish.

General coordinate transformations and scale invariance, i.e. $g_{\mu\nu} \rightarrow f g_{\mu\nu}$, are incompatible in general relativity. The $R\sqrt{-g}$ term in the Einstein-Hilbert action is the only quantity that can be constructed from the metric tensor and its first and second derivatives only, that is invariant under general coordinate transformations. But this term is not even invariant under a global scale transformation $g_{\mu\nu} \rightarrow f g_{\mu\nu}$ for which f is constant. R transforms with Weyl weight -1 and $\sqrt{-g}$ with weight $+2$. There are two ways to proceed to construct a scale invariant action: introducing a new scalar field [134, 135], that transforms with weight -1 , giving rise to so-called scalar-tensor theories, or consider Lagrangians that are quadratic in the curvature scalar. We consider the second. See for example [136, 137] for some recent studies and many references.

Gravity can be formulated under this bigger group, leading to ‘‘Conformal gravity’’, defined in terms of the Weyl tensor, which corresponds to the traceless part of the Riemann tensor:

$$\begin{aligned} S_G &= -\alpha \int d^4x \sqrt{-g} C_{\lambda\mu\nu\kappa} C^{\lambda\mu\nu\kappa} \\ &= -2\alpha \int d^4x \sqrt{-g} \left(R_{\mu\nu} R^{\mu\nu} - \frac{1}{3} R^2 \right) + (\text{boundary terms}), \end{aligned} \quad (3.23)$$

where $C^{\mu\nu\lambda\kappa}$ is the conformal Weyl tensor, and α is a dimensionless gravitational coupling constant. Thus the Lagrangian is quadratic in the curvature scalar and generates field equations that are fourth-order differential equations. Based on the successes of gauge theories with spontaneously broken symmetries and the generation of the Fermi-constant, one may suggest to also dynamically induce the Einstein action with its Newtonian constant as a macroscopic limit of a microscopical conformal theory. This approach has been studied especially by Mannheim and Kazanas, see [138, 139, 140, 141, 142, 143] to solve the CC problem.

These fourth-order equations reduce to a fourth-order Poisson equation:

$$\nabla^4 B(r) = f(r), \quad (3.24)$$

where $B(r) = -g_{00}(r)$ and the source is given by:

$$f(r) = 3(T_0^0 - T_r^r)/4\alpha B(r), \quad (3.25)$$

For a static, spherically symmetric source, conformal symmetry allows one to put $g_{rr} = -1/g_{00}$ and the exterior solution to (3.24) can be written [143]:

$$-g_{00} = 1/g_{rr} = 1 - \beta(2 - 3\beta\gamma)/r - 3\beta\gamma + \gamma r - kr^2. \quad (3.26)$$

Assuming that the quadratic term is negligible at solar system distance scales, the non-relativistic potential can be written:

$$V(r) = -\beta/r + \gamma r/2 \quad (3.27)$$

However, for a spherical source (3.24) can be completely integrated to yield:

$$B(r > R) = -\frac{r}{2} \int_0^R dr' f(r') r'^2 - \frac{1}{6r} \int_0^R dr' f(r') r'^4. \quad (3.28)$$

Compared to the standard second-order equations:

$$\nabla^2\phi(r) = g(r) \quad \rightarrow \quad \phi(r > R) = -\frac{1}{r} \int_0^R dr' g(r') r'^2 \quad (3.29)$$

we see that the fourth-order equations contain the Newtonian potential in its solution, but in addition also a linear potential term that one would like to see dominate over Newtonian gravity only at large distances. The factors β and γ in for example (3.27) are given by:

$$\beta(2 - 3\beta\gamma) = \frac{1}{6} \int_0^R dr' f(r') r'^4 \quad , \quad \gamma = -\frac{1}{2} \int_0^R dr' f(r') r'^2 \quad (3.30)$$

Even if this would be correct, modifying gravity only at large distances cannot solve the cosmological constant problem. The (nearly) vanishing of the vacuum energy and consequently flat and relatively slowly expanding spacetime is a puzzle already at distance scales of say a meter. We could expect deviations of GR at galactic scales, avoiding the need for dark matter, but at solar system scales GR in principle works perfectly fine. It seems hard to improve on this with conformal symmetry, since the world simply is not scale invariant.

We also identified a more serious problem with the scenario of Mannheim and Kazanas described above. In order for the linear term not to dominate already at say solar system distances, the coefficient γ has to be chosen very small. Not only would this introduce a new kind of fine-tuning, it is also simply not allowed to choose these coefficients at will. The linear term will always dominate over the Newtonian $1/r$ -term, in contradiction with the perfect agreement of GR at these scales. See also [144] who raised the same objection.

This scenario therefore does not work.

3.2.1 Λ as Integration Constant

Another option is to reformulate the action principle in such a way that a scale dependent quantity like the scalar curvature, remains undetermined by the field equations themselves. These are the so-called ‘unimodular’ theories of gravity, see e.g. [145, 146]. Note that although the action is not globally scale invariant, Einstein’s equations in the absence of matter and with vanishing cosmological constant is. The dynamical equations of pure gravity in other words, are invariant with respect to global scale transformations, and since we have that $R = 0$, they are scale-free, i.e. they contain no intrinsic length scale.

There is a way to keep the scale dependence undetermined also after including matter which also generates a cosmological constant term. This well-known procedure [147, 148] starts by imposing a constraint on the fluctuations in $g_{\mu\nu}$, such that the determinant of the metric is fixed:

$$\sqrt{-g} = \sigma(x) \quad \rightarrow \quad \delta\sqrt{-g} = 0, \quad (3.31)$$

where $\sigma(x)$ is a scalar density of weight +1. The resulting field equations are:

$$R_{\mu\nu} - \frac{1}{4}g_{\mu\nu}R = -\kappa \left(T_{\mu\nu} - \frac{1}{4}g_{\mu\nu}T \right). \quad (3.32)$$

The covariant derivative $D_\mu G_{\mu\nu} = D_\mu T_{\mu\nu} = 0$ still vanishes and from this one obtains:

$$R - \kappa T = -4\Lambda, \quad (3.33)$$

where Λ now appears as an integration constant and the factor of 4 has been chosen for convenience since substituting this back we recover the normal Einstein equations with cosmological constant.

Recently, some arguments have been put forward in which a unimodular theory is supposed to originate more naturally as a result of ‘the quantum microstructure of spacetime being capable of readjusting itself, soaking up any vacuum energy’, see [149, 38, 150].

Obviously this does not solve much, nor does it provide a better understanding of the cosmological constant. The restriction (3.31) on the variation of the metric has no deeper motivation and the value of the integration constant Λ has to be inserted by hand in order to arrive at the correct value.

Besides, sometimes it is concluded that there are two inequivalent Einstein equations for gravity, describing two theories that are only equivalent classically, but not quantum mechanically. The group of canonical transformations is much larger than that of unitary transformations in Hilbert space, forcing one to quantize in “preferred” coordinates. We do not agree with this point of view. The constraint $g^{\mu\nu}\delta g_{\mu\nu} = 0$ just reflects a choice of coordinates, a certain gauge.

This issue is closely related to the question of the measure of the quantum gravity functional integral (see discussions by B.S. DeWitt [151, 152], ’t Hooft [153] and [154]): Is the integration variable $g_{\mu\nu}$, $g^{\mu\nu}$ or some other function of the metric? The differences in the amplitudes for these theories all appear in the one-loop diagrams, in the form of quartically divergent momentum-independent ghost loops. These all disappear after renormalization and therefore the theories are indistinguishable physically.

3.3 Holography

Gravitational holography [155, 156] limits the number of states accessible to a system. The entropy of a region generally grows with its covering area (in Planck units) rather than with its volume, implying that the dimension of the Hilbert space, i.e. the number of degrees of freedom describing a region, is finite and much smaller than expected from quantum field theory. Considering an infinite contribution to the vacuum energy is not correct because states are counted that do not exist in a holographic theory of gravity.

It is a symmetry principle since there is a projection from states in the bulk-volume, to states on the covering surface.

In [157, 158] it is noted that in effective field theory in a box of size L with UV cutoff M the entropy S scales extensively, as $S \sim L^3 M^3$. A free Weyl fermion on a lattice of size L and spacing $1/M$ has $4^{(LM)^3}$ states and entropy² $S \sim (LM)^3$. The corresponding entropy density $s = S/V$ then is $s = M^3$. In $d = 4$ dimensions quantum corrections to

²For bosons the number of states is not limited by a lattice cutoff alone, so in this argument one has to limit oneself to fermions. For bosons there are an infinite number of states, in contradiction to the conjecture of the Holographic Principle.

the vacuum energy are therefore of order:

$$\rho_{vac} = \frac{\Lambda}{8\pi G} + \langle \rho \rangle = \frac{\Lambda}{8\pi G} + \mathcal{O}(s^{4/3}), \quad (3.34)$$

since both $\langle \rho \rangle$ and s are dominated by ultraviolet modes, (see also [159]). Thus finite s implies finite corrections to $\langle \rho \rangle$.

Using a cutoff M , $E \sim M^4 L^3$ is the maximum energy for a system of size L . States with $L < R_s \sim E$, with R_s the Schwarzschild radius, or $L > M^{-2}$ (in Planckian units) have collapsed into a black-hole. If one simply requires that no state in the Hilbert space exists with $R_s \sim E > L$, then a relation between the size L of the region, providing an IR cutoff, and the UV cutoff M can be derived, in natural units:

$$L^3 M^4 \lesssim L M_P^2 \quad (3.35)$$

This corresponds to the assumption that the effective theory describes all states of the system, except those that have already collapsed to a black hole.

Under these conditions entropy grows no faster than $A^{3/4} \sim L^{3/2}$, with A the area. If these black hole states give no contribution to $\langle \rho \rangle$, we obtain:

$$\langle \rho \rangle \sim s^{4/3} \sim \left(\frac{L^{3/2}}{L^3} \right)^{4/3} \sim L^{-2}. \quad (3.36)$$

In [157] this same scaling was obtained by assuming that $S < A$ as usual, but that the delocalized states have typical Heisenberg energy $1/L$:

$$\langle \rho \rangle \sim \frac{s}{L} \sim \frac{L^2}{L^3 L} \sim L^{-2}. \quad (3.37)$$

Plugging in for L the observed size of the universe today the quantum corrections are only of order 10^{-10} eV^4 .

However, this does not yield the correct equation of state, [159]. During matter dominated epochs, to which WMAP and supernova measurements are sensitive, the horizon size grows as the RW-scale factor, $a(t)^{3/2}$, so the above arguments imply:

$$\Lambda_{eff}(L) \sim a(t)^{-3}, \quad (3.38)$$

or, $w \equiv p/\rho = 0$ at largest scales, since $\rho(t) \sim a(t)^{-3(1+w)}$. The data on the other hand give $w < -0.78$ (95% CL) [19]. In [157, 158] $\Lambda(L)$ is at all times comparable to the radiation + matter energy density, which is also argued to give problems for structure formation [160].

Holography-based scenarios thus naively lead to a cosmological constant that is far less constant than what the data require. This makes a connection between holography and dark energy a lot harder to understand³.

More recently however, another proposal was made [162] where instead L is taken to be proportional to the size of the future event horizon:

$$L(t) \sim a(t) \int_t^\infty \frac{dt'}{a(t')} \quad (3.39)$$

³In [161] in a different context a similar relation between the CC and the volume of the universe is derived, thus suffering from the same drawbacks.

This L describes the size of the largest portion of the universe that any observer will see. This could be a reasonable IR cutoff. It is argued that in this case the equation of state parameter w can be close enough to -1 to agree with the data. This relation is rather ad hoc chosen, and its deeper meaning, if any, still has to be discovered.

A very recent paper studying the implications of the holographic principle for the cosmological constant problem is [163], in which it is argued that holography makes a tiny value for the cosmological constant ‘natural’ in a large universe. This is however a rather empty statement, since a large (observable) universe necessarily must have a tiny value for the cosmological constant. In other words, if the universe is large the effective cosmological constant must be small, and no holographic arguments are needed for that.

3.3.1 Conceptual Issues in de Sitter Space

Another reason to discuss holography in the context of the cosmological constant problem lies in trying to reconcile string theory with the apparent observation of living in a de Sitter spacetime, see [164] for a construction of de Sitter vacua in string theory. The discussion centers around the semi-classical result that de Sitter space has a finite entropy, inversely related to the cosmological constant. As discussed in section (1.5), the entropy is given by:

$$S_{dS} = \frac{A}{4G}, \quad \text{with} \quad A = \frac{12\pi}{\Lambda} \quad (3.40)$$

with A the area of the horizon. It seems natural to interpret this entropy as the number of quantum states necessary to describe a de Sitter universe [39, 165]. This interpretation is common in thermodynamics, and is also given for the entropy of a black hole. Another motivation for this interpretation arises from the fact that the classical phase space of general relativity with asymptotic de Sitter boundary conditions, both in the past and in the future, is compact. A compact phase space yields a finite dimensional Hilbert space when quantized. Thus one may reason that de Sitter space, or better, asymptotic de Sitter space, since the metric fluctuates, should be described by a theory with a finite number of independent quantum states and that a theory of quantum gravity should be constructed with a Hilbert space of finite dimension \mathcal{N} in terms of which the entropy is given by:

$$S_{dS} = \ln \mathcal{N}. \quad (3.41)$$

In this reasoning a cosmological constant should be understood as a direct consequence of the finite number of states in the Hilbert space describing the world. Ergo, the larger the cosmological constant, the smaller the Hilbert space:

$$\Lambda = \frac{3\pi}{\ln \mathcal{N}}. \quad (3.42)$$

Banks [39] therefore argues that the cosmological constant should not be viewed as a derived, calculable quantity, but instead as a fundamental input parameter. In [166] this reasoning is criticized. In any case, finiteness of entropy and Hilbert space leads to several conceptual difficulties.

One is that only compact groups can have non-trivial finite-dimensional unitary representations, but the de Sitter group is non-compact. Therefore, it has been claimed that either de Sitter space has no holographic dual [35], which would make it impossible to have an analog of the successful AdS/CFT dual for de Sitter space, or that in fact the correct symmetry group is not the standard de Sitter group [34]. However, whether these claims hold in the future is unclear, and ways out of this conundrum have been proposed, for example [167]. See also [168, 37] in which a holographic dS/CFT correspondence is formulated, in apparent contradiction to the above claim, since the CFT Hilbert space is in principle infinite dimensional. It appears to be an unsettled question whether this leads to consistent theories, especially since no explicit example is known, in which a dS/CFT emerges directly from string theory, see also the discussion in [36].

Another issue, as hinted on in the first chapter of this thesis, arguments have been made that the standard Einstein-Hilbert action with cosmological constant cannot be quantized for general values of G and Λ , but that it must be derived from a more fundamental theory, which determines these values [34]. The reasoning behind this statement is that although the Hilbert space of quantum gravity in de Sitter space has finite dimension, it is infinite dimensional perturbatively. Perturbation theory is an expansion in powers $G\Lambda$, in four dimensions. In the limit $G\Lambda \rightarrow 0$, \mathcal{N} diverges exponentially, if (3.41) holds. In fact, \mathcal{N} will be non-trivial function of $G\Lambda$, yet it has to take integer values, whereas the latter can vary continuously.

However, it not immediately clear to which states the number \mathcal{N} refers to. Because of the appearance of negative absolute temperatures, it has been argued in [169] that instead of the usual thermodynamic relation:

$$\frac{1}{T} = \frac{\partial S}{\partial E}, \quad \frac{1}{T} = \frac{\partial S}{\partial(-E)} \quad (3.43)$$

the equation on the right should be used to compute the de Sitter temperature. The idea behind this is that the de Sitter entropy should perhaps be ascribed to states behind the horizon, which cannot be observed.

Finally, another difficulty in de Sitter space has to do with formulating ordinary quantum field theory. In quantum field theory, in order to set up an S-matrix, one has to define incoming and outgoing states, and these are only properly defined at spatial infinity. De Sitter space however, is compact, there is no notion of spatial infinity, only of temporal past and future asymptotic regions. Matrix elements constructed this way are therefore no measurable quantities. Note that also the conventional formulation of string theory is based on the existence of an S-matrix.

Another quantum field theory aspect is that there appear to be instabilities in de Sitter space, for example in case of a scalar field, but perhaps also in pure gravity. These issues are discussed in chapter 5 and 6.

Much more needs to be understood about these issues to judge their validity and possible impact on the cosmological constant problem. It is clear however, that the measurements indicating an accelerating universe and a dark energy equation of state $w \approx -1$, have much more far-reaching consequences than ‘just’ the ordinary cosmological constant problem. In other words, a solution to the cosmological constant problem, especially if not by some symmetry, seems to have very deep implications for a broad range of theoretical physics.

3.4 “Symmetry” between Sub- and Super-Planckian Degrees of Freedom

This rather speculative reasoning originates from a comparison with condensed matter physics and is due to Volovik, see for example [170, 171, 172, 173, 174, 175, 176, 177]. The vacuum energy of superfluid ⁴Helium, calculated from an effective theory containing phonons as elementary bosonic particles and no fermions is:

$$\rho_\Lambda = \sqrt{-g}E_{Debye}^4 \quad (3.44)$$

with g the determinant of the acoustic metric, since c is now the speed of sound, and $E_{Debye} = \hbar c/a$, with a the interatomic distance, which plays the role of the Planck length. However, in the condensed matter case, the full theory exists: a second quantized Hamiltonian describing a collection of a macroscopic number of structureless ⁴Helium bosons or ³Helium fermions, in which the chemical potential μ acts as a Lagrange multiplier to ensure conservation of the number of atoms:

$$\begin{aligned} H - \mu N &= \int d\vec{x} \psi^\dagger(\vec{x}) \left[-\frac{\nabla^2}{2m} - \mu \right] \psi(\vec{x}) \\ &+ \int d\vec{x} d\vec{y} V(\vec{x} - \vec{y}) \psi^\dagger(\vec{x}) \psi^\dagger(\vec{y}) \psi(\vec{y}) \psi(\vec{x}). \end{aligned} \quad (3.45)$$

Using this Hamiltonian H to calculate the energy density of the ground state we get:

$$E_{vac} = E - \mu N = \langle \text{vac} | H - \mu N | \text{vac} \rangle \quad (3.46)$$

An overall shift of the energy in H is cancelled in a shift of the chemical potential. Exact calculation shows that not only the low energy degrees of freedom from the effective theory, the phonons, but also the higher energy, “trans-Planckian” degrees of freedom have to be taken into account.

Besides, for a liquid of N identical particles at temperature T in a volume V in equilibrium, the relation between the energy E and pressure P is given by the Gibbs-Duhem equation:

$$E = TS + \mu N - PV. \quad (3.47)$$

Therefore at $T = 0$ the energy density of the ground state becomes:

$$\rho_{vac} \equiv \frac{E_{vac}}{V} = -P_{vac}, \quad (3.48)$$

the same equation of state as for the vacuum state in GR. Using just thermodynamic arguments, it is argued that in the infinite volume, zero temperature limit, this gives exactly zero vacuum energy density as long as there are no external forces, i.e. no pressure acting on the quantum liquid. And assuming there is no matter, no curvature and no boundaries which could give rise to a Casimir effect [171].

The conclusion therefore is that, if these thermodynamic arguments are also valid in a gravitational background for the universe as a whole and up to extremely high energies, one would expect a perfect cancellation between sub- and super-Planckian degrees of freedom contributing to the vacuum energy, resulting in zero cosmological constant.

Moreover, it is also argued that a non-zero cosmological constant arises from perturbations of the vacuum at non-zero temperature. The vacuum energy density would be proportional to the matter energy density, solving the coincidence problem as well.

A similar result is obtained by [178]. In their formulation the world is like a crystal. The atoms of the crystal are in thermal equilibrium and exhibit therefore zero pressure, making the cosmological constant equal to zero.

Both approaches strongly depend on the quantum systems reaching their equilibrium state. However, in the presence of a cosmological constant, the matter in the universe never reaches its equilibrium state [179].

3.5 Interacting Universes, Antipodal Symmetry

This is an approach developed by Linde [180, 181] arguing that the vacuum energy in our universe is so small because there is a global interaction with another universe where energy densities are negative. Consider the following action of two universes with coordinates x_μ and y_α respectively, ($x_\mu, y_\alpha = 0, 1, \dots, 3$) and metrics $g_{\mu\nu}(x)$ and $\bar{g}_{\alpha\beta}(y)$, containing fields $\phi(x)$ and $\bar{\phi}(y)$:

$$S = N \int d^4x d^4y \sqrt{g(x)} \sqrt{\bar{g}(y)} \left[\frac{M_P^2}{16\pi} R(x) + L(\phi(x)) - \frac{M_P^2}{16\pi} R(y) - L(\bar{\phi}(y)) \right], \quad (3.49)$$

and where N is some normalization constant. This action is invariant under general coordinate transformations in each of the universes separately. The important symmetry of the action is $\phi(x) \rightarrow \bar{\phi}(x)$, $g_{\mu\nu}(x) \rightarrow \bar{g}_{\alpha\beta}(x)$ and under the subsequent change of the overall sign: $S \rightarrow -S$. He calls this an antipodal symmetry, since it relates states with positive and negative energies. As a consequence we have invariance under the change of values of the effective potentials $V(\phi) \rightarrow V(\phi) + c$ and $V(\bar{\phi}) \rightarrow V(\bar{\phi}) + c$ where c is some constant. Therefore nothing in this theory depends on the value of the effective potentials in their absolute minima ϕ_0 and $\bar{\phi}_0$. Note that because of the antipodal symmetry $\phi_0 = \bar{\phi}_0$ and $V(\phi_0) = V(\bar{\phi}_0)$.

In order to avoid the troublesome issues of theories with negative energy states, there can be no interactions between the fields $\phi(x)$ and $\bar{\phi}(y)$. Therefore also the equations of motion for both fields are the same and similarly, also gravitons from both universes do not interact.

However some interaction does occur. The Einstein equations are:

$$R_{\mu\nu}(x) - \frac{1}{2} g_{\mu\nu} R(x) = -8\pi G T_{\mu\nu}(x) - g_{\mu\nu} \left\langle \frac{1}{2} R(y) + 8\pi G L(\bar{\phi}(y)) \right\rangle \quad (3.50)$$

$$R_{\alpha\beta}(y) - \frac{1}{2} \bar{g}_{\alpha\beta} R(y) = -8\pi G T_{\alpha\beta}(y) - \bar{g}_{\alpha\beta} \left\langle \frac{1}{2} R(x) + 8\pi G L(\phi(x)) \right\rangle. \quad (3.51)$$

Here $T_{\mu\nu}$ is the energy-momentum tensor of the fields $\phi(x)$ and $T_{\alpha\beta}$ the energy-momentum tensor for the fields $\bar{\phi}(y)$ and the averaging means:

$$\langle R(x) \rangle = \frac{\int d^4x \sqrt{g(x)} R(x)}{\int d^4x \sqrt{g(x)}} \quad (3.52)$$

$$\langle R(y) \rangle = \frac{\int d^4y \sqrt{\bar{g}(y)} R(y)}{\int d^4y \sqrt{\bar{g}(y)}} \quad (3.53)$$

and similarly for $\langle L(x) \rangle$ and $\langle L(y) \rangle$.

Thus there is a global interaction between the universes X and Y : The integral over the whole history of the Y -universe changes the vacuum energy density of the X -universe. These averages could be hard to calculate. Therefore, it is assumed that both universes undergo a long enough period of inflation, such that they become almost flat, and that at late times the fields will settle near the absolute minimum of their potential. As a result, the average of $-L(\phi(x))$ will almost coincide with the value of $V(\phi_0)$, and the average of $R(x)$ coincides with its value at late stages and similarly for $-L(\phi(y))$ and $R(y)$. We arrive at:

$$R_{\mu\nu}(x) - \frac{1}{2}g_{\mu\nu}R(x) = -8\pi Gg_{\mu\nu} [V(\bar{\phi}_0) - V(\phi_0)] - \frac{1}{2}g_{\mu\nu}R(y) \quad (3.54)$$

$$R_{\alpha\beta}(y) - \frac{1}{2}\bar{g}_{\alpha\beta}R(y) = -8\pi Gg_{\alpha\beta} [V(\phi_0) - V(\bar{\phi}_0)] - \frac{1}{2}g_{\alpha\beta}R(x). \quad (3.55)$$

Thus at late stages the effective cosmological constant vanishes:

$$R(x) = -R(y) = \frac{32}{3}\pi G [V(\phi_0) - V(\bar{\phi}_0)] = 0, \quad (3.56)$$

since because of the antipodal symmetry $\phi_0 = \bar{\phi}_0$ and $V(\phi_0) = V(\bar{\phi}_0)$.

This could also be seen as a back-reaction mechanism, from one universe at the other.

Difficulties with this approach are:

1. The form of the theory is completely ad hoc, devised just to make the CC vanish and for that purpose we need to add a new universe with a negative energy density.
2. Theory is completely classical; not obvious how to quantize it, nor whether the cancellation of the CC survives quantum corrections.
3. The cancellation depends on ϕ eventually settling down, in order to calculate the averages. It is not clear how to generalize this.

3.6 Duality Transformations

3.6.1 S-Duality

A different proposal was considered in [182], where S-duality acting on the gravitational field is assumed to mix gravitational and matter degrees of freedom. The purpose is to show that whereas the original metric may be (A)dS, the dual will be flat. Only metrics are considered for which:

$$R^a_b \equiv R^{ca}_{bc} = \Lambda \delta^a_b, \quad (3.57)$$

with Λ the cosmological constant. The mixing between gravitational and matter degrees of freedom is obtained through a new definition of the gravitational dual of the Riemann tensor, including the field strength F_{abcd} of a 3-form field A_{abc} , which equation of motion is simply $F_{abcd} = \omega \epsilon_{abcd}$, with ω some constant, see also section (8.1):

$$\begin{aligned} \tilde{R}_{abcd} &= \frac{1}{2}\epsilon_{abef} \left(R^{ef}_{cd} + F^{ef}_{cd} \right) + \frac{1}{12}\epsilon_{abcd}R, \\ \tilde{F}_{abcd} &= -\frac{1}{2}\epsilon_{abcd}R \end{aligned} \quad (3.58)$$

such that:

$$\begin{aligned}\tilde{\tilde{R}}_{abcd} &= -R_{abcd} \\ \tilde{\tilde{F}}_{abcd} &= -F_{abcd}.\end{aligned}\tag{3.59}$$

The equations of motion for the dual tensors become:

$$\begin{aligned}\tilde{\tilde{R}}^a_b &= 3\omega\delta^a_b \\ \tilde{\tilde{F}}_{abcd} &= -\frac{1}{3}\Lambda\epsilon_{abcd} \equiv \tilde{\omega}\epsilon_{abcd}.\end{aligned}\tag{3.60}$$

Therefore, if the vev ω would vanish, the dual Ricci tensor, in casu the dual cosmological constant would also vanish. Hence the conclusion is that if we would moreover ‘see’ the dual metric, determined by the dual Riemann tensor, we would observe a flat spacetime.

But in order for this to work, one has to limit oneself to spacetimes satisfying (3.57). Besides, the duality relations have only proven to be consistent at linearized level. There is also no particular reason for the vev of ω to vanish.

Note that if one would constrain oneself to metrics which satisfy $R = -4\Lambda$, the trace of the left-hand-side of Einstein’s equation vanishes by definition. In that case also the trace of the energy-momentum tensor should vanish, which in general is not the case. In other words, the vev of the 3-form field would then have to be either zero or already incorporated in Λ , which would render the duality transformations empty. Such a constraint would be too strict, yet is very similar to (3.57).

Note that S-duality is an important concept in string theory. If theories A and B are S-dual then $f_A(\alpha) = f_B(1/\alpha)$. It relates type I string theory to the $SO(32)$ heterotic theory, and type IIB theory to itself.

3.6.2 Hodge Duality

This duality between a r -form and a $(D - r)$ -form in D dimensions is studied in [183], where the cosmological constant is taken to be represented by a 0-form field strength, which is just a constant. This is somewhat related to the unimodular approach of section (3.2.1) in the sense that the intention is to introduce the cosmological constant in a different way in the Einstein-Hilbert action. However, it does not help in solving the cosmological constant problem.

3.7 Summary

A symmetry principle as explanation for the smallness of the cosmological constant in itself is very attractive. A viable mechanism that sets the cosmological constant to zero would be great progress, even if Λ would turn out to be nonzero. Since supersymmetry does not really seem to help, especially some form of scale invariance stands out as a serious option. Needless to say, it is hard to imagine how scale invariance could be used, knowing that the world around us is not scale invariant. Particle masses are small, but many orders of magnitude larger than the observed cosmological constant.

Another option might be that a symmetry condition enforcing ρ_{vac} equal to zero, could be reflected in a certain choice of boundary conditions. In such a scenario, the vacuum

state would satisfy different boundary conditions than excited states. The $x \rightarrow ix$ transformation of section (4) could be an example of this.

4 | Invariance Under Complex Transformations

In this chapter¹ we study a new symmetry argument that results in a vacuum state with strictly vanishing vacuum energy. This argument exploits the well-known feature that de Sitter and Anti- de Sitter space are related by analytic continuation. When we drop boundary and hermiticity conditions on quantum fields, we get as many negative as positive energy states, which are related by transformations to complex space. The proposal does not directly solve the cosmological constant problem, but explores a new direction that appears worthwhile.

4.1 Introduction

The scenario of this chapter has been introduced in [49], and is based on a symmetry with respect to a transformation towards imaginary values of the space-time coordinates: $x^\mu \rightarrow ix^\mu$. This symmetry entails a new definition of the vacuum state, as the unique state that is invariant under this transformation. Since curvature switches sign, this vacuum state must be associated with zero curvature, hence zero cosmological constant. The most striking and unusual feature of the symmetry is the fact that the *boundary conditions* of physical states are not invariant. Physical states obey boundary conditions when the real parts of the coordinates tend to infinity, not the imaginary parts. This is why all physical states, except the vacuum, must break the symmetry. We will argue that a vanishing cosmological constant could be a consequence of the specific boundary conditions of the vacuum, upon postulating this complex symmetry.

We do not address the issue of non-zero cosmological constant, nor the so-called cosmic coincidence problem. We believe that a symmetry which would set the cosmological constant to exactly zero would be great progress.

The fact that we are transforming real coordinates into imaginary coordinates implies, *inter alia*, that hermitean operators are transformed into operators whose hermiticity properties are modified. Taking the hermitean conjugate of an operator requires knowledge of the boundary conditions of a state. The transition from x to ix requires that the boundary conditions of the states are modified. For instance, wave functions Φ that are periodic in real space, are now replaced by waves that are exponential expressions of x , thus periodic in ix . But we are forced to do more than that. Also the creation and annihilation operators will transform, and their commutator algebra in complex space is not a priori clear; it requires careful study.

Thus, the symmetry that we are trying to identify is a symmetry of laws of nature *prior* to imposing any boundary conditions. Demanding invariance under $x_\mu \rightarrow x_\mu + a_\mu$ where

¹Based on our paper [50].

a_μ may be real or imaginary, violates boundary conditions at $\Phi \rightarrow \infty$, leaving only one state invariant: the *physical* vacuum.

4.2 Classical Scalar Field

To set our notation, consider a real, classical, scalar field $\Phi(x)$ in D space-time dimensions, with Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu \Phi)^2 - V(\Phi(x)) , \quad V(\Phi) = \frac{1}{2}m^2\Phi^2 + \lambda\Phi^4 . \quad (4.1)$$

Adopting the metric convention $(-+++)$, we write the energy-momentum tensor as

$$T_{\mu\nu}(x) = \partial_\mu \Phi(x) \partial_\nu \Phi(x) + g_{\mu\nu} \mathcal{L}(\Phi(x)) . \quad (4.2)$$

The Hamiltonian H is

$$H = \int d^{D-1} \vec{x} T_{00}(x) ; \quad T_{00} = \frac{1}{2}\Pi^2 + \frac{1}{2}(\vec{\partial}\Phi)^2 + V(\Phi) ; \quad \Pi(x) = \partial_0 \Phi(x) . \quad (4.3)$$

Write our transformation as $x^\mu = iy^\mu$, after which all coordinates are rotated in their complex planes such that y^μ will become real. For redefined notions in y space, we use subscripts or superscripts y , e.g., $\partial_\mu^y = i\partial_\mu$. The field in y space obeys the Lagrange equations with

$$\mathcal{L}_y = -\mathcal{L} = -\frac{1}{2}(\partial_\mu^y \Phi)^2 + V(\Phi) ; \quad (4.4)$$

$$T_{\mu\nu}^y = -T_{\mu\nu} = \partial_\mu^y \Phi(iy) \partial_\nu^y \Phi(iy) + g_{\mu\nu} \mathcal{L}_y(\Phi(iy)) . \quad (4.5)$$

The Hamiltonian in y -space is

$$H = -(i^{D-1})H_y , \quad H_y = \int d^{D-1} y T_{00}^y ; \quad (4.6)$$

$$T_{00}^y = \frac{1}{2}\Pi_y^2 + \frac{1}{2}(\vec{\partial}_y \Phi)^2 - V(\Phi) , \quad \Pi_y(y) = i\Pi(iy) . \quad (4.7)$$

If we keep only the mass term in the potential, $V(\Phi) = \frac{1}{2}m^2\Phi^2$, the field obeys the Klein-Gordon equation. In the real x -space, its solutions can be written as

$$\Phi(x, t) = \int d^{D-1} p (a(p)e^{i(px)} + a^*(p)e^{-i(px)}) , \quad (4.8)$$

$$\Pi(x, t) = \int d^{D-1} p p^0 (-ia(p)e^{i(px)} + ia^*(p)e^{-i(px)}) ; \quad (4.9)$$

$$p^0 = \sqrt{\vec{p}^2 + m^2} , \quad (px) \stackrel{\text{def}}{=} \vec{p} \cdot \vec{x} - p^0 t , \quad (4.10)$$

where $a(p)$ is just a c-number.

Analytically continuing these solutions to complex space, yields:

$$\Phi(iy, i\tau) = \int d^{D-1} q (a_y(q)e^{i(qy)} + \hat{a}_y(q)e^{-i(qy)}) , \quad (4.11)$$

$$\Pi_y(y, \tau) = i\Pi(iy, i\tau) = \int d^{D-1} q q^0 (-ia_y(q)e^{i(qy)} + i\hat{a}_y(q)e^{-i(qy)}) ; \quad (4.12)$$

$$q^0 = \sqrt{\vec{q}^2 - m^2} , \quad (qy) \stackrel{\text{def}}{=} \vec{q} \cdot \vec{y} - q^0 \tau . \quad (4.13)$$

The new coefficients could be analytic continuations of the old ones,

$$a_y(q) = (-i)^{D-1}a(p) , \quad \hat{a}_y(q) = (-i)^{D-1}a^*(q) , \quad p^\mu = -iq^\mu , \quad (4.14)$$

but this makes sense only if the $a(p)$ would not have singularities that we cross when shifting the integration contour. Note, that, since $D = 4$ is even, the hermiticity relation between $a_y(q)$ and $\hat{a}_y(q)$ is lost. We can now consider solutions where we restore them:

$$\hat{a}_y(q) = a_y^*(q) , \quad (4.15)$$

while also demanding convergence of the q integration. Such solutions would not obey acceptable boundary conditions in x -space, and the fields would be imaginary rather than real, so these are unphysical solutions. The important property that we concentrate on now, however, is that, according to Eq. (4.5), these solutions would have the opposite sign for $T_{\mu\nu}$.

Of course, the field in y -space appears to be tachyonic, since m^2 is negative. In most of our discussions we should put $m = 0$. A related transformation with the objective of $T_{\mu\nu} \rightarrow -T_{\mu\nu}$ was made by Kaplan and Sundrum in [184]. Non-Hermitian Hamiltonians were also studied by Bender et al. in for example [185, 186, 187, 188]. Another approach based on similar ideas which tries to forbid a cosmological constant can be found in [189].

4.3 Gravity

Consider Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = -8\pi GT_{\mu\nu}. \quad (4.16)$$

Writing

$$x^\mu = i y^\mu = i(\vec{y}, \tau) , \quad g_{\mu\nu}^y(y) \rightarrow g_{\mu\nu}(x = iy), \quad (4.17)$$

and defining the Riemann tensor in y space using the derivatives ∂_μ^y , we see that

$$R_{\mu\nu}^y = -R_{\mu\nu}(iy) . \quad (4.18)$$

Clearly, in y -space, we have the equation

$$R_{\mu\nu}^y - \frac{1}{2}g_{\mu\nu}^y R^y + \Lambda g_{\mu\nu}^y = +8\pi GT_{\mu\nu}(iy) = -8\pi GT_{\mu\nu}^y. \quad (4.19)$$

Thus, Einstein's equations are invariant except for the cosmological constant term.

A related suggestion was made in [190]. In fact, we could consider formulating the equations of nature in the full complex space $z = x + iy$, but then everything becomes complex. The above transformation is a one-to-one map from real space \mathfrak{R}^3 to the purely imaginary space \mathfrak{S}^3 , where again real equations emerge.

The transformation from real to imaginary coordinates naturally relates deSitter space with anti-deSitter space, or, a vacuum solution with positive cosmological constant to a vacuum solution with negative cosmological constant. Only if the cosmological constant is zero, a solution can map into itself by such a transformation. None of the excited states can have this invariance, because they have to obey boundary conditions, either in real space, or in imaginary space.

4.4 Non-relativistic Particle

The question is now, how much of this survives in a quantum theory. The simplest example to be discussed is the non-relativistic particle in one space dimension. Consider the Hamiltonian

$$H = \frac{p^2}{2m} + V(x), \quad (4.20)$$

where $p = -i\partial/\partial x$. Suppose that the function $V(x)$ obeys

$$V(x) = -V(ix), \quad V(x) = x^2 V_0(x^4), \quad (4.21)$$

with, for instance, $V_0(x^4) = e^{-\lambda x^4}$. Consider a wave function $|\psi(x)\rangle$ obeying the wave equation $H|\psi\rangle = E|\psi\rangle$. Then the substitution

$$x = iy, \quad p = -ip_y, \quad p_y = -i\frac{\partial}{\partial y}, \quad (4.22)$$

gives us a new function $|\psi(y)\rangle$ obeying

$$H_y|\psi(y)\rangle = -E|\psi(y)\rangle, \quad H_y = \frac{p_y^2}{2m} + V(y). \quad (4.23)$$

Thus, we have here a symmetry transformation sending the hamiltonian H into $-H$. Clearly, $|\psi(y)\rangle$ cannot in general be an acceptable solution to the usual Hamilton eigenvalue equation, since $|\psi(y)\rangle$ will not obey the boundary condition $|\psi(y)|^2 \rightarrow 0$ if $y \rightarrow \pm\infty$. Indeed, hermiticity, normalization, and boundary conditions will not transform as in usual symmetry transformations.

Yet, this symmetry is not totally void. If $V = 0$, a state $|\psi_0\rangle$ can be found that obeys both the boundary conditions at $x \rightarrow \pm\infty$ and $y \rightarrow \pm\infty$. It is the ground state, $\psi(x) = \text{constant}$. To be precise, this state is only normalizable if the boundary condition at $x \rightarrow \pm\infty$ is replaced by a periodic boundary condition $\psi(x) = \psi(x + L)$, which also removes the other solution satisfying $E = 0$, namely $\psi(x) = \alpha x$. (The important thing is that the bound $E = 0$ on the physical states follows by comparing the solutions on the real axis with the solutions on the imaginary axis.)

The state $\psi(x) = \text{constant}$ obeys both boundary conditions because of its invariance under transformations $x \rightarrow x + a$, where a can be any complex number. Because of our symmetry property, it obeys $E = -E$, so the energy of this state has to vanish. Since it is the only state with this property, it must be the ground state. Thus, we see that our complex symmetry may provide for a mechanism that generates a zero-energy ground state, of the kind that we are looking for in connection with the cosmological constant problem.

In general, if $V(x) \neq 0$, this argument fails. The reason is that the invariance under complex translations breaks down, so that no state can be constructed obeying all boundary conditions in the complex plane. In our treatment of the cosmological constant problem, we wish to understand the physical vacuum. It is invariant under complex translations, so there is a possibility that a procedure of this nature might apply.

As noted by Jackiw [191], there is a remarkable example in which the potential does not have to vanish. We can allow for any well-behaved function that depends only on $x^4 = y^4$. For example, setting $m = 1$,

$$V(x) = 2x^6 - 3x^2 = x^2(2x^4 - 3), \quad (4.24)$$

with ground state wavefunction $\exp(-x^4/2)$, indeed satisfies condition (4.21), which guarantees zero energy eigenvalue. Note that this restricts the transformation to be discrete, since otherwise it crosses the point $x = \sqrt{i}y$ where the potential badly diverges. Boundary conditions are still obeyed on the real and imaginary axis, but not for general complex values, see figure 4.1.

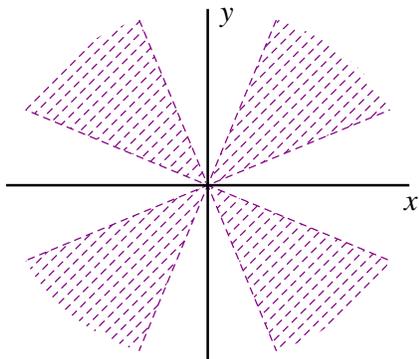


Figure 4.1: Region in complex space where the potential is well-defined; the shaded region indicates where boundary conditions are not obeyed.

Moreover, as Jackiw also pointed out (see also [192]), this example is intriguing since it reminds us of supersymmetry. Setting again $m = 1$ for clarity of notation, the Hamiltonian

$$H = \frac{1}{2}(p + iW')(p - iW') \quad (4.25)$$

with W the superpotential and a prime denoting a derivative with respect to the fields, has a scalar potential

$$V = \frac{1}{2}(W'W' - W''). \quad (4.26)$$

If W satisfies condition (4.21), the Hamiltonian possesses a zero energy eigenfunction e^{-W} , which obeys the correct boundary conditions in x and y . The Hamiltonian in this example is the bosonic portion of a supersymmetric Hamiltonian, so our proposal might be somehow related to supersymmetry.

We need to know what happens with hermiticity and normalizations. Assume the usual hermiticity properties of the bras, kets and the various operators in x space. How do these properties read in y space? We have

$$\begin{aligned} x &= x^\dagger & p &= p^\dagger, \\ y &= -y^\dagger & p_y &= -p_y^\dagger, \end{aligned} \quad (4.27)$$

but the commutator algebra is covariant under the transformation:

$$\begin{aligned} [p, x] &= -i & p &= -i\partial/\partial x, \\ [p_y, y] &= -i & p_y &= -i\partial/\partial y. \end{aligned} \quad (4.28)$$

Therefore, the wave equation remains the same locally in y as it is in x , but the boundary condition in y is different from the one in x . If we would replace the hermiticity properties of y and p_y in Eq. (4.27) by those of x and p , then we would get only states with $E \leq 0$.

4.5 Harmonic Oscillator

An instructive example is the $x \rightarrow y$ transformation, with $x = iy$, in the harmonic oscillator. The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2, \quad (4.29)$$

for which one introduces the conventional annihilation and creation operators a and a^\dagger :

$$a = \sqrt{\frac{m\omega}{2}} \left(x + \frac{ip}{m\omega} \right), \quad a^\dagger = \sqrt{\frac{m\omega}{2}} \left(x - \frac{ip}{m\omega} \right); \quad (4.30)$$

$$H = \omega \left(a^\dagger a + \frac{1}{2} \right). \quad (4.31)$$

In terms of the operators in y -space, we can write

$$a_y = \sqrt{\frac{m\omega}{2}} \left(y + \frac{ip_y}{m\omega} \right), \quad \hat{a}_y = \sqrt{\frac{m\omega}{2}} \left(y - \frac{ip_y}{m\omega} \right) = -a_y^\dagger; \quad (4.32)$$

$$a_y = -ia^\dagger, \quad \hat{a}_y = -ia, \quad H = -\omega \left(\hat{a}_y a_y + \frac{1}{2} \right). \quad (4.33)$$

If one were to replace the correct hermitian conjugate of a_y by \hat{a}_y instead of $-a_y^\dagger$, then the Hamiltonian (4.33) would take only the eigenvalues $H = -H_y = \omega(-n - \frac{1}{2})$. Note that these form a natural continuation of the eigenstates $\omega(n + \frac{1}{2})$, as if n were now allowed only to be a negative integer.

The ground state, $|0\rangle$ is not invariant. In x -space, the y ground state would be the non-normalizable state $\exp(+\frac{1}{2}m\omega x^2)$, which of course would obey ‘good’ boundary conditions in y -space.

4.6 Second Quantization

The examples of the previous two sections, however, are not the transformations that are most relevant for the cosmological constant. We wish to turn to imaginary coordinates, but not to imaginary oscillators. We now turn our attention to second-quantized particle theories, and we know that the vacuum state will be invariant, at least under all complex translations. Not only the hermiticity properties of field operators are modified in the transformation, but now also the commutation rules are affected. A scalar field $\Phi(x)$ and its conjugate, $\Pi(x)$, often equal to $\dot{\Phi}(x)$, normally obey the commutation rules

$$[\Pi(\vec{x}, t), \Phi(\vec{x}', t)] = -i\delta^3(\vec{x} - \vec{x}'), \quad (4.34)$$

where the Dirac deltafunction $\delta(x)$ may be regarded as

$$\delta(x) = \sqrt{\frac{\lambda}{\pi}} e^{-\lambda x^2}, \quad (4.35)$$

in the limit $\lambda \uparrow \infty$. If \vec{x} is replaced by $i\vec{y}$, with \vec{y} real, then the commutation rules are

$$[\Pi(i\vec{y}, t), \Phi(i\vec{y}', t)] = -i\delta^3(i(\vec{y} - \vec{y}')) , \quad (4.36)$$

but, in Eq. (4.35) we see two things happen:

(i) This delta function does not go to zero unless its argument x lies in the right or left quadrant of Fig. 4.2. Now, this can be cured if we add an imaginary part to λ , namely $\lambda \rightarrow -i\mu$, with μ real. Then the function (4.35) exists if $x = r e^{i\theta}$, with $0 < \theta < \frac{1}{2}\pi$. But then,

(ii) If $x = iy$, the sign of μ is important. If $\mu > 0$, replacing $x = iy$, the delta function becomes

$$\delta(iy) = \sqrt{\frac{-i\mu}{\pi}} e^{-i\mu y^2} \rightarrow -i\delta(y) , \quad (4.37)$$

which would be $+i\delta(y)$ had we chosen the other sign for μ .

We conclude that the sign of the square root in Eq. (4.35) is ambiguous.

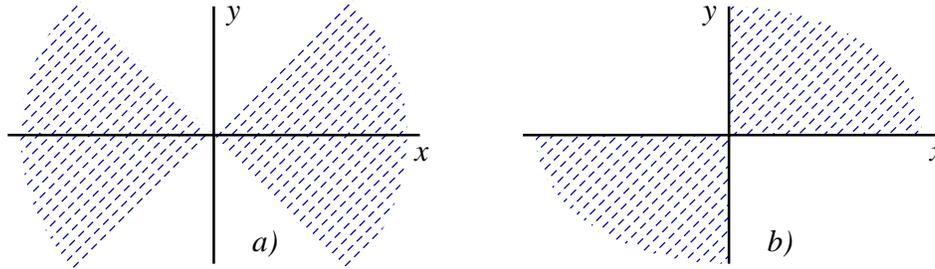


Figure 4.2: Region in complex space where the Dirac delta function is well-defined, (a) if λ is real, (b) if μ is real and positive.

There is another way to phrase this difficulty. The commutation rule (4.34) suggests that either the field $\Phi(\vec{x}, t)$ or $\Pi(\vec{x}, t)$ must be regarded as a distribution. Take the field Π . Consider test functions $f(\vec{x})$, and write

$$\Pi(f, t) \stackrel{\text{def}}{=} \int f(\vec{x}) \Pi(\vec{x}, t) d^3\vec{x} ; \quad [\Pi(f, t), \Phi(\vec{x}, t)] = -if(\vec{x}) . \quad (4.38)$$

As long as \vec{x} is real, the integration contour in Eq. (4.38) is well-defined. If, however, we choose $x = iy$, the contour must be taken to be in the complex plane, and if we only wish to consider real y , then the contour must be along the imaginary axis. This would be allowed if $\Pi(\vec{x}, y)$ is holomorphic for complex \vec{x} , and the end points of the integration contour should not be modified.

For simplicity, let us take space to be one-dimensional. Assume that the contour becomes as in Fig. 4.3a. In the y space, we have

$$\Pi(f, t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(iy) \Pi(iy) d(iy) ; \quad [\Pi(f, t), \Phi(iy, t)] = -if(iy) . \quad (4.39)$$

so that

$$[\Pi(iy, t), \Phi(iy', t)] = -\delta(y - y') . \quad (4.40)$$

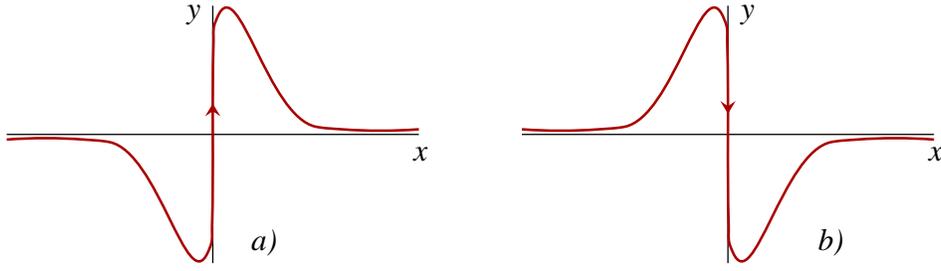


Figure 4.3: Integration contour for the commutator algebra (4.38), (a) and (b) being two distinct choices.

Note now that we could have chosen the contour of Fig. 4.3b instead. In that case, the integration goes in the opposite direction, and the commutator algebra in Eq. (4.40) receives the opposite sign. Note also that, if we would be tempted to stick to one rule only, the commutator algebra would receive an overall minus sign if we apply the transformation $x \rightarrow iy$ twice.

The general philosophy is now that, with these new commutation relations in y -space, we could impose conventional hermiticity properties in y -space, and then consider states as representations of these operators. How do individual states then transform from x -space to y -space or *vice versa*? We expect to obtain non-normalizable states, but the situation is worse than that. Let us again consider one space-dimension, and begin with defining the annihilation and creation operators $a(p)$ and $a^\dagger(p)$ in x -space:

$$\Phi(x, t) = \int \frac{dp}{\sqrt{2\pi \cdot 2p^0}} (a(p)e^{i(px)} + a^\dagger(p)e^{-i(px)}) , \quad (4.41)$$

$$\Pi(x, t) = \int \frac{dp\sqrt{p^0}}{\sqrt{2} \cdot 2\pi} (-ia(p)e^{i(px)} + ia^\dagger(p)e^{-i(px)}) \quad (4.42)$$

$$p^0 = \sqrt{\vec{p}^2 + m^2} , \quad (px) \stackrel{\text{def}}{=} \vec{p} \cdot \vec{x} - p^0 t \quad (4.43)$$

$$a(p) = \int \frac{dx}{\sqrt{2\pi \cdot 2p^0}} (p^0\Phi(x, t) + i\Pi(x, t)) e^{-i(px)} , \quad (4.44)$$

$$a^\dagger(p) = \int \frac{dx}{\sqrt{2\pi \cdot 2p^0}} (p^0\Phi(x, t) - i\Pi(x, t)) e^{i(px)} . \quad (4.45)$$

Insisting that the commutation rules $[a(p), a^\dagger(p')] = \delta(p - p')$ should also be seen in y -space operators:

$$[a_y(q), \hat{a}_y(q')] = \delta(q - q') , \quad (4.46)$$

we write, assuming $p^0 = -iq^0$ and $\Pi = -i\partial\Phi/\partial\tau$ for free fields,

$$\Phi(iy, i\tau) = \int \frac{dq}{\sqrt{2\pi \cdot 2q^0}} (a_y(q)e^{i(qy)} + \hat{a}_y(q)e^{-i(qy)}) \quad (4.47)$$

$$\Pi(iy, i\tau) = \int \frac{dq\sqrt{q^0}}{\sqrt{2} \cdot 2\pi} (-a_y(q)e^{i(qy)} + \hat{a}_y(q)e^{-i(qy)}) , \quad (4.48)$$

$$q^0 = \sqrt{\vec{q}^2 - m^2}, \quad (qy) \stackrel{\text{def}}{=} \vec{q} \cdot \vec{y} - q^0 \tau \quad (4.49)$$

$$a_y(q) = \int \frac{dy}{\sqrt{2\pi \cdot 2q^0}} (q^0 \Phi(iy, i\tau) - \Pi(iy, i\tau)) e^{i(qy)}, \quad (4.50)$$

$$\hat{a}_y(q) = \int \frac{dy}{\sqrt{2\pi \cdot 2q^0}} (q^0 \Phi(iy, i\tau) + \Pi(iy, i\tau)) e^{-i(qy)}, \quad (4.51)$$

so that the commutator (4.46) agrees with the field commutators (4.40). In most of our considerations, we will have to take $m = 0$; we leave m in our expressions just to show its sign switch.

In x -space, the fields Φ and π are real, and the exponents in Eqs (4.47)—(4.51) are all real, so the hermiticity relations are $a_y^\dagger = a_y$ and $\hat{a}_y^\dagger = \hat{a}_y$. As in the previous sections, we replace this by

$$\hat{a}_y = a_y^\dagger. \quad (4.52)$$

The Hamiltonian for a free field reads

$$\begin{aligned} H &= i \int_{-\infty}^{\infty} dy \left(\frac{1}{2} \Pi(iy)^2 - \frac{1}{2} (\partial_y \Phi(iy))^2 + \frac{1}{2} m^2 \Phi(iy)^2 \right) = \\ &= -i \int dq q^0 \left(\hat{a}_y(q) a_y(q) + \frac{1}{2} \right) = -i \int dq q^0 (n + \frac{1}{2}). \end{aligned} \quad (4.53)$$

Clearly, with the hermiticity condition (4.52), the Hamiltonian became purely imaginary, as in Section 4.2. Also, the zero point fluctuations still seem to be there. However, we have not yet addressed the operator ordering. Let us take a closer look at the way individual creation and annihilation operators transform. We now need to set $m = 0$, $p^0 = |p|$, $q^0 = |q|$. In order to compare the creation and annihilation operators in real space-time with those in imaginary space-time, substitute Eqs. (4.47) and (4.48) into (4.44), and the converse, to obtain

$$a(p) = \int \int \frac{dx dq}{2\pi \sqrt{4p^0 q^0}} \left\{ (p^0 - iq^0) a_y(q) e^{(q-ip)x} + (p^0 + iq^0) \hat{a}_y(q) e^{(-q-ip)x} \right\} \quad (4.54)$$

$$a_y(q) = \int \int \frac{dy dp}{2\pi \sqrt{4p^0 q^0}} \left\{ (q^0 + ip^0) a(p) e^{(-iq-p)y} + (q^0 - ip^0) a^\dagger(p) e^{(-iq+p)y} \right\} \quad (4.55)$$

The difficulty with these expressions is the fact that the x - and the y -integrals diverge. We now propose the following procedure. Let us limit ourselves to the case that, in Eqs. (4.50) and (4.51), the y -integration is over a finite box only: $|y| < L$, in which case $a_y(q) \sqrt{2q^0}$ will be an entire analytic function of q . Then, in Eq. (4.54), we can first shift the integration contour in complex q -space by an amount ip up or down, and subsequently rotate the x -integration contour to obtain convergence. Now the square roots occurring explicitly in Eqs. (4.54) and (4.55) are merely the consequence of a choice of normalization, and could be avoided, but the root in the definition of p^0 and q^0 are much more problematic. In principle we could take any of the two branches of the roots. However, in our transformation procedure we actually *choose* $q^0 = -ip^0$ and the second parts of Eqs. (4.54) and (4.55) simply cancel out. Note that, had we taken the other sign, i.e. $q^0 = ip^0$, this would have affected the expression for $\Phi(iy, i\tau)$ such,

that we would still end up with the same final result. In general, the x -integration yields a delta function constraining q to be $\pm ip$, but q^0 is chosen to be on the branch $-ip^0$, in both terms of this equation (q^0 normally does not change sign if q does). Thus, we get, from Eqs. (4.54) and (4.55), respectively,

$$a(p) = i^{1/2} a_y(q), \quad q = ip, \quad q^0 = ip^0, \quad (4.56)$$

$$a_y(q) = i^{-1/2} a(p), \quad p = -iq, \quad p^0 = -iq^0, \quad (4.57)$$

so that $a(p)$ and $a_y(q)$ are analytic continuations of one another. Similarly,

$$a^\dagger(p) = i^{1/2} \hat{a}_y(q), \quad \hat{a}_y(q) = i^{-1/2} a^\dagger(p), \quad p = -iq, \quad p^0 = -iq^0. \quad (4.58)$$

There is no Bogolyubov mixing between a and a^\dagger . Note that these expressions agree with the transformation law of the Hamiltonian (4.53).

Now that we have a precisely defined transformation law for the creation and annihilation operators, we can find out how the states transform. The vacuum state $|0\rangle$ is defined to be the state upon which all annihilation operators vanish. We now see that this state is invariant under all our transformations. Indeed, because there is no Bogolyubov mixing, all N particle states transform into N particle states, with N being invariant. The vacuum is invariant because 1) unlike the case of the harmonic oscillator, Section 4.5, creation operators transform into creation operators, and annihilation operators into annihilation operators, and because 2) the vacuum is translation invariant.

The Hamiltonian is transformed into $-i$ times the Hamiltonian (in the case $D = 2$); the energy density T_{00} into $-T_{00}$, and since the vacuum is the only state that is invariant, it must have $T_{00} = 0$ and it must be the only state with this property.

4.7 Pure Maxwell Fields

This can now easily be extended to include the Maxwell action as well. In flat spacetime:

$$S = - \int d^3x \frac{1}{4} F_{\mu\nu}(x) F^{\mu\nu}(x), \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (4.59)$$

The action is invariant under gauge transformations of the form

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \xi(x). \quad (4.60)$$

Making use of this freedom, we impose the Lorentz condition $\partial_\mu A^\mu = 0$, such that the equation of motion $\partial_\mu F^{\mu\nu} = 0$ becomes $\square A^\mu = 0$. As is well known, this does not completely fix the gauge, since transformations like (4.60) are still possible, provided $\square \xi = 0$. This remaining gauge freedom can be used to set $\nabla \cdot \vec{A} = 0$, denoted Coulomb gauge, which sacrifices manifest Lorentz invariance. The commutation relations are

$$[E^i(x, t), A_j(x', t)] = i \delta_{ij}^{tr}(\vec{x} - \vec{x}'), \quad (4.61)$$

where

$$E^k = \frac{\partial \mathcal{L}}{\partial \dot{A}_k} = -\dot{A}^k - \frac{\partial A_0}{\partial x^k}, \quad (4.62)$$

is the momentum conjugate to A^k , which we previously called Π , but it is here just a component of the electric field. The transverse delta function is defined as

$$\delta_{ij}^{tr}(\vec{x} - \vec{x}') \equiv \int \frac{d^3p}{(2\pi)^3} e^{i\vec{p}(\vec{x}-\vec{x}')} \left(\delta_{ij} - \frac{p_i p_j}{\vec{p}^2} \right), \quad (4.63)$$

such that its divergence vanishes. In Coulomb gauge, \vec{A} satisfies the wave equation $\square \vec{A} = 0$, and we write

$$\vec{A}(x, t) = \int \frac{d^3p}{(2\pi)^3 \sqrt{2p^0}} \sum_{\lambda=1}^2 \vec{\varepsilon}(p, \lambda) \left(a(p, \lambda) e^{i(p x)} + a^\dagger(p, \lambda) e^{-i(p x)} \right), \quad (4.64)$$

where $\vec{\varepsilon}(p, \lambda)$ is the polarization vector of the gauge field, which satisfies $\vec{\varepsilon} \cdot \vec{p} = 0$ from the Coulomb condition $\nabla \cdot \vec{A} = 0$. Moreover, the polarization vectors can be chosen to be orthogonal $\vec{\varepsilon}(p, \lambda) \cdot \vec{\varepsilon}(p, \lambda') = \delta_{\lambda\lambda'}$ and satisfy a completeness relation

$$\sum_{\lambda} \varepsilon_m(p, \lambda) \varepsilon_n(p, \lambda) = \left(\delta_{mn} - \frac{p_m p_n}{\vec{p}^2} \right). \quad (4.65)$$

The commutator between the creation and annihilation operators becomes

$$[a(p, \lambda), a^\dagger(p', \lambda')] = \delta(\vec{p} - \vec{p}') \delta_{\lambda\lambda'}, \quad (4.66)$$

in which the polarization vectors cancel out due to their completeness relation.

In complex space, the field A_μ thus transforms analogously to the scalar field, with the only addition that the polarization vectors $\vec{\varepsilon}_\mu(p)$ will now become function of complex momentum \vec{q} . However, since they do not satisfy a particular algebra, like the creation and annihilation operators, they do not cause any additional difficulties. The commutation relations between the creation and annihilation operators behave similarly as in the scalar field case, since the second term in the transverse delta function (4.63), and the polarization vector completeness relation (4.65), is invariant when transforming to complex momentum.

Thus we find

$$F_{\mu\nu}(x, t) F^{\mu\nu}(x, t) \rightarrow -F_{\mu\nu}(iy, i\tau) F^{\mu\nu}(iy, i\tau), \quad (4.67)$$

and again T_{00} flips sign, as the energy-momentum tensor reads:

$$T_{\mu\nu} = -F_{\mu\alpha} F_\nu^\alpha + \frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} \eta_{\mu\nu}. \quad (4.68)$$

In term of the E and B fields, which are given by derivatives of A_μ , $E_i = F_{0i}$, $B_k = \frac{1}{2} \varepsilon_{ijk} F_{jk}$, we have:

$$T_{00} = \frac{1}{2} (E^2 + B^2) \rightarrow -T_{00}, \quad (4.69)$$

which indicates that the electric and magnetic fields become imaginary. A source term $J^\mu A_\mu$ can also be added to the action (4.59), if one imposes that $J^\mu \rightarrow -J^\mu$, in which case the Maxwell equations $\partial_\mu F^{\mu\nu} = J^\nu$ are covariant.

Implementing gauge invariance in imaginary space is also straightforward. The Maxwell action and Maxwell equations are invariant under $A_\mu(x, t) \rightarrow A_\mu(x, t) + \partial_\mu \xi(x, t)$. In complex spacetime this becomes

$$A_\mu(iy, i\tau) \rightarrow A_\mu(iy, i\tau) - i\partial_\mu(y, \tau)\xi(iy, i\tau) \quad (4.70)$$

and the Lorentz condition

$$\partial_\mu(x, t)A^\mu(x, t) = 0 \quad \rightarrow \quad -i\partial(y, \tau)A^\mu(iy, i\tau). \quad (4.71)$$

In Coulomb gauge the polarization vectors satisfy

$$\vec{\varepsilon}(q) \cdot \vec{q} = 0, \quad (4.72)$$

with imaginary momentum q .

Unfortunately, the Maxwell field handled this way will not be easy to generalize to Yang-Mills fields. The Yang-Mills field has cubic and quartic couplings, and these will transform with the wrong sign. One might consider forcing vector potentials to transform just like space-time derivatives, but then the kinetic term emerges with the wrong sign. Alternatively, one could suspect that the gauge group, which is compact in real space, would become non-compact in imaginary space, but this also leads to undesirable features.

4.8 Relation with Boundary Conditions

All particle states depend on boundary conditions, usually imposed on the real axis. One could therefore try to simply view the $x \rightarrow ix$ symmetry as a one-to-one mapping of states with boundary conditions imposed on $\pm x \rightarrow \infty$ to states with boundary conditions imposed on imaginary axis $\pm ix \rightarrow \infty$. At first sight, this mapping transforms positive energy particle states into negative energy particle states. The vacuum, not having to obey boundary conditions would necessarily have zero energy. However, this turns out not to be sufficient.

Solutions to the Klein-Gordon equation, with boundary conditions imposed on imaginary coordinates are of the form:

$$\Phi_{\text{im}}(x, t) = \int \frac{dp}{\sqrt{2\pi \cdot 2p^0}} (a(p)e^{(px)} + \hat{a}(p)e^{-(px)}) , \quad p^0 = \sqrt{p^2 + m^2} \quad (4.73)$$

written with the subscript ‘im’ to remind us that this is the solution with boundary conditions on the imaginary axis. With these boundary conditions, the field explodes for real valued $x \rightarrow \pm\infty$, whereas for the usual boundary conditions, imposed on the real axis, the field explodes for $ix \rightarrow \pm\infty$. Note that for non-trivial a and \hat{a} , this field now has a non-zero complex part on the real axis, if one insists that the second term is the Hermitian conjugate of the first, as is usually the case. This is a necessary consequence of this setup. However, we insist on writing $\hat{a} = a^\dagger$ and, returning to three spatial dimensions, we write for $\Phi_{\text{im}}(x, t)$ and $\Pi_{\text{im}}(x, t)$:

$$\Phi_{\text{im}}(\vec{x}, t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p^0}} (a_p e^{(px)} + a_p^\dagger e^{-(px)}) ,$$

$$\begin{aligned}\dot{\Phi}_{\text{im}}(\vec{x}, t) = \Pi_{\text{im}}(\vec{x}, t) &= \int \frac{d^3 p}{(2\pi)^3} (-) \sqrt{\frac{p^0}{2}} (a_p e^{(px)} - a_p^\dagger e^{-(px)}), \\ p^0 &= \sqrt{\vec{p}^2 + m^2}, \quad (px) \stackrel{\text{def}}{=} \vec{p} \cdot \vec{x} - p^0 t,\end{aligned}\quad (4.74)$$

and impose the normal commutation relations between a and a^\dagger :

$$[a_p, a_{p'}^\dagger] = (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{p}'). \quad (4.75)$$

Using Eqn. (4.75), the commutator between Φ_{im} and Π_{im} at equal times, becomes:

$$[\Phi_{\text{im}}(\vec{x}), \Pi_{\text{im}}(\vec{x}')] = \delta^{(3)}(\vec{x} - \vec{x}'), \quad (4.76)$$

which differs by a factor of i from the usual relation, and by a minus sign, compared to Eqn. (4.40). The energy-momentum tensor is given by

$$T_{\mu\nu \text{im}} = \partial_\mu \Phi_{\text{im}} \partial_\nu \Phi_{\text{im}} - \frac{1}{2} \eta_{\mu\nu} \partial^k \Phi_{\text{im}} \partial_k \Phi_{\text{im}}, \quad (4.77)$$

and thus indeed changes sign, as long as one considers only those contributions to a Hamiltonian that contain products of a and a^\dagger :

$$H_{\text{im}}^{\text{diag}} = \int \frac{d^3 p}{(2\pi)^3} p^0 \left(-a_p^\dagger a_p - \frac{1}{2} [a_p, a_p^\dagger] \right) = -H. \quad (4.78)$$

However, the remaining parts give a contribution that is rapidly diverging on the imaginary axis

$$T_{\mu\nu \text{im}}^{\text{non-diag}} = a^2 e^{2(px)} + (a^\dagger)^2 e^{-2(px)}, \quad (4.79)$$

but which blows up for $\pm x \rightarrow \infty$. Note that when calculating vacuum expectation values, these terms give no contribution.

To summarize, one can only construct such a symmetry, changing boundary conditions from real to imaginary coordinates, in a very small box. This was to be expected, since we are comparing hyperbolic functions with their ordinary counterparts, $\sinh(x)$ vs. $\sin(x)$, and they are only identical functions in a small neighborhood around the origin.

4.9 Related Symmetries

So far, we have discussed the implications of the transformation $x_\mu \rightarrow ix_\mu$ and $g_{\mu\nu} \rightarrow g_{\mu\nu}$. Alternatively, one might want to consider the related discrete transformation $g_{\mu\nu} \rightarrow -g_{\mu\nu}$. The combined transformation $x_\mu \rightarrow ix_\mu$ and $g_{\mu\nu} \rightarrow -g_{\mu\nu}$ is a coordinate transformation to imaginary spacetime. Under the transformation $g_{\mu\nu} \rightarrow -g_{\mu\nu}$, the connection and Ricci tensor are invariant, the Ricci scalar is not:

$$\begin{aligned}\Gamma_{\mu\nu}^\lambda &\rightarrow \Gamma_{\mu\nu}^\lambda \\ R_{\mu\nu} &\rightarrow R_{\mu\nu} \\ R &\rightarrow -R\end{aligned}\quad (4.80)$$

Therefore, promoting this transformation to a symmetry, a cosmological constant term in the Einstein-Hilbert action is no longer allowed. Only terms that transform with a minus sign are allowed in the action. Einstein's equation transforms as:

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} &= -8\pi GT_{\mu\nu} \rightarrow \\ R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} &= -8\pi GT_{\mu\nu} \end{aligned} \quad (4.81)$$

and is invariant up to the cosmological constant term, assuming invariance of $T_{\mu\nu}$ under this transformation. Note the sign change compared to the action. In the field equations we should only allow terms that are invariant under the proposed transformation, whereas in the action only those terms are allowed that transform with a minus sign.

Adding matter however, again causes problems. Consider for example the Maxwell action:

$$S_{Max} = \int d^4x \sqrt{-g} \left(-\frac{1}{4}g^{\mu\alpha}g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta} + g^{\mu\nu}J_\mu A_\nu \right) \quad (4.82)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (4.83)$$

This shows that the kinetic term for the EM-field is invariant, but that the coupling of the source term J_μ is not. A kinetic term for gauge bosons therefore would not be allowed, unless one also demands

$$A_\mu \rightarrow \pm i A_\mu, \quad J_\mu \rightarrow \mp J_\mu. \quad (4.84)$$

With these additional transformations, the kinetic term is allowed as well as the source term.

The standard energy-momentum tensor for this action is:

$$T_{\mu\nu} = -F_{\mu\alpha}F_\nu^\alpha + \left(\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - J^\alpha A_\alpha \right) g_{\mu\nu}, \quad (4.85)$$

where the terms originating from the variation of the kinetic term are invariant. One minus sign because of the contraction with $g_{\mu\nu}$ and one minus sign from the $A_\mu \rightarrow \pm i A_\mu$ transformation. This was to be expected, since the energy momentum tensor is derived from the variation of the matter action, with respect to $g_{\mu\nu}$: $T_{\mu\nu} \propto \delta\mathcal{L}/\delta g_{\mu\nu}$, neutralizing the minus sign in the action.

For a scalar field:

$$S = \int d^4x \sqrt{-g} (g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - m^2 \phi^2) \quad (4.86)$$

we have to face the same situation as in the previous section; the kinetic part is allowed, but the potential term is not. That mass squared terms have to transform to minus mass squared can of course easily be seen just from special relativity:

$$p^2 = g^{\mu\nu} p_\mu p_\nu = m^2 \rightarrow -m^2. \quad (4.87)$$

For fermions the situation is a bit more complicated. We normally require:

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}, \quad (4.88)$$

so under the transformation $g^{\mu\nu} \rightarrow -g^{\mu\nu}$ the gamma matrices would instead have to be defined as $\gamma^\mu \rightarrow \pm i\gamma^\mu$.

The free Dirac Lagrangean in curved spacetime reads:

$$\mathcal{L} = \sqrt{-g} (i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi) \quad (4.89)$$

with $D_\mu = \partial_\mu - \frac{i}{4}\omega_\mu^{mn}\sigma_{mn}$ and σ_{mn} is defined in terms of flat space gamma-matrices: $\sigma_{mn} = i(\gamma_n\gamma_m - \gamma_m\gamma_n)/2$. Moreover, ω_μ^{mn} is the spin-connection.

This appears to show the same behavior as the free scalar field. The kinetic term transforms with a minus sign (because of the two gamma-matrices in it), whereas the mass term again forces us to consider imaginary masses.

Furthermore, the fermionic current:

$$J^\mu \equiv \bar{\psi}\gamma^\mu\psi \rightarrow -J^\mu, \quad (4.90)$$

becomes negative definite as one might have expected.

The coupling between spin-1 and spin- $\frac{1}{2}$ particles is normally obtained by introducing the covariant derivative:

$$p_\mu \rightarrow p_\mu - eA_\mu \quad \text{in QM} \quad i\partial_\mu \rightarrow i\partial_\mu - eA_\mu \quad (4.91)$$

and this is independent of the sign of the metric tensor. However, the requirement $A_\mu \rightarrow \pm iA_\mu$, appears to force us to consider imaginary charges, which was to be expected, because of (4.84) and consistency of Maxwell's equations.

The main difficulty in this approach therefore is that masses as well as general potential terms like $\lambda\phi^4$ and other interaction, are not allowed. This transformation therefore is similar to that discussed in the previous sections.

Yet another alternative is to consider performing both transformations, i.e. $g_{\mu\nu} \rightarrow -g_{\mu\nu}$ and $x_\mu \rightarrow ix_\mu$. Under this transformation the different components of Einstein's equation transform as:

$$\begin{aligned} \Gamma &\rightarrow -i\Gamma \\ R_{\mu\nu} &\rightarrow -R_{\mu\nu} \\ R &\rightarrow R \\ \Lambda g_{\mu\nu} &\rightarrow -\Lambda g_{\mu\nu} \end{aligned} \quad (4.92)$$

Therefore, this symmetry does not forbid a cosmological constant term in the Einstein-Hilbert action. Einstein's equation transforms as:

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} &= -8\pi GT_{\mu\nu} \rightarrow \\ -R_{\mu\nu} + \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} &= +8\pi GT_{\mu\nu}, \end{aligned} \quad (4.93)$$

and is invariant, assuming that also $T_{\mu\nu}$ transforms with a minus sign. This does not seem to be helpful to control the cosmological constant term. Moreover, the scalar field action is invariant, but the Maxwell action (4.82) is not.

This is a peculiar transformation, since one would perhaps naively expect that ordinary flat space quantum field theory would be invariant under the combined transformations $\eta_{\mu\nu} \rightarrow -\eta_{\mu\nu}$ and $x_\mu \rightarrow ix_\mu$.

4.9.1 Energy \rightarrow $-$ Energy

Another approach in which negative energy states are considered has been recently proposed in [184]. Here the discrete symmetry $E \rightarrow -E$ is imposed explicitly on the matter fields by adding to the Lagrangian an identical copy of the normal matter fields, but with an overall minus sign:

$$\mathcal{L} = \sqrt{-g} \left(M_{Pl}^2 R - \Lambda_0 + \mathcal{L}_{\text{matt}}(\psi, D_\mu) - \mathcal{L}_{\text{matt}}(\hat{\psi}, D_\mu) + \dots \right), \quad (4.94)$$

where Λ_0 is the bare cosmological constant. The Lagrangian with fields $\hat{\psi}$ occurring with the wrong sign is referred to as the ghost sector. The two matter sectors have equal but opposite vacuum energies, and therefore cancelling contributions to the cosmological constant.

Crucial in this reasoning is that there is no coupling other than gravitational between the normal matter fields and their ghost counterparts, otherwise the Minkowski vacuum would not be stable. Note that this approach is quite similar to Linde's antipodal symmetry, we discussed in section (3.5), although there the symmetry does not involve gravity and the coupling of the two sectors is suppressed, simply because they live in different universes. Here any particle to ghost-particle coupling constant g in e.g a term $g\phi^2\hat{\phi}^2$ has to be postulated to be exactly zero.

Moreover, in order to ensure stability of the vacuum, and prevent a rapid decay of the vacuum due to negative energy particles, also new Lorentz symmetry violating physics is required to suppress processes where normal matter particles and ghosts emerge from the vacuum; a process mediated by an off-shell graviton. There is no other way to suppress the phase space integral over ghost momenta [193]. In addition, one also has to assume that the ghost sector is rather empty, compared to the normal matter sector, in order not to spoil standard cosmology with such an exotic type dark matter.

The gravitational coupling moreover has to be sufficiently small in order to suppress the gravitationally induced interactions between the two sectors and to make sure that the quantum gravitational corrections to the bare cosmological constant are kept very small. Processes for example in which out of nothing two gravitons and two ghosts appear. It is therefore necessary to impose a UV cutoff on these contributions of order 10^{-3} eV, corresponding to a length scale of about 100 microns².

4.10 Summary

It is natural to ascribe the extremely tiny value of the cosmological constant to some symmetry. Until now, the only symmetry that showed promises in this respect has been supersymmetry. It is difficult, however, to understand how it can be that supersymmetry is obviously strongly broken by all matter representations, whereas nevertheless the vacuum state should respect it completely. This symmetry requires the vacuum fluctuations of bosonic fields to cancel very precisely against those of the fermionic field, and it is hard to see how this can happen when fermionic and bosonic fields have such dissimilar spectra.

²In section (7.5) a proposal by one of the authors of [184] is discussed in which such a cutoff is argued to arise from the graviton not being a point-like particle but having this finite size.

The symmetry proposed in this chapter is different. It is suspected that the field equations themselves have a larger symmetry than the boundary conditions for the solutions. It is the boundary conditions, and the hermiticity conditions on the fields, that force all physical states to have positive energies. If we drop these conditions, we get as many negative energy as positive energy states, and indeed, there may be a symmetry relating positive energy with negative energy. This is the most promising beginning of an argument why the vacuum state must have strictly vanishing gravitational energy.

The fact that the symmetry must relate real to imaginary coordinates is suggested by the fact that De Sitter and Anti-De Sitter space are related by analytic continuation, and that their cosmological constants have opposite sign.

Unfortunately, it is hard to see how this kind of symmetry could be realized in the known interaction types seen in the sub-atomic particles. At first sight, all mass terms are forbidden. However, we could observe that all masses in the Standard Model are due to interactions, and it could be that fields with positive mass squared are related to tachyonic fields by our symmetry. The one scalar field in the Standard Model is the Higgs field. Its self interaction is described by a potential $V_1(\Phi) = \frac{1}{2}\lambda(\Phi^\dagger\Phi - F^2)^2$, and it is strongly suspected that the parameter λ is unnaturally small. Our symmetry would relate it to another scalar field with opposite potential: $V_2(\Phi_2) = -V_1(\Phi_2)$. Such a field would have no vacuum expectation value, and, according to perturbation theory, a mass that is the Higgs mass divided by $\sqrt{2}$. Although explicit predictions would be premature, this does suggest that a theory of this kind could make testable predictions, and it is worth-while to search for scalar fields that do not contribute to the Higgs mechanism at LHC, having a mass somewhat smaller than the Higgs mass. We are hesitant with this particular prediction because the negative sign in its self interaction potential could lead to unlikely instabilities, to be taken care of by non-perturbative radiative corrections.

The symmetry we studied in this chapter would set the vacuum energy to zero and has therefore the potential to explain a vanishing cosmological constant. In light of the recent discoveries that the universe appears to be accelerating [7, 10, 17], one could consider a slight breaking of this symmetry. This is a non-trivial task that we will have to postpone to further work. Note however, that our proposal would only nullify exact vacuum energy with equation of state $w = -1$. Explaining the acceleration of the universe with some dark energy component other than a cosmological constant, quintessence for example, therefore is not ruled out within this framework.

The considerations reported about in this chapter will only become significant if, besides Maxwell fields, we can also handle Yang-Mills fields, fermions, and more complicated interactions. As stated, Yang-Mills fields appear to lead to difficulties. Fermions, satisfying the linear Dirac equation, can be handled in this formalism. Just as is the case for scalar fields, one finds that mass terms are forbidden for fermions, but we postpone further details to future work. Radiative corrections and renormalization group effects will have to be considered. To stay in line with our earlier paper, we still consider arguments of this nature to explain the tiny value of the cosmological constant unlikely to be completely successful, but every alley must be explored, and this is one of them.

5 | Type II: Back-reaction

It may be argued that any cosmological constant will be automatically cancelled, or screened, to a very small value by back-reaction effects on an expanding space. The effective cosmological constant then is very small today, simply because the universe is rather old. Often these effects are studied in an inflationary background, where a primordial cosmological constant is most dominant. In an inflationary background, the effects of fields that are not conformally invariant will be most dominant

Two familiar massless particles are not conformally invariant, massless minimally coupled scalars and gravitons. In this chapter we will be concerned with back-reaction by a scalar field, discuss the general setting of the renormalization group running of a cosmological constant. In the next chapter we will discuss in detail a back-reaction model in a purely quantum gravitation setting, without any matter fields.

In chapter 1 we discussed a no-go theorem, derived by Weinberg [40], that we will see at work in this chapter. The theorem states that the vacuum energy density cannot be cancelled dynamically, using a scalar field, without fine-tuning in any effective four-dimensional theory with constant fields at late times, that satisfies the following conditions:

1. General Covariance;
2. Conventional four-dimensional gravity is mediated by a *massless* graviton;
3. Theory contains a finite number of fields below the cutoff scale;
4. Theory contains no negative norm states.

Under these rather general assumptions the theorem states that the potential for the compensator field, which should adjust the vacuum energy to zero, has a runaway behavior. This means that there is no stationary point for the potential of the scalar field that should realize the adjustment, providing a severe difficulty for such models.

5.1 Scalar Field, Instabilities *in* dS-Space

The first attempts to cancel dynamically a ‘bare’ cosmological constant were made by referring to instabilities in the case of a scalar field in de Sitter space [194]. A massless minimally coupled scalar field ϕ has no de Sitter-invariant vacuum state, the two-point function in such a state does not exist, because of an IR-divergence [195]. As a consequence, the expectation value of ϕ^2 is time-dependent:

$$\langle \phi^2 \rangle = \frac{H^3}{4\pi^3} t. \tag{5.1}$$

However, this breaking of de Sitter invariance is not reflected by the energy-momentum tensor, since $T_{\mu\nu}$ only contains derivatives and hence is not sensitive to long-wavelength modes. This changes if one includes interactions. Consider for example a $\lambda\phi^4$. Then:

$$\langle T_{\mu\nu} \rangle \sim \lambda \langle \phi^2 \rangle^2 g_{\mu\nu} \propto t^2. \quad (5.2)$$

So in this case it is possible for $\langle T_{\mu\nu} \rangle$ to grow for some time, until higher order contributions become important. The infrared divergence results in a mass for the field which in turn stops the growth of $\langle T_{\mu\nu} \rangle$, see for example [194, 196], comparable to what happens in scalar field theory in flat spacetime, with a cubic self-interaction term.

Another example of an instability with scalar particles in De Sitter space was given by Myhrvold in [197] with an $\lambda\phi^4$ self-interaction term. In this case, spacetime curvature makes the particle decay into three particles, which again decay, in a runaway process. The interaction is crucial to break conformal invariance, without this breaking there are always stable de Sitter solutions [198].

One of the first illustrative, but unsuccessful attempts to use such instabilities to screen the cosmological constant, was made by Dolgov [199]. He used a rather simple classical model for back-reaction:

$$\mathcal{L} = \frac{1}{2} (\partial_\alpha \phi \partial^\alpha \phi - \xi R \phi^2), \quad \rightarrow \quad \nabla_\alpha \nabla^\alpha \phi + \xi R \phi = 0, \quad (5.3)$$

where R is the scalar curvature, assumed to be positive, and ξ is taken to be a *negative* constant. The energy-momentum tensor is:

$$\begin{aligned} T_{\mu\nu}(\phi) = & \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} g_{\mu\nu} \nabla_\kappa \phi \nabla^\kappa \phi - \xi \phi^2 (R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R) \\ & - \xi \nabla_\mu \nabla_\nu \phi^2 + \xi g_{\mu\nu} \nabla_\kappa \nabla^\kappa \phi^2, \end{aligned} \quad (5.4)$$

and the energy density in the scalar field ρ_ϕ :

$$\rho_\phi = \frac{1}{2} \dot{\phi}^2 + 3\xi \left(\frac{\dot{a}}{a} \right)^2 \phi^2 + 6\xi \left(\frac{\dot{a}}{a} \right) \dot{\phi} \phi. \quad (5.5)$$

In terms of the scale factor a , the equation of motion for ϕ reads:

$$\ddot{\phi} + 3 \frac{\dot{a}}{a} \dot{\phi} + 6\xi \left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a} \right)^2 \right) \phi = 0. \quad (5.6)$$

Together with the Friedmann equation, describing the evolution of the scale factor:

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{1}{3} (\Lambda_0 + 8\pi G \rho_\phi), \quad (5.7)$$

where $\Lambda_0 = 3H^2$ stands for the effective value of the cosmological constant during a de Sitter phase, this gives a pair of non-linear coupled equations, that provide growing solutions for ϕ . At early times:

$$\phi(t) = \phi_0 e^{\gamma t}, \quad \text{with} \quad \gamma = \frac{3}{2} H \left[\left(1 + \frac{16}{3} |\xi| \right)^{\frac{1}{2}} - 1 \right], \quad (5.8)$$

with t the comoving time in flat RW-coordinates. The energy density in the scalar field is negative and increases:

$$\begin{aligned}\rho_\phi &= -A\phi^2 \\ A &= \frac{1}{2} (6H^2|\xi| + 12\gamma H|\xi| - \gamma^2) > 0.\end{aligned}\tag{5.9}$$

The scalar curvature R and Hubble parameter H become are:

$$R = \frac{8\pi[4\lambda + (6\xi - 1)\dot{\phi}^2]}{M_P^2 + 8\pi\xi(6\xi - 1)\phi^2}, \quad H^2 = \frac{8\pi\lambda + \frac{1}{2}\dot{\phi}^2 + 6H\xi\phi\dot{\phi}}{3M_P^2 - 8\pi\xi\phi^2}\tag{5.10}$$

where λ has mass-dimension 4. The gravitational back-reaction on ϕ slows down the explosive rise of (5.8). At late times the field grows linearly in time $\phi \sim \sigma t$, where σ is some constant, and the scalefactor $a(t)$ grows as a power of time:

$$a(t) \sim t^\beta, \quad \text{with} \quad \beta = \frac{2|\xi| + 1}{4|\xi|},\tag{5.11}$$

and, consequently, $H \sim t^{-1}$ and $R \sim t^{-2}$.

Most importantly, the scalar field energy-momentum tensor at late times approaches the form of a cosmological constant term:

$$8\pi G\langle T_{\mu\nu} \rangle \sim -\Lambda_0 g_{\mu\nu} + \mathcal{O}(t^{-2}).\tag{5.12}$$

so the leading back-reaction term cancels the cosmological constant originally present. The kinetic energy of the growing ϕ -field acts to cancel the cosmological constant. The order (t^{-2}) part of the energy-density is:

$$\rho_\phi + \frac{\Lambda_0}{8\pi G} \sim \frac{3(2|\xi| + 1)2(1 - 2|\xi|)}{128\pi\xi^2 t^2},\tag{5.13}$$

which makes the effective cosmological constant nowadays extremely small.

Unfortunately, not only the cosmological constant term is driven to zero, since ϕ also couples to R , Newton's constant is also screened:

$$G_{eff} = \frac{G_0}{1 + 8\pi G_0|\xi|\phi^2} \sim \frac{1}{t^2},\tag{5.14}$$

where G_0 is the ‘‘bare’’ value of G at times where $\phi = 0$.

Other models of this kind were also studied by Dolgov, see [200, 201] but these proved to be unstable, leading quickly to a catastrophic cosmic singularity. All these models do evade Weinberg's no-go theorem, since the fields are not constant at late times.

Models such as described above, where a screening is based on a term $\xi R\phi^2$ have to be handled with extreme care, since such a $\xi R\phi^2$ term can be obtained, or transformed away by making a field transformation, and therefore is unphysical. The metric $g_{\mu\nu}$ can be rescaled with a ϕ -dependent scale factor, yielding a new metric. In general, under a field transformation $g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$ the (gravitational) action transforms as:

$$\mathcal{L} \rightarrow \mathcal{L} + (G^{\mu\nu} - 8\pi GT^{\mu\nu}) \delta g_{\mu\nu} + \mathcal{O}(\delta g^2)\tag{5.15}$$

If we now make the transformation $\delta g_{\mu\nu}$ dependent on a field ϕ :

$$\delta g_{\mu\nu} = \frac{1}{4}\lambda\phi^2 g_{\mu\nu} + \dots \quad (5.16)$$

we see that a term linear in the curvature scalar appears:

$$\mathcal{L} \rightarrow \mathcal{L} - \lambda (R\phi^2 + 2\pi GT\phi^2) + \mathcal{O}(\lambda^2) \quad (5.17)$$

Subsequent transformation on ϕ can bring its kinetic term in the canonical form. In the same way, a $\xi R\phi^2$ -term can be transformed away, at the cost of introducing higher dimensional self-interaction terms in the scalar field potential. A different way to see that such a term has no physical significance is to do perturbative quantum gravity, see [202]. At tree level, one can substitute the equations of motion, which in pure gravity gives $R = 0$, and if coupled to a scalar field $R = T_\mu^\mu$. At higher loop orders these become the quantum corrected equations of motion.

Moreover, a priori, it becomes unclear to which metric matter is coupled. In other words, this brings an ambiguity in the definition of the metric.

Models where a screening of a cosmological constant depend on such terms therefore can never lead to a solution of the cosmological constant problem.

5.1.1 ‘Cosmon’ Screening

This is a different version of trying to screen the cosmological constant by a scalar field χ , called ‘cosmon’ field [203, 204, 205]. It is assumed that all particle masses are determined by this scalar field. Moreover, the renormalization group equation for this field is assumed to be such, that at present it can play the role of quintessence. The value of the χ field increases with time.

One starts with the following effective action, after integrating out all standard model fields and fluctuations:

$$S = \int d^4x \sqrt{g} \left[-\frac{1}{2}\chi^2 R + \frac{1}{2} \left(\sigma \left(\frac{\chi}{m} \right) - 6 \right) \partial^\mu \chi \partial_\mu \chi + V(\chi) \right], \quad (5.18)$$

where σ is a parameter depending on χ and the potential is taken to be:

$$V(\chi) = m^2 \chi^2. \quad (5.19)$$

Note that in the action above, no cosmological constant is written. Since everything is integrated out, a cosmological term could have been introduced, but would necessarily be of near zero value, in order for this effective theory to make any sense.

The normal Einstein-Hilbert term can be retrieved by making a field transformation:

$$g_{\mu\nu} \rightarrow \left(\frac{M_P}{\chi} \right)^2 g_{\mu\nu}, \quad (5.20)$$

after which the effective action (5.18) transforms to:

$$S = \int d^4x \sqrt{g} \left[M_P^2 R + \frac{1}{2} \frac{M_P^2}{\chi^2} \left(\sigma \left(\frac{\chi}{m} \right) - 6 \right) \partial^\mu \chi \partial_\mu \chi + \frac{M_P^2}{\chi^2} V(\chi) \right]. \quad (5.21)$$

Adding a cosmological constant to this action, and performing the inverse of the above Weyl scaling, leads to:

$$S_\Lambda = M_P^2 \Lambda \quad \text{in (5.21)} \quad \rightarrow \quad S_\Lambda = \Lambda \chi^2 \quad \text{in (5.18)}. \quad (5.22)$$

This term is small for small χ^2 , but becomes larger and larger as the value of χ increases. A suggested additional transformation of the χ field:

$$\phi/M_P = \ln(\chi^4/V(\chi)) = \ln(\chi^2/m^2) \quad \rightarrow \quad \chi^2 = m^2 e^{\phi/m}, \quad (5.23)$$

does not help much. In this case the value of ϕ increases, and the cosmological constant term again becomes larger and larger.

In ref. [203, 204, 205], on the other hand, the transformation (5.23) is used on the χ -field which, together with the Weyl rescaling of the metric, transforming $V(\chi)$ to:

$$V(\chi, \phi) \rightarrow M_P^2 \chi^2 e^{-\phi/m}, \quad (5.24)$$

which suggests that $V(\chi)$ decreases for increasing ϕ . However, using the same transformation (5.23) again to remove the χ^2 term, we are left with

$$V = M_P^2 m^2, \quad (5.25)$$

a strange constant term, which again is of the form of a cosmological constant. Just putting $\chi^2 = M_P^2$, and therefore arguing that $V(\phi) \rightarrow 0$ for increasing ϕ does not seem to be correct.

Positing a renormalization group equation, according to which the kinetic term, parameterized by $\sigma(\chi)$ in (5.18), runs with energy E as:

$$\frac{\partial \sigma}{\partial \chi} = \beta_\sigma = E \sigma^2, \quad \sigma = \frac{1}{E \ln(\chi_c/\chi)}, \quad (5.26)$$

determines late time cosmology. It is suggested that postulating another renormalization group equation, for a parameter g specifying a potential of the form $V = g\chi^4$:

$$\chi \frac{dg}{d\chi} = -Ag, \quad A > 0, \quad (5.27)$$

will ensure asymptotically vanishing dark energy. However, not only is there no deeper reason for such a particular, renormalization group equation, with a minus-sign, this also still does not control the cosmological constant, as we saw before.

Note that the ‘renormalization group equation’ for the χ -field is important to make any statements about late time cosmology. The running of the effective cosmological constant at large values of χ , characterized as a $\lambda(\chi/m)\chi^4$ potential term is especially important and would generally lead to a late time cosmological constant that is too large. A ‘hidden fine-tuning’ needs to be reintroduced in the assumption that the effective potential in (5.18) has a flat direction.

Clearly, more is needed for this scenario to solve the cosmological constant problem.

5.1.2 Radiative Stability in Scalar Field Feedback Mechanism

Another approach deserves to be mentioned here, which will also be listed under back-reaction mechanisms. This concerns a model that does not solve the cosmological constant problem, but is intended to provide a way to protect a zero or small cosmological constant against radiative corrections, without using a symmetry, [206, 207]. This is achieved using a scalar field with a non-standard, curvature dependent kinetic term, such that in the limit where the scalar curvature goes to zero, the kinetic term vanishes.

$$\begin{aligned}
 S &= \int d^4x \sqrt{-g} \left(\frac{R}{2\kappa^2} + \alpha R^2 + L_{kin} - V(\phi) \right) \\
 L_{kin} &= \frac{\kappa^{-4} K^q}{2q f^{2q-1}},
 \end{aligned}
 \tag{5.28}$$

where q is a constant that has to be $q > 1/2$ for stability reasons, and f is a function of the scalar curvature R , postulated to vanish at $R = 0$ and that behaves near $R = 0$ as:

$$f(R) \sim (\kappa^4 R^2)^m, \tag{5.29}$$

with κ the Planck length. The parameter α is assumed to be $\alpha > 0$ to stabilize gravity at low energies, m is an integer that satisfies $2(m-1) > q(2q-1)$ and $K \equiv -\kappa^4 \partial^\mu \phi \partial_\mu \phi$.

The lowest value of $V(\phi)$, and thus the true value of the vacuum energy density, is assumed to be negative in this approach, not zero, but the peculiar dynamics makes the universe settle down to a near zero energy state. The scalar field stops rolling and its kinetic terms diverges.

The two main problems with this scenario are: 1) This specific kinetic term is chosen by hand, not motivated by a more fundamental theory, 2) all other fields settle to their ground state faster than the vacuum energy, making the universe empty, and reheating is necessary, to thermally populate the universe again.

Other models where some dynamical feedback mechanism is proposed based on a non-standard kinetic term can be found in [208, 209, 210, 205]. An interesting conjecture is made on the existence of a conformal fixed point, possibly related to dilatation symmetry [203]. However, all these models still need fine-tuning in masses and coupling constants, and it is unclear whether they are experimentally viable, see [207].

5.1.3 Dilaton

A natural scalar field candidate to screen the cosmological constant could be the dilaton, which appears in string theory and compactified supergravity theories. In the presence of a dilaton, all mass scales arise multiplied with an exponential:

$$V_0(\phi) \sim M^4 e^{4\lambda\phi}, \tag{5.30}$$

with ϕ the dilaton, and λ a coupling constant. The minimum of this potential is obtained for the value $\phi_0 = -\infty$, which leads to the so-called ‘dilaton runaway problem’: couplings depend typically on ϕ , and these tend to go to zero, or sometimes infinity, in this limit. Moreover, all mass scales have this similar scaling behavior, so particle

masses also vanish. Besides, the dilaton itself is nearly massless when it reaches the minimum of its potential, leading to long-range interactions that are severely constrained. Note that the difference with quintessence is that the hypothetical quintessence scalar field is assumed to couple only very weakly to the standard model fields, contrary to the dilaton. Both scenarios however predict varying ‘constants’ of nature, such as the fine structure constant α . The current limits on varying α are quite severe:

$$\frac{\Delta\alpha}{\alpha} \stackrel{z=2}{\approx} 10^{-5} \quad (5.31)$$

In summary, the dynamical cancellation of a cosmological constant term by back-reaction effects of scalar fields is hard to realize. Next, let’s focus on the other possible back-reaction mechanism, a purely gravitational one.

5.2 Instabilities of dS-Space

Gravitational waves propagating in some background spacetime affect the dynamics of this background. Imposing the transverse-tracefree gauge condition removes all gauge freedom, and the independent components of the perturbation become equivalent to a pair of massless scalar fields [211, 195].

This back-reaction can be described by an effective energy-momentum tensor $\tau_{\mu\nu}$.

5.2.1 Scalar-type Perturbations

In [212, 213], Brandenberger and coworkers study back-reaction effects of scalar gravitational perturbations. It is suggested that this could possibly solve the CC problem.

Gravitational perturbation theory is formulated in terms of an expansion in $g_{\mu\nu}$, see also section 7.5 of [214]:

$$g_{\mu\nu}(\alpha) = \sum_n \frac{1}{n!} \alpha^n g_{\mu\nu}^{(n)}, \quad (5.32)$$

with $g_{\mu\nu}^{(0)} \equiv g_{\mu\nu}(0)$ the background metric. Perturbation equations for $g_{\mu\nu}^{(n)}$ for the vacuum Einstein equation, temporarily ignoring matter:

$$G_{\mu\nu} = 0, \quad (5.33)$$

are obtained by differentiating the Einstein tensor $G_{\mu\nu}(\alpha)$, constructed from $g_{\mu\nu}(\alpha)$, n times with respect to α at $\alpha = 0$. The zeroth and first order Einstein equations for $g_{\mu\nu}^{(0,1)}$ cancel (this is the so-called ‘background field method’):

$$G_{\mu\nu}[g^{(0)}] = 0 \quad \text{and} \quad G_{\mu\nu}^{(1)}[g^{(1)}] = 0, \quad (5.34)$$

where $G_{\mu\nu}^{(1)}$ is the linearized Einstein tensor constructed from the background metric $g_{\mu\nu}^{(1)}$. The second order equation shows the important effect:

$$G_{\mu\nu}^{(1)}[g^{(2)}] = -G_{\mu\nu}^{(2)}[g^{(1)}], \quad (5.35)$$

where $G_{\mu\nu}^{(2)}[g^{(1)}]$ denotes the second-order Einstein tensor, constructed from $g_{\mu\nu}^{(1)}$.

This implies that $G_{\mu\nu}^{(2)}[g^{(1)}]$ plays the role of an ‘effective energy-momentum tensor’, associated with the perturbation $g_{\mu\nu}^{(1)}$, acting as a source for the second order metric perturbation $g_{\mu\nu}^{(2)}$.

When gravity is coupled to matter, $T_{\mu\nu} \neq 0$, but the same steps are taken. The zeroth and first order terms are assumed to satisfy the equations of motion. Next, the spatial average is taken of the remaining terms (a ‘coarse-grain viewpoint’) and the resulting equations are regarded as correction terms for a new homogeneous metric $g_{\mu\nu}^{(0,br)}$, where the superscript $(0, br)$ denotes the order in perturbation theory and the fact that back-reaction is taken into account:

$$G_{\mu\nu} \left(g_{\alpha\beta}^{(0,br)} \right) = -8\pi G \left[T_{\mu\nu}^{(0)} + \tau_{\mu\nu} \right] \quad (5.36)$$

and $\tau_{\mu\nu}$ contains terms resulting from averaging of the second order metric and matter perturbations:

$$\tau_{\mu\nu} = \langle T_{\mu\nu}^{(2)} - \frac{1}{8\pi G} G_{\mu\nu}^{(2)} \rangle. \quad (5.37)$$

In other words, the first-order perturbations are regarded as contributing an extra energy-momentum tensor to the zeroth-order equations of motion; the effective energy-momentum tensor of the first-order equations renormalizes the zeroth-order energy-momentum tensor.

Now work in longitudinal gauge and take for simplicity the matter to be described by a single scalar field φ with potential V . Then there is only one independent metric perturbation variable denoted $\phi(x, t)$. The perturbed metric is:

$$ds^2 = -(1 + 2\phi)dt^2 + a(t)^2(1 - 2\phi)\delta_{ij}dx^i dx^j. \quad (5.38)$$

and $g^{(0,br)}$ is a Robertson-Walker metric. Denoting by $\delta\varphi$ the matter perturbation:

$$\varphi(x, t) = \varphi_0(t) + \delta\varphi(x, t), \quad (5.39)$$

the τ_{00} and τ_{ij} elements of the effective energy momentum tensor are:

$$\begin{aligned} \tau_{00} &= \frac{1}{8\pi G} \left[12H\langle\phi\dot{\phi}\rangle - 3\langle(\dot{\phi})^2\rangle + 9a^{-2}\langle(\nabla\phi)^2\rangle \right] \\ &+ \langle(\delta\dot{\phi})^2\rangle + a^{-2}\langle(\nabla\delta\varphi)^2\rangle \\ &+ \frac{1}{2}V''(\varphi_0)\langle\delta\varphi^2\rangle + 2V'(\varphi_0)\langle\phi\delta\varphi\rangle, \end{aligned} \quad (5.40)$$

where we have deliberately made a split, to capture in the first line only contributions from the metric perturbations $\delta\phi$, and in the lower lines contributions from the matter perturbations $\delta\varphi$ (note however, that there are also cross-terms). Similarly, we write

$$\begin{aligned} \tau_{ij} &= a^2\delta_{ij} \left(\frac{1}{8\pi G} \left[(24H^2 + 16\dot{H})\langle\phi^2\rangle + 24H\langle\dot{\phi}\phi\rangle \right. \right. \\ &+ \left. \left. \langle(\dot{\phi})^2\rangle + 4\langle\phi\ddot{\phi}\rangle - \frac{4}{3}a^{-2}\langle(\nabla\phi)^2\rangle \right] \right) \\ &+ 4\dot{\varphi}_0^2\langle\phi^2\rangle + \langle(\delta\dot{\varphi})^2\rangle - a^{-2}\langle(\nabla\delta\varphi)^2\rangle - 4\dot{\varphi}_0\langle\delta\dot{\varphi}\phi\rangle \\ &- \frac{1}{2}V''(\varphi_0)\langle\delta\varphi^2\rangle + 2V'(\varphi_0)\langle\phi\delta\varphi\rangle \end{aligned} \quad (5.41)$$

with $H \equiv \dot{a}/a$ the Hubble parameter, $V' = \partial V/\partial\varphi$ and $\langle \rangle$ denotes spatial averaging. In the long-wavelength limit, considering modes with wavelengths longer than the Hubble radius, and ignoring terms $\propto \dot{\phi}^2$ on the basis that such terms are only important during times when the equation of state changes [215], this gives:

$$\tau_{00} = \frac{1}{2}V''(\varphi_0)\langle\delta\varphi^2\rangle + 2V'(\varphi_0)\langle\phi\delta\varphi\rangle \quad (5.42)$$

and

$$\begin{aligned} \tau_{ij} &= a^2\delta_{ij}\left(\frac{1}{8\pi G}\left[(24H^2 + 16\dot{H})\langle\phi^2\rangle\right] + 4\dot{\varphi}_0^2\langle\phi^2\rangle\right) \\ &- \frac{1}{2}V''(\varphi_0)\langle\delta\varphi^2\rangle + 2V'(\varphi_0)\langle\phi\delta\varphi\rangle. \end{aligned} \quad (5.43)$$

In case of slow-roll inflation, with φ the inflaton, these modes contribute:

$$\tau_0^0 \approx \left(2\frac{V''V^2}{V'^2} - 4V\right)\langle\phi^2\rangle \approx \frac{1}{3}\tau_i^i, \quad (5.44)$$

and:

$$p \equiv -\frac{1}{3}\tau_i^i \approx -\tau_0^0 \quad (5.45)$$

showing the main result, that the equation of state of the dominant infrared contribution to the effective energy-momentum tensor $\tau_{\mu\nu}$ which describes back-reaction, takes the form of a negative CC:

$$p_{br} = -\rho_{br}, \quad \rho_{br} < 0. \quad (5.46)$$

This leads to the speculation that gravitational back-reaction may lead to a dynamical cancellation mechanism for a bare CC, since $\tau_0^0 \propto \langle\phi^2\rangle$, which is proportional to IR phase space and this diverges in a De Sitter universe. Long wavelength modes are those with wavelength longer than H , and as more and more modes cross the horizon, $\langle\phi^2\rangle$ grows. To end inflation this way, however, takes an enormous number of e-folds, see [215] for a recent discussion.

This approach is strongly debated in the literature as it is not obvious whether one can consistently derive the equations of motion in this way, see for example [216, 217, 218, 219, 220]. We believe that the obtained result that the back-reaction of long wavelength gravitational perturbations screens the cosmological constant, is incorrect.

A direct reason to be sceptical is purely intuitive. ‘Long wavelength’ in this framework refers to energy-momentum modes that at every instant are much larger than the Hubble radius of interest. The Hubble radius defines the observable portion of the universe. How can one understand what happens locally? Take a box of a cubic meter¹. In this box one could measure the cosmological constant and one would find its near zero value. Why? How can modes with wavelengths larger than the Hubble radius, have any effect at all on what we measure in this box?

Note that we can always choose our coordinates such that locally at a given point P , $g'_{\mu\nu}(x'_P) = \eta_{\mu\nu}$ and $\partial g'_{\mu\nu}/\partial x'_\alpha = 0$ evaluated at $x = x_P$, simply constructing a

¹This is a very useful example often employed by Gerard 't Hooft to try to understand the fundamental mechanism of an approach.

local inertial frame at the point P . The second and higher order derivatives of the metric can of course not be made to vanish and measure the curvature. The long wavelength perturbations are small enough that we do not notice any deviation from homogeneity and isotropy, not even at cosmic distances, but are argued to be large enough to alter the dynamics of our universe, and determine the small value of the cosmological constant that we in principle could measure at distances of, say, a meter². This sounds contradictory.

One of the delicate issues in the above derivation is gauge invariance. As pointed out in [220], the spatially averaged metric is not a local physical observable: averaging over a fixed time slice, the averaged value of the expansion will not be the same as the expansion rate at the averaged value of time, because of the non-linear nature of the expansion with time. In other words, locally this ‘achieved renormalization’, i.e. the effect of the perturbations, is identical to a coordinate transformation of the background equations and not a physical effect. A similar conclusion was obtained in [221].

Brandenberger and co-workers have subsequently tried to improve their analysis by identifying a local physical variable which describes the expansion rate [222, 223]. This amounts to adding another scalar field that acts as an independent physical clock. Within this procedure they argue that back-reaction effects are still significant in renormalizing the cosmological constant.

However, in [224], arguments are given, where the above ‘derivation’ (5.36,5.37) of the effective energy-momentum tensor goes wrong. We will briefly review these here.

The central equation is (5.35). The point now is twofold: one is that $G_{\mu\nu}^{(2)}[g^{(1)}]$ is a highly gauge-dependent quantity, which cannot straightforwardly be inserted as a new source of energy-momentum, and two, if $G_{\mu\nu}^{(2)}[g^{(1)}]$ is small, its effects can be calculated from (5.35), but if it is large enough to apparently alter the dynamics of the universe, the third and higher contributions to $g_{\mu\nu}(\alpha)$ will also be large and calculating this back-reaction to second order perturbation theory does not give reliable results.

This mechanism to drive the cosmological constant to zero, therefore does not work, but back-reaction effects continue to be a hot topic in cosmology. Back-reaction of inhomogeneities has been invoked to explain an accelerating universe without a cosmological constant, for example by Kolb et al. [225, 226, 227]. In a decelerating, matter dominated universe, the back-reaction induced effective cosmological constant is positive. There is no consensus on this issue, although many critical studies have been undertaken to show that these back-reaction effects do not lead to an accelerated expansion. Simply put, it is just not possible to obtain an accelerating cosmology from a decelerating universe, just by means of back-reaction [228, 229, 230, 231]. For a large list of references and a critical examination, see [232]. Amusingly, if these back-reaction effects would lead to observable physical effects, they would include a renormalization that might be in conflict with CMB measurements [232, 228].

A more specific criticism is that the super-Hubble perturbations are assumed to satisfy the second order equation:

$$G_{ab}^{(1)}[g^{(0)}] = \langle -G_{ab}^{(2)}[g^{(1)}] \rangle, \quad (5.47)$$

while the first order perturbations are again assumed to satisfy the equations of motion.

²The scale of a meter is used just to make the argument more intuitive. The only constraint of course, is that it should be larger than roughly a millimeter.

The brackets indicate a spacetime averaging, such that the linear terms in the second order perturbation are eliminated, see [233] for details. In (5.47) the term on the right hand side may be very large, yet it can only be used when the wavelength of the perturbation is much smaller than the curvature of the background spacetime [233, 224].

In the next chapter we will discuss another back-reaction model in a purely quantum gravitational setting. At first sight, this appears to be more promising.

5.3 Running Λ from Renormalization Group

As discussed in chapter one, the cosmological constant in a field theory is expected to run with renormalization scale μ as any other dimensionful parameter. These are rather straightforward calculations and can be found in textbooks on quantum field theory. We follow the derivation as given in [234]

Consider the following Lagrangean:

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}m_0^2\phi^2 - \frac{g_0}{4!}\phi^4 + \lambda_0 + \rho_0\phi \quad (5.48)$$

where λ_0 is just a constant, that becomes a vacuum term in the energy-momentum tensor with dimension GeV^4 . All parameters m_0 , λ_0 and the source function ρ_0 are bare constants, that will be renormalized by quantum corrections. Wave-function renormalization appears through a scale change of this ρ_0 . We do not need that here, so we will set $\rho_0 = 0$ from now on. For the time being, we also set $g_0 = 0$, considering just the free scalar field. The first step is to derive the running of the vacuum energy density with masses m .

The vacuum transition amplitude, from the remote past to the distant future in Euclidean space can then be written as the functional integral:

$$\langle 0 + | 0 - \rangle = \int [d\phi] \exp \left(- \int d_E^n x \left[\frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}m_0^2\phi^2 - \lambda_0 \right] \right) \quad (5.49)$$

Using this expression, the variation with mass parameter m_0 is:

$$\frac{\partial}{\partial m_0^2} \langle 0 + | 0 - \rangle = -\frac{1}{2} \int d_E^n x \langle 0 + | \phi_0(x)^2 | 0 - \rangle = -\frac{1}{2} \int d_E^n x \Delta_E(0) \langle 0 + | 0 - \rangle, \quad (5.50)$$

with Δ_E is the Green's function, for two coincident points:

$$\Delta_E(0) = \int \frac{d_E^n k}{(2\pi)^n} \frac{1}{k^2 + m_0^2}, \quad (5.51)$$

and is written as a normalized expectation value in (5.50), when $\langle 0 + | 0 - \rangle$ differs from unity. In Feynman diagram language, this gives just a vacuum bubble, a loop diagram with no external legs.

This Green's function can be evaluated in the usual way, exponentiating the denominator:

$$\frac{1}{k^2 + m_0^2} = \int_0^\infty ds e^{-s(k^2 + m_0^2)} \quad (5.52)$$

and thus:

$$\Delta_E(0) = \frac{1}{(4\pi)^{n/2}} \int_0^\infty ds s^{-n/2} e^{-sm_0^2} = \frac{(m_0^2)^{\frac{n}{2}-1}}{(4\pi)^{n/2}} \Gamma\left(1 - \frac{n}{2}\right) \quad (5.53)$$

Rewriting (5.50) as:

$$\frac{\partial}{\partial m_0^2} \ln \langle 0 + | 0 - \rangle = -\frac{1}{2} \int d_E^n x \langle \Delta_E(0) \rangle, \quad (5.54)$$

this equation can trivially be integrated and gives:

$$\langle 0 + | 0 - \rangle = \exp\left(-\int d_E^n x \mathcal{E}\right) \quad (5.55)$$

with:

$$\mathcal{E} = \frac{m_0^n}{(4\pi)^{n/2} n} \Gamma\left(1 - \frac{n}{2}\right) - \lambda_0 \quad (5.56)$$

and by fine-tuning λ_0 , the vacuum energy density \mathcal{E} can be given any value. The special value $\mathcal{E} = 0$ would furthermore set $\langle 0 + | 0 - \rangle = 1$. Note that because of dimensional reasons, no other constant can appear in (5.56).

The expression (5.56) is just the lowest order vacuum energy density and of course represents the infinite sum of zero-point energies, as can be seen by performing the p^4 contour integration in the Fourier representation for $\Delta_E(0)$, eqn. (5.51) and integrating with respect to m_0^2 . This gives:

$$\mathcal{E} = \int \frac{d^{n-1}k}{(2\pi)^{n-1}} \frac{1}{2} \sqrt{k^2 + m_0^2} - \lambda_0, \quad (5.57)$$

a familiar expression, which we evaluated in chapter one.

Here we can use (5.56) instead. In four dimensions the leading term is again a quartic divergence, which can be traced back to the poles of the Gamma-function. It has poles at $n = 2$ and $n = 4$, reflecting the logarithmic divergence in two dimensions and quadratic divergence in four dimensions of the momentum integral (5.51). This pole can be made explicit using the recursion relation:

$$\Gamma(z+2) = z(z+1)\Gamma(z), \quad \text{and} \quad \Gamma(1) = 1. \quad (5.58)$$

The pole in four dimensions becomes:

$$\Gamma\left(1 - \frac{n}{2}\right) \simeq \frac{2}{n-4}. \quad (5.59)$$

Where we used the so-called minimal subtraction scheme (MS) in which only the pole term is removed from the Gamma-function. The residue is evaluated at $n = 4$ and μ^{n-4} varies as to maintain the correct dimensionality. The following renormalization is obtained for the vacuum term:

$$\lambda_0 = \mu^{n-4} \left(\frac{1}{2} \frac{m_0^4}{(4\pi)^2} \frac{1}{n-4} + \lambda \right) \quad (5.60)$$

This makes the four-dimensional vacuum energy finite at one loop order.

The counter term in (5.60) has to have integer powers of m_0^2 , since it corresponds to a subtraction of divergent contributions coming from high momentum, or large k^2 of integrals behaving like m_0^2/k^2 . The renormalization group equation then reads:

$$\mu \frac{d\lambda}{d\mu} = (4-n)\lambda - \frac{1}{2} \frac{m^4}{(4\pi)^2}, \quad (5.61)$$

written in terms of renormalized mass, which in the free theory, that we used here, is the same as the bare mass. In general of course they are not, but renormalization group equations, as eqn. (5.61), should always be written in terms of renormalized quantities. Note that the dominating effect in this equation is simply the canonical term, required by the dimension of λ . The higher orders contribute even less.

Substituting (5.60) in (5.56), taking the limit $n \rightarrow 4$ and using that:

$$\Gamma(2 - \frac{d}{2}) = \Gamma(\epsilon/2) = \frac{2}{\epsilon} - \gamma + \mathcal{O}(\epsilon), \quad (5.62)$$

with $\gamma \simeq 0.5772$ the Euler-Mascheroni constant, the renormalized vacuum energy density becomes:

$$\mathcal{E} = \frac{1}{4} \frac{m^4}{(4\pi)^2} \left[\ln \left(\frac{m^2}{4\pi\mu^2} \right) + \gamma - \frac{3}{2} \right] - \lambda. \quad (5.63)$$

With Interactions

Retaining the interaction term $\frac{g}{4!}\phi^4$, there is an additional renormalization of the vacuum energy. To first order in g_0 , the correction to the vacuum amplitude is:

$$\begin{aligned} \langle 0 + |0-\rangle^{(1)} &= -\frac{g_0}{4!} \int d_E^n x \langle 0 + |\phi_0^4(x)|0-\rangle^{(0)} \\ &= -3\frac{g_0}{4!} \int d_E^n x \Delta_E^2(0) \langle 0 + |0-\rangle^{(0)}. \end{aligned} \quad (5.64)$$

In Feynman diagram language, this corresponds to the disconnected double vacuum bubble, with a 4-point vertex.

The higher order contributions generate the following structure:

$$\langle 0 + |0-\rangle \simeq \langle 0 + |0-\rangle^{(0)} \exp \left(-3\frac{g_0}{4!} \int d_E^n x \Delta_E^2(0) \right), \quad (5.65)$$

plus other corrections. The expression for \mathcal{E} based on (5.65), is modified to:

$$\mathcal{E} = \frac{m_0^n}{(4\pi)^{n/2}} \frac{1}{n} \Gamma \left(1 - \frac{n}{2} \right) + \frac{g_0}{8} \left[\frac{m_0^{n-2}}{(4\pi)^{n/2}} \Gamma \left(1 - \frac{n}{2} \right) \right]^2 - \lambda_0. \quad (5.66)$$

Besides, now we also have to include mass renormalization, since due to the self-interaction term for the scalar field, the bare and renormalized mass are no longer the same. To order g the mass renormalization is given by:

$$m_0^n = m^n \left(1 - \frac{n}{2} \frac{g}{(4\pi)^2} \frac{1}{n-4} \right), \quad (5.67)$$

as g is dimensionless.

Substituting this expression, we obtain:

$$\begin{aligned} \mathcal{E} &= \frac{m^n}{(4\pi)^{n/2}} \frac{1}{n} \Gamma\left(1 - \frac{n}{2}\right) - \frac{1}{2} \mu^{n-4} \frac{m^4}{(4\pi)^2} \frac{1}{n-4} \\ &+ \frac{1}{2} \mu^{n-4} \frac{gm^4}{(4\pi)^4} \left[\left(\frac{m^2}{4\pi\mu^2} \right)^{\frac{n}{2}-2} \frac{1}{2} \Gamma\left(1 - \frac{n}{2}\right) - \frac{1}{n-4} \right]^2 \\ &+ \frac{1}{2} \mu^{n-4} \frac{m^4}{(4\pi)^2} \frac{1}{n-4} \left[1 - \frac{g}{(4\pi)^2} \frac{1}{n-4} \right] - \lambda_0. \end{aligned} \quad (5.68)$$

The first two lines are finite for $n \rightarrow 4$, so vacuum energy can again be rendered finite by renormalizing the cosmological constant:

$$\lambda_0 = \mu^{n-4} \left(\frac{1}{2} \frac{m^4}{(4\pi)^2} \frac{1}{n-4} \left[1 - \frac{g}{(4\pi)^2} \frac{1}{n-4} \right] + \lambda \right). \quad (5.69)$$

Note that although there is a double pole at $n = 4$, the residue of this pole does not depend on n , except for the overall factor μ^{n-4} . This ensures that the counterterm only depends on integer powers of the mass m .

The modified renormalization group equation for λ becomes:

$$\begin{aligned} \mu \frac{d\lambda}{d\mu} &= (4-n)\lambda - \frac{1}{2} \frac{m_0^4}{(4\pi)^2} \left[1 + 2 \frac{g}{(4\pi)^2} \frac{1}{n-4} \right] \\ &= (4-n)\lambda - \frac{1}{2} \frac{m^4}{(4\pi)^2}, \end{aligned} \quad (5.70)$$

showing that in fact, there is no correction to the renormalization group equation for λ to first order in g .

The generalization of this result to arbitrary order in g reads:

$$\mu \frac{d\lambda}{d\mu} = -(n-4)\lambda + m^4 \beta_\lambda(g), \quad (5.71)$$

with $\beta_\lambda(g)$ independent of n .

Can Renormalization Group Running Nevertheless be of Use to the Cosmological Constant Problem?

In [235, 236, 179, 237, 238, 239] an approach is studied, viewing Λ as a parameter subject to renormalization group running. The cosmological constant becomes a scaling parameter $\Lambda(\mu)$, where μ is often identified with the Hubble parameter at the corresponding epoch, in order to make the running of Λ smooth enough to agree with all existing data, [240]. The question is whether this running can screen the cosmological constant.

Renormalization group equations typically give logarithmic corrections, which makes it hard to see how this can ever account for the suppression of a factor of 10^{120} needed

for the cosmological constant. In the above Refs., On dimensional grounds the renormalization group equations are written:

$$\begin{aligned} \frac{d\lambda}{d\ln\mu} &= \sum_n A_n \mu^{2n}, & \rightarrow & \quad \lambda(\mu) = \sum_n C_n \mu^{2n} \\ \frac{d}{d\ln\mu} \left(\frac{1}{G} \right) &= \sum_n B_n \mu^{2n}, & \rightarrow & \quad \frac{1}{G(\mu)} = \sum_n D_n \mu^{2n} \end{aligned} \quad (5.72)$$

with A_n, B_n, C_n and D_n coefficients, that depend on the particle content of the standard model. The right-hand-sides of the equations on the left define the β -functions for λ and G^{-1} , and all terms in this β -function for λ are of the form $\mu^{2n} m_i^{4-2n}$, with m_i the masses of different particles of the standard model. For the cosmological constant, the common A_0 -term, proportional to m_i^4 , has to be ignored in order to describe a phenomenologically successful cosmology [241, 240]. The dominant term is then the second one, which makes the screening still very slow, but stronger than logarithmic. We write:

$$\frac{d\lambda}{d\ln\mu} = \frac{\sigma}{(4\pi)^2} \mu^2 m_i^2 + \dots, \quad (5.73)$$

where $\sigma = \pm 1$ depending on whether bosonic or fermionic fields dominate below the Planck mass and $\sigma = 0$ if $\mu < m_i$. Integrating this, we find:

$$\lambda(\mu) = C_0 + C_1 \mu^2; \quad C_1 = \frac{\sigma}{2(4\pi)^2} m_i^2. \quad (5.74)$$

We can compare m_i^2 with M_P^2 by introducing a parameter ν :

$$\nu = \frac{\sigma}{12\pi} \frac{m_i^2}{M_P^2} \quad (5.75)$$

where the pre-factor $1/(12\pi)$ is a convention used in refs. [235, 236, 179, 237, 238, 239]. Note that for a standard FRW model, where λ is spacetime independent, we have $\nu = 0$. The renormalization group parameter μ is identified with the Hubble parameter $H(t)$, such that the renormalization condition can be set at $\mu_0 = H_0$, i.e. $\lambda(\mu_0) = \lambda_0$, with λ_0 the present day observed value of the cosmological constant. This identification is inspired by the numerical agreement that $H_0^2 M_P^2 \approx 10^{-47} \text{ GeV}^4 = \lambda_0$, since $H_0 \approx 10^{-33} \text{ eV}$. With this renormalization condition, the integration constant C_0 becomes:

$$C_0 = \lambda_0 - \frac{3\nu}{8\pi} M_P^2 \mu_0^2, \quad C_1 = \frac{3\nu}{8\pi} M_P^2 \quad (5.76)$$

For all standard model particles the mass $m_i > \mu$, so they decouple, and the beta-function in (5.73) effectively becomes:

$$\frac{d\lambda}{d\ln\mu} = \frac{\sigma}{(4\pi)^2} \mu^4 \quad (5.77)$$

resulting in coefficients:

$$C_0 = \lambda_0 - C_1^{SM} \mu_0^4. \quad (5.78)$$

Since $\mu_0 = 10^{-33} \text{ eV}$, this implies that there is basically no running at all, for masses well below the Planck scale.

However, since G and Λ now effectively become time-dependent, the energy conservation equation in a FRW universe (see (2.8)):

$$\dot{\rho} + 3H(\rho + p) = 0, \quad (5.79)$$

now becomes:

$$(\rho + \Lambda)\dot{G} + G\dot{\Lambda} = 0. \quad (5.80)$$

A cosmological model can be set up, since we have, in addition to (5.74, 5.80):

$$\rho + \Lambda = \frac{3H^2}{\pi G}, \quad (5.81)$$

which can be solved for $G(H, \nu)$:

$$G(H, \nu) = \frac{G_0}{1 + \nu \ln(H^2/H_0^2)}, \quad (5.82)$$

and $G_0 = G(H_0) = 1/M_P^2$.

The crucial term in the running of G is the canonical one:

$$\mu \frac{d}{d\mu} \frac{1}{G} = \sum_i a_i M_i^2 + \dots \quad (5.83)$$

had this term not been taken into account, the running of G and Λ with H would become [242]:

$$\begin{aligned} G(H) &= \frac{G_0}{1 + \alpha G_0 (H^2 - H_0^2)}, \\ \lambda(H) &= \lambda_0 + \frac{3\alpha}{16\pi} (H^4 - H_0^4) \end{aligned} \quad (5.84)$$

and the running would be so slow, that it results in essentially no phenomenology at all.

Keeping this term, the running is still very slow, but it could possibly be measured as a quintessence or phantom dark energy and be consistent with all data, as long as $0 \leq |\nu| \ll 1$ [242]. As a solution to the cosmological constant problem, it obviously cannot help.

Besides, based on RG-group analysis, it is argued in [243, 239] that there may be a UV fixed point at which gravity becomes asymptotically free. If there would be an IR fixed point at which $\Lambda_{eff} = 0$ this could shed some new light on the cosmological constant problem. This scaling also effects G , making it larger at larger distances.

5.3.1 Triviality as in $\lambda\phi^4$ Theory

The Einstein Hilbert action with a cosmological constant can be rewritten as [244]:

$$S = -\frac{3}{4\pi} \int d^4x \sqrt{-\hat{g}} \left(\frac{1}{12} R(\hat{g}) \phi^2 + \frac{1}{2} \hat{g}^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - \frac{\lambda}{4!} \phi^4 \right) \quad (5.85)$$

after rescaling the metric tensor as:

$$g_{\mu\nu} = \varphi^2 \hat{g}_{\mu\nu}, \quad ds^2 = \varphi^2 \hat{d}s^2 \quad (5.86)$$

and defining:

$$\phi = \frac{\varphi}{\sqrt{G}}, \quad \Lambda = \frac{\lambda}{4G}. \quad (5.87)$$

Now it is suggested that the same arguments first given by Wilson [245], that are valid in ordinary $\lambda\phi^4$ -theory, might also hold here and that this term is suppressed quantum mechanically.

It is noted that perturbative running as in normal $\lambda\phi^4$ -theory is by far not sufficient, one has:

$$\mu \frac{d\Lambda}{d\mu} = \beta_1 \Lambda^2, \quad \Lambda = \frac{\Lambda_0}{1 - \Lambda \beta_1 \log \mu}, \quad (5.88)$$

the purely gravitational renormalization group equation for a cosmological constant. However, the hope is that there might be some non-perturbative suppression. Similar ideas have been contemplated by Polyakov, [246].

5.4 Screening as a Consequence of the Trace Anomaly

As we discussed, Weinberg's no-go theorem is widely applicable to screening mechanism. However, it was noted in e.g. [247], that the symmetry (1.48) may be broken by conformal anomalies, after which the Lagrangian obtains an additional term proportional to $\sqrt{g}\phi\Theta_\mu^\mu$, where Θ_μ^μ is the effect of the conformal anomaly.

However, as already noted by Weinberg [40], this does not provide a loophole to get around the no-go theorem. The reason is that, although the field equation for ϕ now looks like:

$$\frac{\partial \mathcal{L}}{\partial \phi} = \sqrt{-g} (T_\mu^\mu + \Theta_\mu^\mu), \quad (5.89)$$

which may suggest an equilibrium value for ϕ with zero trace, this is not sufficient for a flat space solution. The Einstein equation for a constant metric now becomes:

$$0 = \frac{\partial \mathcal{L}_{eff}}{\partial g_{\mu\nu}} \propto e^{2\phi} \mathcal{L}_0 + \phi \Theta_\mu^\mu, \quad (5.90)$$

and the extra factor of ϕ shows that these two conditions are not the same. The reason is that the term Θ_μ^μ does not simply end up in T_μ^μ .

More sophisticated are proposals in which it is argued that the quantum trace anomaly of massless conformal fields in 4 dimensions leads to a screening of the cosmological constant [248, 249, 250, 251, 252]. The idea is similar in spirit to the one in the previous section, (6). One tries to find a renormalization group screening of the cosmological constant in the IR, but instead of taking full quantum gravity effects, only quantum effects of the conformal factor are considered. See also [253] for a related earlier study.

Recall that if an action is invariant under conformal transformations:

$$g_{\mu\nu}(x) \rightarrow \Omega^2(x) g_{\mu\nu}(x) = \bar{g}_{\mu\nu}(x), \quad (5.91)$$

one directly finds, after varying with respect to $g_{\mu\nu}$ for the trace of the energy momentum tensor:

$$T_{\mu}^{\mu}[g_{\mu\nu}(x)] = -\frac{\Omega(x)}{\sqrt{-g}} \frac{\delta S[\bar{g}_{\mu\nu}]}{\delta \Omega(x)} \Big|_{\Omega=1} = 0. \quad (5.92)$$

This shows that if the classical action is invariant under conformal transformations, then the classical energy momentum tensor is traceless.

Masses and dimensionful couplings explicitly violate conformal invariance and even if one starts out with a classical theory that is invariant under conformal transformations, quantum corrections generally introduce mass scales. This is closely related with the scaling behavior of the action, and hence with the renormalization group. The vacuum fluctuations of massless fields, for instance, generate such a conformal anomaly.

Such a non-vanishing trace of the energy-momentum tensor couples to a new spin-0 degree of freedom, reflected in the conformal sector, or trace of the metric. It is argued, that fluctuations of this conformal sector of the metric grow logarithmically at distance scales of the order of the horizon. Moreover, their effect on the renormalization group flow is argued to lead to an IR stable, conformally invariant fixed point of gravity, where scale invariance is restored. The anomalous trace leads to an effective action that is non-local in the full metric for gravitational interactions, although the trace anomaly itself is given by a sum of local terms that are fourth order in curvature invariants.

The authors conclude that the effective cosmological constant in units of Planck mass decreases at large distances and that $G_N \Lambda \rightarrow 0$ at the IR fixed point in the infinite volume limit.

However, the cosmological constant problem manifests itself already at much smaller distances and moreover, it is unclear whether this scenario is compatible with standard cosmological observations. Moreover, like the other approaches in this chapter, it relies heavily on quantum effects having a large impact at enormous distance scales. As argued in the previous section, it is debatable whether these effects can be sufficiently significant.

5.5 Summary

Finding a viable mechanism that screens the original possibly large cosmological constant to its small value today, is a very difficult task. Weinberg's no-go theorem puts severe limits on this approach. Back-reaction effects moreover, are generally either very weak, or lead to other troublesome features like a screened Newton's constant. These models are often studied on an inflationary background to power the gravitational back-reaction effects, but still typically require an enormous number of e-folds, to see any effect at all.

Another drawback is that it is hard to understand the impact of these back-reaction effects on local physics.

The underlying idea however that the effective cosmological constant is small simply because the universe is old, is attractive and deserves full attention.

6 | Do Infrared Gravitons Screen the Cosmological Constant?

Knowing that the proposals of the previous chapter do not work, one may want to study a purely quantum gravitational backreaction mechanism. It indeed had been argued that during inflation this process is sufficient to screen the cosmological constant to zero. This could therefore be of considerable importance to the cosmological constant problem. Unfortunately however, we conclude that this screening has negligible effects on the cosmological constant.

6.1 Introduction

In a series of papers, Tsamis and Woodard [254, 255, 256, 257, 258, 259, 260], arrived at the remarkable result that the back-reaction of inflationary produced gravitons is sufficient to cancel a ‘bare’ cosmological constant from roughly 10^{16} GeV down to zero. In a Newtonian sense, it is argued that the interaction energy density between these gravitons will screen such a huge cosmological constant. Their actual calculations involve two loop quantum gravitational processes and are therefore very complicated.

In the remainder, we will be concerned with two questions. One is whether indeed this interaction energy generates negative vacuum energy and screens the cosmological constant, and the other, if so, whether this effect can become as large as claimed by Tsamis and Woodard in the above cited papers.

Our motivation for a close study of this proposal is not only the cosmological constant problem, but also to understand whether the standard paradigm that quantum gravitational effects cannot have a major influence on large distances, may be incorrect. Note also the previous results from Taylor and Veneziano [261], who argued that at large distances quantum gravitational effects act to slightly *increase* the value of a positive cosmological constant.

In general, the dominant infrared effects come from the lightest particles self-interacting with lowest canonical field dimension. Gravitons are massless for any value of Λ , but for $\Lambda = 0$ their lowest self-interaction term consists of two derivatives distributed between three graviton fields ($\sqrt{-g}R$ -term) and this is why conventional quantum gravity is indeed very weak in the infrared. However, when $\Lambda \neq 0$, the lowest dimensional self-interaction term is of dimension three, a three-point vertex with no derivatives (corresponding to the $\Lambda\sqrt{-g}$ -term). The IR behavior of the theory with cosmological constant could therefore perhaps be very different from that without. Tsamis and Woodard christen it Quantum Cosmological Gravity, or QCG for short [254].

They argue that on an inflationary background, the infrared divergences are enhanced. The spatial coordinates expand exponentially with increasing time, and so their Fourier

conjugates, the spatial momenta, are redshifted to zero. The IR effects originate from the low end of the momentum spectrum, so they are strengthened when this sector is more densely populated.

Since other particles are either massive, in which case they decouple from the infrared, or conformally invariant, and therefore do not feel the de Sitter redshift, gravitons must completely dominate the far IR.

Of course, quantum gravitational effects are very weak. The typical strength of quantum gravitational effects during inflation at scale M is:

$$G\Lambda = 8\pi \left(\frac{M}{M_P} \right)^4, \quad (6.1)$$

which for GUT-scale inflation becomes $G\Lambda = 10^{-11}$ and for electroweak-scale inflation $G\Lambda = 10^{-67}$. The hope is that these small numbers are overcome by a very large number of e-folds of inflation.

6.2 A Review of the Scenario

It has been shown by Grishchuk [262] that in an expanding spacetime gravitons are being produced, a process most efficient during inflation. When massless, virtual particles are produced out of the vacuum, those with wavelengths larger than H^{-1} will not recombine and annihilate, but are able to escape to infinity. This is similar to a black hole emitting Hawking radiation. For a derivation one could also check [65]. A more intuitive picture can be obtained using the energy-time uncertainty principle [263], which we will now discuss.

Consider a particle with mass m and co-moving wavevector \vec{k} in a spacetime with scalefactor $a(t)$:

$$E(\vec{k}, t) = \sqrt{m^2 + \|\vec{k}\|^2/a^2(t)}. \quad (6.2)$$

The Heisenberg uncertainty principle restricts how long a virtual pair of particles with $\pm\vec{k}$ can live. The lifetime, Δt , is given by the integral:

$$\int_t^{t+\Delta t} dt' E(\vec{k}, t') \sim 1. \quad (6.3)$$

The smaller the mass of a particle, the longer it survives, and for the fully massless case, in de Sitter spacetime with $a(t) = \exp(Ht)$:

$$\int_t^{t+\Delta t} dt' E(\vec{k}, t') \Big|_{m=0} = [1 - e^{-H\Delta t}] \frac{k}{Ha(t)} \sim 1. \quad (6.4)$$

Thus growth of $a(t)$ increases the time a ‘virtual’ particle of fixed m and \vec{k} can exist and, during inflation, particles with zero mass and wavelength $k \lesssim Ha(t)$ can exist forever.

Conformal invariance suppresses the number of particles that is being produced by a factor of $1/a(t)$ [263], so massless minimally coupled scalars and gravitons, which are not conformally invariant, will be most abundant. The energy-momentum carried by

these gravitons is proportional to H^4 . In TW's work inflation is assumed to start at an energy of roughly 10^{16} GeV, because of a large cosmological constant. So the energy density in infrared gravitons is quite small, compared to the vacuum energy density. More importantly, their energy-momentum tensor can never be interpreted as vacuum energy-density, since the latter is given by $Cg_{\mu\nu}$, with C some constant. Besides, the graviton energy-density is positive and could not screen a positive cosmological constant.

A priori, it could have been a nice idea to argue that a Bogolyubov transformation could bring one from one vacuum to another, with a different value for the cosmological constant. Similar to a black hole emitting Hawking radiation and thereby decreasing its mass, one could think that the inflationary produced gravitons would diminish the cosmological constant. This analogy however fails. The Hawking effect is a one-loop effect, and at one loop order, there can only be at best a small constant renormalization of Λ . The Hawking effect can be derived from a Bogolyubov transform of the vacuum state, whereas the inflationary produced gravitons cannot in such a way produce a decrease in the cosmological constant. The difference is exactly that a black hole is formed from an object of a certain mass in spacetime, whereas the cosmological constant is a property of spacetime itself.

The argument of Tsamis and Woodard however, is that the *interaction energy density* between different gravitons, takes on the form $T_{\mu\nu} = Cg_{\mu\nu}$, with C a positive constant in a $(-+++)$ -metric, thus acting as a negative vacuum energy density. Per graviton pair, this interaction energy density is negligible as well, and, moreover, the gravitons produced this way soon leave the Hubble volume. Pairs produced in different Hubble volumes are not in causal contact with each other, and hence have no gravitational interaction with each other. However, the argument is that at the end of inflation the past lightcone is growing larger and larger, so eventually all the inflationary produced gravitons do come into causal contact with each other, making the effect large enough to completely screen the huge bare cosmological constant to zero. Their calculations indicate that roughly 10^7 e-folds are necessary to build up a large enough effect, after which inflation stops rather suddenly over the course of a few (less than ten) e-folds.

Two additional questions arise in general in these back-reaction mechanisms. First, is it possible at all to generate a mechanism that makes the value of the cosmological constant time-dependent and therefore, observer-dependent? And secondly, how can such a mechanism account for a small counterterm $\delta\Lambda$ of order $1/\text{cm}^2$? We will return to both of these questions in section (6.5). But, as is well-known, and was discussed in section (1.5.1), in order to make the cosmological constant time-dependent, one has to introduce a new field. This is typically a scalar field to maintain Lorentz invariance. Let us stress here that Tsamis and Woodard do not introduce a scalar field. However, it appears that their mechanism could be modelled using a growing scalar field to act as the growing interaction energy density of the produced gravitons.

6.2.1 A Newtonian Picture

To get a better feeling for the suggested proposal, consider first the rough Newtonian derivation, as given by Woodard in [264]. We will present our comments in the next section.

The energy-density in inflationary produced infrared gravitons is:

$$\rho_{IR} \sim H^4, \quad (6.5)$$

with H the Hubble parameter during inflation: $H^2 = \Lambda/3$. Starting with initial radius H^{-1} , the physical radius of the universe, is exponentially growing with co-moving time t :

$$r(t) \sim H^{-1}e^{Ht}. \quad (6.6)$$

The total energy per volume, denoted by M is taken to be:

$$M(t) \sim r^3(t)\rho_{IR} \sim He^{3Ht}. \quad (6.7)$$

Note that this assumes a continuous production of gravitons to balance the growing volume. We will return to this relation in the next section, for now let's continue the discussion as given in [264].

If this mass would self-gravitate, the Newtonian interaction energy would be:

$$E_N = -\frac{GM^2(t)}{r(t)} \sim -GH^3e^{5Ht}, \quad (6.8)$$

assuming that the gravitons are on average a distance $r(t)$ apart. This gives an interaction energy density, $E_N/r^3 = -GH^6e^{2Ht}$. However, most of the inflationary produced gravitons will not be in causal contact with each other, as they soon leave their Hubble volume. If one assumes, as Tsamis and Woodard do, that their potentials $V = -GM/r$ remain, the rate at which they accumulate is estimated to be:

$$\frac{dV}{dt} \sim -GH^3e^{2Ht}. \quad (6.9)$$

The growth of the Newtonian interaction energy density during a short time interval is:

$$\begin{aligned} \rho(t) &\sim \rho_{IR}V(t) \sim -GH^6Ht \\ &= -\frac{\Lambda}{8\pi G}(G\Lambda)^2Ht, \end{aligned} \quad (6.10)$$

using $H^2 \sim \Lambda$. After a long period of inflation, the potential's accumulation rate is assumed to be very small, such that:

$$|\dot{\rho}(t)| \ll H|\rho(t)|, \quad (6.11)$$

and, using energy conservation:

$$\dot{\rho}(t) = -3H(\rho(t) + p(t)), \quad (6.12)$$

this implies that the induced interaction pressure $p(t)$ must be nearly opposite to the interaction energy density. This would imply, that the interaction energy approaches the form of negative vacuum energy and hence, according to Tsamis and Woodard, would screen the positive cosmological constant.

6.3 Evaluation

The back-reaction of the inflationary produced gravitons is a purely non-local effect. For a local observer, inside a Hubble volume, there are too few gravitons to have any observable effect on the cosmological constant. About one infrared pair emerges per Hubble time in each Hubble volume [265]. An important assumption made in the above section was that even after the gravitons have left their Hubble volume their potentials remain. It is the accumulation of these potentials that, in a Newtonian sense, causes the build-up of negative vacuum energy density, and hence the screening.

The result that has led to a large effect on the cosmological constant, appears to hinge on certain assumptions. In terms of the Newtonian argument - which was admitted to be somewhat intuitive - one could suspect that the graviton energy density in an expanding universe decreases by a factor of a^{-4} , like any other massless relativistic species. If this were the correct prescription, one would have to replace the total energy per volume, denoted M , as given in (6.7) by:

$$M(t) \sim r^3(t)\rho_{IR}e^{-4Ht} \sim He^{-Ht}, \quad (6.13)$$

The total energy density would decrease! The number of Hubble volumes grows like e^{3Ht} but the energy density per graviton in that case decreases like e^{-4Ht} and thus can never grow larger. The interaction energy density is proportional to M^2 and we have to correct (6.8), unless one assumes an ever increasing production rate of gravitons, to:

$$V = -\frac{GM^2(t)}{r(t)} \sim -GH^3e^{-3Ht}, \quad (6.14)$$

which obviously cannot have any effect in screening the cosmological constant.

Another issue is that in order to make the statement that the interaction energy density acts like negative vacuum energy, the standard equation expressing conservation of energy has been used. This equation is derived from the Einstein equations, more specifically, from demanding that the energy-momentum tensor is covariantly conserved: $\nabla_\mu T^{\mu\nu} = 0$. However, in this setup the energy density is not a local energy density and it is therefore not a priori clear that its energy momentum tensor is covariantly conserved, nor that it satisfies the *local* Einstein equations.

Furthermore, where the Newtonian argument concludes that after a long time the interaction energy density does not change much anymore, $\dot{\rho} = 0$, the question must be raised at which point this equation sets in. This would be the point where, using (6.12), the potential energy due to the gravitons satisfies the equation of state for vacuum energy density: $p = -\rho$.

This seems remarkable: if the interaction energy density generates an energy-momentum tensor $Cg_{\mu\nu}$ characteristic of vacuum energy, than one might argue that it would do so from the beginning. If the interaction energy density between a single graviton pair cannot be interpreted as vacuum energy density, than the interaction energy density of N graviton pairs also cannot be interpreted as vacuum energy density.

The real answers to these questions must, of course, come from the more explicit calculations, which will be scrutinized in the following sections.

Another issue is that gravitational potentials are of course not gauge invariant, which implies that at this point the physical reality of the interaction energy density is not

obvious. In the full quantum gravitational framework, this amounts to imposing the correct renormalization condition, and we will therefore pay special care to this issue in the following sections.

6.4 The Full Quantum Gravitational Calculation

To be more concrete, the actual computation to calculate the IR effects, involves the expectation value of the invariant element in the presence of a homogeneous and isotropic, initially free de Sitter vacuum:

$$\langle 0 | g_{\mu\nu}(t, \vec{x}) dx^\mu dx^\nu | 0 \rangle. \quad (6.15)$$

The easiest way to do this is first to calculate the amputated expectation value, and then add the external leg.

The production of gravitons is a one-loop effect, so their back-reaction at the metric starts at two-loop. The calculation therefore focusses on the two-loop 1-point function. The effect of the one-loop 1-point function is absorbed into a local counterterm plus a time-dependent redefinition of the coordinate system [257]. Note that this is the general strategy for removing the tadpole diagrams. They are removed by a substitution:

$$h_{\mu\nu} \rightarrow h_{\mu\nu} + a_{\mu\nu} \quad (6.16)$$

with $a_{\mu\nu}$ necessarily time dependent, see e.g. [28]¹. So a time dependent cosmological counterterm absorbs their effect. This is a very convenient bookkeeping device, since otherwise the graviton propagators would look very ugly. We will return to this later, since in their work, Tsamis and Woodard actually calculate the two-loop tadpoles, without subtracting them with a counterterm, see figure (6.1).

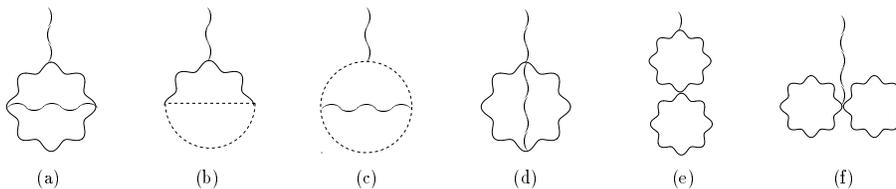


Figure 6.1: Two-loop contributions to the background geometry. Gravitons reside on wavy lines and ghosts on segmented lines, from [255].

The classical background in conformal coordinates is:

$$ds^2 = -dt^2 + e^{2Ht} d\vec{x} \cdot d\vec{x} = \Omega^2 (-du^2 + d\vec{x} \cdot d\vec{x}) \quad (6.17)$$

$$\Omega \equiv \frac{1}{Hu} = \exp(Ht) = R(t) \quad (6.18)$$

with $H^2 \equiv \frac{1}{3}\Lambda$ and $R(t)$ the scalefactor. For convenience, perturbation theory is formulated in terms of a ‘pseudo-graviton’ field $\psi_{\mu\nu}$:

$$g_{\mu\nu} \equiv \Omega^2 \tilde{g}_{\mu\nu} \equiv \Omega^2 (\eta_{\mu\nu} + \kappa \psi_{\mu\nu}) \quad (6.19)$$

¹Note that in flat spacetime, in for example the Higgs mechanism, this can be accomplished by a purely constant shift.

where $\kappa^2 \equiv 16\pi G$.

Because of homogeneity and isotropy of the dynamics and the initial state, the amputated 1-point function, can be written in terms of two functions of conformal time u :

$$D_{\mu\nu}^{\rho\sigma} \langle 0 | \kappa \psi_{\rho\sigma}(x) | 0 \rangle = a(u) \bar{\eta}_{\mu\nu} + c(u) \delta_\mu^0 \delta_\nu^0, \quad (6.20)$$

where $D_{\mu\nu}^{\rho\sigma}$ is the gauge fixed kinetic operator, and a bar on $\eta_{\mu\nu}$ indicates that temporal components of this tensor are deleted:

$$\eta_{\mu\nu} = \bar{\eta}_{\mu\nu} + \delta_\mu^{(0)} \delta_\nu^{(0)}. \quad (6.21)$$

The explicit form for $D_{\mu\nu}^{\rho\sigma}$ is found to be:

$$\begin{aligned} D_{\mu\nu}^{\rho\sigma} = & \left(\frac{1}{2} \bar{\delta}_\mu^{(\rho} \bar{\delta}_\nu^{\sigma)} - \frac{1}{4} \eta_{\mu\nu} \eta^{\rho\sigma} - \frac{1}{2} \delta_\mu^0 \delta_\nu^0 \delta_0^\rho \delta_0^\sigma \right) D_A \\ & + \delta_{(\mu}^0 \bar{\delta}_{\nu)}^{(\rho} \delta_0^{\sigma)} D_B + \delta_\mu^0 \delta_\nu^0 \delta_0^\rho \delta_0^\sigma D_C. \end{aligned} \quad (6.22)$$

The pseudo-graviton kinetic operator $D_{\mu\nu}^{\rho\sigma}$ splits in two parts, a term proportional to $D_A \equiv \Omega(\partial^2 + \frac{2}{u^2})\Omega$, which is the kinetic operator for a massless minimally coupled scalar, and a part proportional to $D_B = D_C \equiv \Omega\partial^2\Omega$, the kinetic operator for a conformally coupled scalar.

After attaching the external leg one obtains the full 1-point function, which has the same form, but with different components:

$$\langle 0 | \kappa \psi_{\mu\nu}(x) | 0 \rangle = A(u) \bar{\eta}_{\mu\nu} + C(u) \delta_\mu^0 \delta_\nu^0. \quad (6.23)$$

The functions $A(u)$ and $C(u)$ obey the following differential equations:

$$\begin{aligned} -\frac{1}{4} D_A [A(u) - C(u)] &= a(u) \\ D_C C(u) &= 3a(u) + c(u) \end{aligned} \quad (6.24)$$

The functions $a(u)$ and $A(u)$ on the one hand, and $c(u)$ and $C(u)$ on the other, are therefore related by retarded Green's functions $G_{A,C}^{ret}$ for the massless minimally coupled and conformally coupled scalars:

$$\begin{aligned} A(u) &= -4G_A^{ret}[a](u) + G_C^{ret}[3a + c](u), \\ C(u) &= G_C^{ret}[3a + c](u) \end{aligned} \quad (6.25)$$

In terms of the functions $A(u)$ and $C(u)$ the invariant element in comoving coordinates reads:

$$\hat{g}_{\mu\nu}(t, \vec{x}) dx^\mu dx^\nu = -\Omega^2 [1 - C(u)] du^2 + \Omega^2 [1 + A(u)] d\vec{x} \cdot d\vec{x}. \quad (6.26)$$

Comparing (6.17) with (6.26) Tsamis and Woodard make the following identification:

$$\begin{aligned} R(t) &= \Omega \sqrt{1 + A(u)}, \\ dt &= -\Omega \sqrt{1 - C(u)} du, \\ d(Ht) &= -\sqrt{1 - C(u)} d[\ln(Hu)]. \end{aligned} \quad (6.27)$$

Using this we can find the time dependence of the effective Hubble parameter:

$$H_{eff}(t) = \frac{d}{dt} \ln(R(t)) = \frac{H}{\sqrt{1-C(u)}} \left(1 - \frac{\frac{1}{2}u \frac{d}{du} A(u)}{1+A(u)} \right). \quad (6.28)$$

Note that a priori the signs in (6.27) are arbitrary. However, since:

$$\frac{1}{Hu} \equiv e^{Ht} \quad \rightarrow \quad Hdt = -\frac{1}{Hu} du, \quad (6.29)$$

for constant H , so the minus sign in relating dt and du is correct. A plus sign would take us back in time. Indeed, using that, we find that the effective Hubble parameter in (6.28) increases.

The plus-sign in the identification of the scale factor $R(t)$, is a priori less justified. Replacing it with a minus-sign would similarly lead to an increasing Hubble parameter. For now let us stick with the convention of Tsamis and Woodard.

One then infers that the backreaction of the IR gravitons acts to screen the bare cosmological constant, originally present. The improved results² in terms of:

$$\epsilon \equiv \left(\frac{\kappa H}{4\pi} \right)^2 = \frac{G\Lambda}{3\pi} = \frac{8}{3} \left(\frac{M}{M_P} \right)^4 \quad (6.30)$$

with $G\Lambda \sim 10^{-11}$ for GUT-scale inflation and $G\Lambda \sim 10^{-67}$ for EW-scale inflation, turn out to be:

$$A(u) = \epsilon^2 \left\{ \frac{172}{9} \ln^3(Hu) + (\text{subleading}) \right\} + \mathcal{O}(\epsilon^3), \quad (6.31)$$

$$C(u) = \epsilon^2 \left\{ \frac{57}{3} \ln^2(Hu) + (\text{subleading}) \right\} + \mathcal{O}(\epsilon^3) \quad (6.32)$$

Using (6.27) and (6.31) one finds:

$$Ht = - \left\{ 1 - \frac{19}{2} \epsilon^2 \ln^2(Hu) + \dots \right\} \ln(Hu) \quad (6.33)$$

using that the correction is smaller than unity. This can be inverted to give:

$$\ln(Hu) = - \left(1 + \frac{19}{2} (\epsilon Ht)^2 + \dots \right) Ht, \quad (6.34)$$

and it follows that $\ln(Hu) \sim -Ht$, for as long as perturbation theory is valid. The equivalence is a direct result of using conformal time u :

$$\frac{1}{Hu} \equiv e^{Ht} \quad \rightarrow \quad \ln(Hu) \equiv -Ht. \quad (6.35)$$

This implies that for as long as perturbation theory is valid, H , and therefore Λ , remain constant. Significant changes in the value of Λ must be non-perturbative.

However, we can now use the above equivalence to write $A(u)$ as:

$$A(u) = -\frac{172}{9} \epsilon^2 (Ht)^3 + \dots \quad (6.36)$$

²Papers before 1997 yield different results.

and one arrives at:

$$\begin{aligned} H_{eff}(t) &\approx H + \frac{1}{2} \frac{d}{dt} \ln(1 + A), \\ &\approx H \left\{ 1 - \frac{\frac{86}{3} \epsilon^2 (Ht)^2}{1 - \frac{172}{9} \epsilon^2 (Ht)^3} \right\} \end{aligned} \quad (6.37)$$

The induced energy density, which acts to screen the original cosmological constant present gives:

$$\begin{aligned} \rho(t) &\approx \frac{\Lambda}{8\pi G} \left\{ -\frac{1}{H} \frac{\dot{A}}{1+A} + \frac{1}{4H^2} \left(\frac{\dot{A}}{1+A} \right)^2 \right\} \\ &\approx \frac{\Lambda}{8\pi G} \left\{ -\frac{\frac{172}{3} \epsilon^2 (Ht)^2}{1 - \frac{172}{9} \epsilon^2 (Ht)^3} + \left(\frac{\frac{86}{3} \epsilon^2 (Ht)^2}{1 - \frac{172}{9} \epsilon^2 (Ht)^3} \right)^2 \right\} \end{aligned} \quad (6.38)$$

This can be written more intuitively, to better see the magnitude of the effect as follows:

$$H_{eff}(t) = H \left\{ 1 - \epsilon^2 \left[\frac{1}{6} (Ht)^2 + (\text{subleading}) \right] + \mathcal{O}(\kappa^6) \right\} \quad (6.39)$$

and the induced energy density and pressure, in powers of H :

$$\begin{aligned} \rho(t) &= \frac{\Lambda}{8\pi G} + \frac{(\kappa H) H^4}{2^6 \pi^4} \left\{ -\frac{1}{2} \ln^2 a + \mathcal{O}(\ln a) \right\} + \mathcal{O}(\kappa^4) \\ p(t) &= -\frac{\Lambda}{8\pi G} + \frac{(\kappa H) H^4}{2^6 \pi^4} \left\{ \frac{1}{2} \ln^2 a + \mathcal{O}(\ln a) \right\} + \mathcal{O}(\kappa^4), \end{aligned} \quad (6.40)$$

where in order to derive the expression for the pressure p , again the stress-energy conservation equation (6.12) with $\dot{\rho} = 0$, is used. We have argued that in our opinion, it is not a priori clear that this equation is satisfied. The energy density and pressure generate a non-local energy-momentum tensor, that is not covariantly conserved, since Λ has become time-dependent. At each instant of time, the Einstein equations will be satisfied with a constant Λ and a covariantly conserved energy-momentum tensor. However, this tensor will than also be constant.

However, let us continue the discussion. Since the effect is so weak, the number of e-folds of inflation needed is enormous. Recall that this quantity is defined as follows. For a pure de Sitter phase, we have:

$$\frac{\ddot{a}}{a} = \left(\frac{\dot{a}}{a} \right)^2 = \frac{\Lambda}{3} \equiv H_\Lambda^2, \quad \rightarrow \quad a(t) = a_1 e^{H_\Lambda(t-t_1)} \quad (6.41)$$

with $a(t) = a_1$ at $t = t_1$. During an inflationary period Δt , the size of the universe increases by a factor:

$$Z = e^{H_\Lambda \Delta t}. \quad (6.42)$$

To solve the flatness problem one must require $Z > 10^{30}$. The number N of e-folds now is defined as $Z \equiv e^N$, or, assuming inflation lasts from time t_0 to time t_1 :

$$N = \int_{t_0}^{t_1} H dt. \quad (6.43)$$

In order for $Z > 10^{30}$, $N > 69$. In the scenario by TW, this number is much bigger, but that is not in conflict with any known result.

The number of e-foldings needed to make the backreaction effect large enough to end inflation is argued to be:

$$N \sim \left(\frac{9}{172}\right)^{\frac{1}{3}} \left(\frac{3\pi}{G\Lambda}\right)^{\frac{2}{3}} = \left(\frac{81}{11008}\right)^{\frac{1}{3}} \left(\frac{M_P}{M}\right)^{\frac{8}{3}} \quad (6.44)$$

where M is the mass scale at inflation and M_P is the Planck mass. For inflation at the GUT scale this gives $N \sim 10^7$ e-foldings. This enormously long period of inflation, much longer than in typical inflation models, is a direct consequence of the fact that gravity is such a weak interaction. It results in a universe that is much bigger compared to ordinary scenarios of inflation, but that in itself is not a problem.

It is argued that the effect might be strong enough to effectively kill the ‘bare’ cosmological constant, as long as such a long period of inflation is acceptable. There do exist arguments that the number of e-folds is limited to some 85, see [266] for details, but these are far from established. Moreover, this bound is achieved on the assumption that at late times the acceleration is given by a pure cosmological constant, assuming that at late times the universe enters an asymptotic de Sitter phase which can store only a limited amount of entropy in field theoretic degrees of freedom. Tsamis and Woodard however, argue that eventually the cosmological constant will be screened all the way to zero, in which case this bound would not be applicable.

However, as we already argued, the manipulations in equations (6.33) and (6.34) are only allowed for as long as the Hubble constant is indeed constant. Perturbative techniques break down when the effect becomes too strong. This makes it very difficult to conclude what happens after a large number of e-folds. For as long as the calculations might be reliable, nothing really happens.

Tsamis and Woodard argue that the breaking down of perturbation theory is rather soft, since each elementary interaction remains weak. Furthermore, a technique following Starobinski [267] is used in which non-perturbative aspects are absorbed in a stochastic background that obeys the classical field equations [259].

It is then argued [259] that eventually the screening must overcompensate the original bare cosmological constant, leading to a period of deflation. This happens because the screening at any point derives from a coherent superposition of interactions from within the past lightcone and the invariant volume of the past lightcone grows faster as the expansion slows down. Now thermal gravitons are produced that act as a thermal barrier, that grows hotter and denser as deflation proceeds. Incoming virtual IR modes scatter off this barrier putting a halt to the screening process. The barrier dilutes and the expansion takes over again.

Throughout the above calculation, the ‘primordial’ cosmological constant Λ was used. The mechanism, however, is argued to screen the cosmological constant, which implies that the effective cosmological constant should be used instead, once the Hubble constant starts to decrease. The strength of the effect would then be even weaker, since this is controlled by $G\Lambda$. Moreover, although more gravitons would enter the past lightcone, once inflation starts to end, all these gravitons are redshifted to insignificance. It therefore appears impossible to use this mechanism to even end inflation, let alone to argue that today’s cosmological constant is zero.

6.5 Renormalization in Perturbative QG

As we already indicated above, there also appears to be a more fundamental problem with this scenario. This has to do with the very existence of growing infrared divergences and the possibility, according to Tsamis and Woodard of quantum gravity being capable of inducing strong effects at large distances. Crucial in this argument is the renormalization subtraction one should use in perturbative gravity.

The results obtained by Tsamis and Woodard are based on two-loop 1-point functions. They expand the full metric as:

$$g_{\mu\nu} \equiv \hat{g}_{\mu\nu} + \kappa h_{\mu\nu} \quad (6.45)$$

where the split between $g_{\mu\nu}$ and $h_{\mu\nu}$ is determined by requiring that the vacuum expectation value of $h_{\mu\nu}$ vanishes. The background field $\hat{g}_{\mu\nu}$ is assumed to satisfy the classical equations of motion $\hat{R}_{\mu\nu} = \Lambda \hat{g}_{\mu\nu}$. It is then argued that two-loop processes lead to growing infrared divergences which break De Sitter invariance, such that the vacuum expectation value of $h_{\mu\nu}$ becomes time dependent, instead of just some number times the De Sitter metric. Any such number could have been incorporated in the cosmological counterterm $\delta\Lambda$ [257], and no screening effect could be seen.

The growing IR divergences they get, indicates that one arrives at a state that is filled with long wavelength particles. These are the gravitons that continue to redshift beyond the horizon as inflation and the graviton production mechanism continues. However, these gravitons pass the horizon and cannot causally influence the spacetime of the observer. Hence, they should also not be accounted for in the setup. Note in this respect also that the Bunch-Davies vacuum used by Tsamis and Woodard to define perturbation theory, is the natural vacuum choice for particles of large mass $m \gg H^{-1}$. In that case, the Compton wavelength of the particle is small compared to the local radius of curvature of spacetime. However, Bunch-Davies vacuum is in general not a good choice when dealing with massless particles, as this generally leads to unphysical infrared divergences [30, 268].

In fact, Tsamis and Woodard use the cosmological counterterm to remove primary divergences in the amputated 1-point function only, to make the initial Hubble constant time-independent: $H_{eff}(0) = H$. The functions $A(u)$ and $C(u)$ in (6.26) and their first derivatives are therefore zero at $u = H^{-1}$. They then choose [257] the condition for $\delta\Lambda$ in terms of the initial values of $a(u)$ and $c(u)$ to be:

$$\left(\frac{d}{du}\right) A(u)|_{u=H^{-1}} = a(H^{-1}) - c(H^{-1}) = 0. \quad (6.46)$$

They notice that time-dependent coordinate transformations “can be used to impose an independent condition on $a(u)$ and $c(u)$ on any instant u . A nice example would be $a(u) = -c(u)$ ”. Note that this latter condition enforces exact de Sitter invariance, with $a(u)$ and $c(u)$ constants and, because of the initial condition $\dot{\Lambda}_{eff}(0+)$, equal to zero [257]. Insisting that the symmetries of De Sitter spacetime survive, necessarily implies that there can be no screening.

Another disturbing point lies in the fact that the cosmological counterterm is assumed to be time-independent. As stated before, usually tadpoles are removed by the substitution:

$$h_{\mu\nu} \rightarrow h_{\mu\nu} + a_{\mu\nu}, \quad (6.47)$$

with $a_{\mu\nu}$ taken to be such that the linear terms in the Lagrangean vanish. In gravity, this $a_{\mu\nu}$ generally has to be time-dependent in order to achieve this. With a time-dependent cosmological counterterm all tadpole diagrams vanish and they can cause no physical effect.

A different setup would therefore be to impose from the start:

$$\Gamma = \mathcal{L} + \delta\mathcal{L}, \quad \frac{\partial\Gamma}{\partial g} = 0 \quad (6.48)$$

with $\delta\mathcal{L}$ the counter Lagrangean. This condition would immediately subtract off all tadpole diagrams to any order in perturbation theory. Not imposing this condition would result in very ugly looking graviton propagators, with tadpole diagrams attached to them. However, this is not taken into account by Tsamis and Woodard; they use the ‘clean’ De Sitter propagator, but without imposing the condition (6.48).

Of course, when one tries to do perturbation theory on an initially de Sitter background and forces the corrections to be not de Sitter invariant, large IR divergences will occur. They will be growing, since the difference from the background becomes larger and larger. This is however not a physical effect.

We believe that in such a setup the cosmological constant does not become a time-dependent parameter, let alone that it be screened to zero. At a somewhat deeper level, since these gravitons all cross the horizon as long as inflation proceeds, one might ask whether it is possible at all that boundary effects may change the value of the cosmological constant. We do not believe that this is possible. One can always limit oneself to a box of dimensions smaller than the Hubble scale of the universe at that time and measure the cosmological constant. It is a locally determined parameter that should not depend on boundary effects.

6.6 Summary

In this chapter we have examined an interesting proposal by Tsamis and Woodard, that intends to use the quantum gravitational back-reaction of inflationary produced gravitons to screen the cosmological constant. However, we conclude that the supposed screening is at best completely negligible in magnitude and offers no solution to the cosmological constant problem. The infrared divergences are non-physical and can be removed by a proper renormalization subtraction. Besides, we believe that the induced energy density is not of the form of vacuum energy density.

These conclusions concur with our intuition that IR effects in gravity are virtually non-existent. Newton’s constant has dimension $[\text{GeV}]^2$, which implies by dimensional analysis that it will always be accompanied by (at least) two derivatives, cutting off any possible large distance effects.

Another remark to be made, is that one is free to add a counterterm $\delta\Lambda$ of strength $1/\text{cm}^2$. When the Hubble radius of the universe is of this length, this term will be the dominating energy density. However, the mechanism with inflationary produced gravitons at those scales is even weaker, and so it seems that for the present day cosmological constant, this scenario, even aside from our other objections, cannot work.

7 | Type III: Violating the Equivalence Principle

An intriguing way to try to shed light on the cosmological constant problem is to look for violations of the equivalence principle of general relativity. The near zero cosmological constant could be an indication that vacuum energy contrary to ordinary matter-energy sources does not gravitate.

The approach is based not on trying to eliminate any vacuum energy, but to make gravity numb for it. This requires a modification of some of the building blocks of general relativity. General covariance (and the absence of ghosts and tachyons) requires that gravitons couple universally to all kinds of energy. Moreover, this also fixes uniquely the low energy effective action to be the Einstein-Hilbert action. If gravity were not mediated by an exactly massless state, this universality would be avoided. One might hope that vacuum energy would then decouple from gravity, thereby eliminating its gravitational relevance and thus eliminating the cosmological constant problem.

7.1 Extra Dimensions, Braneworld Models

Since the Casimir effect troubles our notion of a vacuum state, the cosmological constant problem starts to appear when considering distances smaller than a millimeter or so. This size really is a sort of turn-over scale. Somehow all fluctuations with sizes between a Planck length and a millimeter are cancelled or sum up to zero. Therefore, extra dimensions with millimeter sizes might provide a mechanism to understand almost zero 4D vacuum energy, since in these scenarios gravity is changed at distances smaller than a millimeter.

A lot of research in this direction in recent years has been devoted to braneworld models in $D = 4 + N$ dimensions, with N extra spatial dimensions. In this setting the cosmological constant problem is at least as severe as in any other, but new mechanisms of cancelling a vacuum energy can be thought of. The general idea is that the observed part of our world is confined to a hypersurface, a 3-brane, embedded in a higher dimensional spacetime. The standard model fields are restricted to live on this 3-brane, while only gravitons can propagate in the full higher dimensional space. To reproduce the correct 4-dimensional gravity at large distances three approaches are known. Usually one takes the extra, unseen dimensions to cover a finite volume, by compactifying them. One of the earliest approaches was by Rubakov and Shaposhnikov [269] who unsuccessfully tried to argue that the 4D cosmological constant is zero, since 4D vacuum energy only curves the extra dimensions.

Besides, it is conceivable that the need to introduce dark matter, and a very small cosmological constant or some other form of dark energy to explain an accelerating universe

nowadays, is in fact just a signal of general relativity breaking down at very large distance scales. General relativity however, works very well on scales from 10^{-1} mm to at least 10^{14} cm, the size of the solar system. Our challenge is to modify gravity in the IR regime in such a way that the results of GR are not spoiled on those intermediate distances at which it works correctly.

Extra-dimensional models, like the early Kaluza-Klein scenarios, generically have additional degrees of freedom, often scalar fields, that couple to the four dimensional energy-momentum tensor and modify four-dimensional gravity. A four dimensional massless graviton has two physical degrees of freedom, a five dimensional one five, just like a massive 4-dimensional one¹. There are however, strong experimental constraints on such scalar-tensor theories of gravity. One can for example calculate the slowing down of binary pulsars due to the radiation of these gravi-scalars as they are sometimes called. It was shown in [270, 271] that in case of a 5-dimensional bulk, this leads to a modification of the quadrupole formula by 20 %, while observations agree with the quadrupole formula to better than $\frac{1}{2}$ %. In case of more extra dimensions, there will be more gravi-scalars and the problem only gets worse.

Usually, this is circumvented by giving these scalars a mass. In infinitely large, uncompactified extra dimensions there is another way out, since in these models the gravi-scalar represents a non-normalizable and therefore unphysical mode.

In this chapter we will first briefly review the Randall-Sundrum models and show why they cannot solve the cosmological constant problem. Next we focus on the DGP-model with infinite volume extra dimensions. This is a very interesting setup, but it also illustrates very well the difficulties one faces in deconstructing a higher dimensional model to a viable 4D world meeting all the GR constraints. A rather more speculative but perhaps also more promising approach is subsequently discussed, in which Lorentz invariance is spontaneously broken to yield a Higgs mechanism analog for gravity. Before concluding with a summary, we discuss yet another option, where one considers the graviton to be a composite particle.

7.2 Randall-Sundrum Models, Warped Extra Dimensions

There are in fact two different models known as Randall-Sundrum models, dubbed RS-I and RS-II. We begin with RS-I.

This model consists of two 3-branes at some distance from each other in the extra dimension. One brane, called the “hidden brane” has positive tension, while the other one, the “visible brane”, on which we live, has negative tension. Both branes could have gauge theories living on them. All of the Standard Model fields are localized on the brane, and only gravity can propagate through the entire higher dimensional space.

The equation of motion looks as follows:

$$M_*\sqrt{G}\left(R_{AB}-\frac{1}{2}G_{AB}R\right)-M_*\Lambda\sqrt{G}G_{AB}$$

¹In general, the total number of independent components of a rank 2 symmetric tensor in D dimensions is $D(D+1)/2$, however, only $D(D-3)/2$ of those correspond to physical degrees of freedom of a D -dimensional massless graviton; the remaining extra components are the redundancy of manifestly gauge and Lorentz invariant description of the theory.

$$= T_{hid}\sqrt{g_{hid}}g_{\mu\nu}^{hid}\delta_A^\mu\delta_B^\nu\delta(y) + T_{vis}\sqrt{g_{vis}}g_{\mu\nu}^{vis}\delta_A^\mu\delta_B^\nu\delta(y-y_0), \quad (7.1)$$

where

$$g_{\mu\nu}^{hid}(x) = G_{\mu\nu}(x, y = 0), \quad g_{\mu\nu}^{vis}(x) = G_{\mu\nu}(x, y = y_0), \quad (7.2)$$

and M_* is the higher dimensional Planck scale and Λ denotes the bulk cosmological constant.

The y -direction is compactified on an orbifold S_1/\mathbf{Z}_2 . With the above assumptions for the brane tensions and bulk CC, it can be shown that when the bulk is taken to be a slice of AdS_5 , there exists the following static solution, with a flat 4D-metric:

$$ds^2 = e^{-|y|/L}\eta_{\mu\nu}dx^\mu dx^\nu + dy^2 \quad (7.3)$$

with L the size of the extra dimension. The minus sign in the exponential factor occurs because of the assumption that our visible brane has a negative tension. As a result of this ‘warp-factor’, all masses on the visible brane are suppressed, compared to their natural value. For the Higgs mass for example, one obtains:

$$m^2 = e^{-y_0/L}m_0^2 \quad (7.4)$$

a small hierarchy in y_0/L therefore leads to a large hierarchy between m and m_0 , which would solve the ‘ordinary’ hierarchy problem, of the quadratically diverging Higgs mass.

Moreover, despite the fact that the brane tension on the visible brane is negative, it is possible that it still has a flat space solution. Fine-tuning is necessary to obtain this result, and besides, this solution is not unique. Other, non-flat space solutions also exist. Therefore, this cannot help in solving the cosmological constant problem, but it is interesting to see that a 4D cosmological constant can be made to curve only extra dimensions.

Alternatively, the extra dimensions can be kept large, uncompactified, but warped, as in the Randall-Sundrum type-II models. In this case the size of the extra dimensions can be infinite, but their volume $\int dy\sqrt{G} \sim L$, is still finite. Note that this cannot be obtained by a coordinate transformation of the RS-I model, with the hidden brane at infinity. The warp-factor causes the graviton wavefunction to be peaked near the brane, or, in other words, gravity is localized, such that at large 4D-distances ordinary general relativity is recovered.

The action now reads:

$$S = \frac{1}{2}M_*^3 \int d^4x \int_{-\infty}^{+\infty} dy\sqrt{G}(R_5 - 2\Lambda_5) + \int d^4x\sqrt{g}(\mathcal{E}_4 + \mathcal{L}_{SM}), \quad (7.5)$$

where \mathcal{E}_4 denotes the 4D brane tension and Λ_5 the bulk cosmological constant, which is assumed to be negative. The equation of motion derived from this action, ignoring now \mathcal{L}_{SM} is:

$$M_*\sqrt{G} \left(R_{AB} - \frac{1}{2}G_{AB}R \right) = -M_*^3\Lambda_5\sqrt{G}G_{AB} + \mathcal{E}_4\sqrt{g}g_{\mu\nu}\delta_A^\mu\delta_B^\nu\delta(y), \quad (7.6)$$

indicating that the brane is located at $y = 0$. This equation has the same flat space solution as above, with warp-factor $\exp(-|y|)$, where L now is defined as:

$$L \equiv \sqrt{-\frac{3}{2\Lambda_5}}, \quad \mathcal{E}_4 = \frac{3M_*^3}{L}, \quad (7.7)$$

but, again, at the expense of fine-tuning Λ_5 and \mathcal{E}_4 .

Gravity in the 4D subspace reduces to GR up to some very small Yukawa-type corrections. Unfortunately however, with regard to the cosmological constant problem, the model suffers from the same drawbacks as RS-I. All fundamental energy scales are at the TeV level, but the vacuum energy density in our 4D-world is much smaller.

For a recent overview of brane cosmology in such scenarios, see [272, 273].

7.2.1 Self-Tuning Solutions

Transmitting any contribution to the CC to the bulk parameters, in such a way that a 4D-observer does not realize any change in the 4D geometry seems quite suspicious. It would become more interesting if this transmission would occur automatically, without the necessity of re-tuning the bulk quantities by hand every time the 4D vacuum energy changes. Models that realize this are called *self-tuning* models (see for example [274] for an overview).

The literature is full of such proposals, however we have found that all of them contain ‘hidden’ fine-tunings. The real difficulty is always that no mechanism can be provided to single out flat 4D metrics from slightly curved 4D (A)dS solutions.

One of the first ideas were presented in [275, 276], where it is argued that the brane tension is compensated for by a change in integration constants for fields living in the bulk. These models live in five dimensions and contain a scalar field ϕ in the bulk, that is assumed to couple to the brane vacuum energy with a potential $\Lambda e^{-\kappa\phi}$. For any value of the brane tension a solution with a flat brane metric can be found, with a warped bulk geometry. However, expanding and contracting solutions are also allowed [277] and moreover, the flat solutions suffers from having naked singularities in the bulk. Subsequently, efforts have been made to hide the singularities behind event horizons in [278, 279], but in this case self-tuning of vacuum energy is lost, and the model contains hidden fine-tunings in order to preserve a flat brane metric [280, 281, 282]. These models therefore clearly do not work.

A related approach, considering a warped higher dimensional geometry, is studied in refs. [283, 284, 285, 286]. It is argued that once a cosmological constant vanishes in the UV, there exist solutions such that it will not be regenerated along the renormalization group flow. Any vacuum energy is cancelled by a decreasing warp factor, ensuring a flat space solution on the brane. The renormalization group scale in 4D gauge theory is interpreted as the compactification radius of the full gravity theory. However, these are not the only solutions and there exists no argument why they should be preferred. Note however, that this is quite contrary to ordinary renormalization group behavior, as studied in section 5.3.

Since five-dimensional models do not seem to work, much focus has been put on six-dimensional approaches, with a brane located on a conical singularity. The four-dimensional vacuum energy on the brane creates a deficit angle in the bulk, and it is argued that any change in brane tension will be compensated for by a change in deficit angle, see [287, 288, 289, 290, 291, 292].

The occurrence of this deficit angle can be seen explicitly as follows [293], where for now we restrict ourselves to two extra dimensions, and for simplicity no Einstein-Hilbert

term on the brane. The bulk Einstein equations are:

$$\sqrt{-G} \left(\mathcal{R}_{AB} - \frac{1}{2} G_{AB} \mathcal{R} \right) = \frac{\mathcal{E}_4}{4M_*^4} G_{A\mu} \eta^{\mu\nu} G_{\nu B} \delta^2(X^a - Y^a), \quad (7.8)$$

taking the 3-brane to be embedded at the point Y^a and small Latin superscripts denote coordinates running only over the two extra dimensions. Assuming now a metric ansatz:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu + \mathcal{G}_{ab}(X^4, X^5) dX^a dX^b, \quad (7.9)$$

Einstein's equation splits up to:

$$\begin{aligned} \sqrt{\mathcal{G}} R^{(2)} &= -\frac{\mathcal{E}_4}{2M_*^4} \delta^2(X^a - Y^a) \\ R_{ab} - \frac{1}{2} R^{(2)} \mathcal{G}_{ab} &= 0 \end{aligned} \quad (7.10)$$

with $R^{(2)}$ the two-dimensional scalar curvature. The first of these equations, has as a solution \mathcal{G}_{ab} corresponding to a conical geometry on the two-dimensional space, with the tip of the cone at Y^a , and deficit angle given by:

$$\delta = \frac{\mathcal{E}_4}{4M_*^4}, \quad (7.11)$$

which for many values of M_* is only smaller than 2π if the brane tension is fine-tuned, which is just the re-incarnation of the cosmological constant problem². So apart from exact supersymmetry at TeV energies, one also has to require the higher dimensional Planck mass to be of at least TeV level.

Another important issue is that in fact as far as we know, all the models still have hidden fine-tunings. To make these explicit, consider the bulk action of [287]:

$$S_6 = \int d^6x \sqrt{G} \left(\frac{1}{2} M_6^4 R - \Lambda_6 - \frac{1}{4} F_{ab} F^{ab} \right) \quad (7.12)$$

where now there is a 2-form field strength in the bulk, which is taken to generate a magnetic flux: $F_{ij} = \sqrt{\gamma} B_0 \epsilon_{ij}$, with γ_{ij} the higher 2-dimensional metric, B_0 just a constant and all other components of F_{ij} are assumed to vanish. A flat solution with metric:

$$ds^2 = G_{ab} dX^a dX^b = \eta_{\mu\nu} dx^\mu dx^\nu + R_0^2 (d\theta^2 + \alpha^2 \sin^2 \theta d\phi^2), \quad (7.13)$$

can be obtained, with deficit angle $\delta \equiv 2\pi(1 - \alpha)$, where Greek letters run over four-dimensional coordinates x , where Latin letters run also over all higher coordinates y . This flat metric is obtained when:

$$2\Lambda_6 = \frac{1}{2} B^2, \quad \frac{1}{R_0^2} = \frac{B^2}{2M_4^4}. \quad (7.14)$$

In order to have a self-tuning solution, the deficit angle is adjusted whenever there is a change in four-dimensional vacuum energy. However, it can easily be seen that this is

²Typically, from stringtheory the brane tension of a 3-brane is given by: $T = 1/[(2\pi)^3 (l_s)^4 g_s]$ with l_s the string length, typically a Planck length, and g_s the string coupling. Normally, with $T \sim M_W^4$, with M_W the weak scale, as often assumed in these models, this implies that $1/g_s$ has to be extremely small, not even in the perturbative regime. However, in this setup also the string length l_s is modified.

not the case, since the magnetic flux is given by a closed form, which after integration must be the same, before and after the change in Λ . The deficit angle δ is related to Λ_6 as follows:

$$\alpha^2 = \frac{\Phi_B^2 \Lambda_6}{M_4^4}, \quad (7.15)$$

with Φ_B the magnetic flux integrated over the extra dimensional space. There obviously is no self-tuning, the right-hand-side has to be static if one maintains the same Planck mass before and after a change in the 4D cosmological constant. Similar conclusions were reached by other authors [294, 274, 295]. In [295] on rather general terms it is shown that many of these models fail, since the effective four-dimensional theory has a finite number of fields below the cutoff energy, and face Weinberg's no-go theorem again, which we discussed in section (1.5.1).

Moreover, a severe drawback that so far all these models face is that this scenario does not exclude '*nearby curved solutions*' [274]. This simply means that also solutions for neighboring values of some bulk parameters are allowed, which result in a curved 4D space, either expanding or contracting. Besides, there are additional problems such as a varying effective Planck mass, or varying masses for fields on the brane. And last but not least, the flat space solution generically has naked singularities in the bulk. We will return to this issue in section (7.3), which suffers from the same problem. So far no self-tuning scenario without these drawbacks has been found. One of the most recent papers in this direction is [296] in which it is acknowledged that this scenario has serious flaws.

The approach of section (7.3) benefits from the fact that general relativity is modified at large distances. Among other things, this implies that the low energy effective theory does not have a finite number of fields below some low energy cutoff. This might provide a way out of the drawbacks the models in this section suffer from. But, in fact, it turns out that also this more sophisticated model generalized to N extra dimensions, suffers from the same serious drawbacks.

7.2.2 Extra Time-like Dimensions

For completeness let us here also briefly mention approaches using extra time-like dimensions in $D = 11$ supergravity, e.g. [297]. It is argued that classical vacuum solutions can be obtained with zero cosmological constant and without massless ghosts or tachyons in the low energy limit. There are however many different solutions with different characteristics so the predictive power seems to be minimal. Moreover, the usual problems arise after supersymmetry breaking, and after taking higher orders into account.

7.3 Infinite Volume Extra Dimensions

In [298, 299, 300, 301, 43, 302], a model based on infinite volume extra dimensions is presented. We will first give a qualitative idea of what this model looks like, before going into more details. In the literature it has become known as the DGP-model, after the founding fathers, Dvali, Gabadadze and Porrati.

In this setup there is just one 3-brane to which all standard model fields are confined. Only gravity can propagate in the full higher dimensional spacetime. It is assumed that the higher-dimensional theory is supersymmetric and that SUSY is spontaneously broken on the brane. These breaking effects can be localized on the brane only, without affecting the bulk, because the infinite volume gives a large enough suppression factor. Apart from that, an unbroken R-parity might be assumed to forbid any negative vacuum energy density in the bulk. The bulk cosmological constant therefore vanishes.

Gravity in the bulk is taken to be very strong; the higher dimensional Planck mass M_* , is assumed to be of order 10^{-3} eV. As a result, four dimensional gravity is modified in the ultraviolet at distances $r \lesssim M_*^{-1} \sim 0.1$ mm, where the effective field theory description of gravity breaks down. Since gravity is so strong at these small distances, its short-distance nature depends on the UV-completion of the theory. This UV-completion is not known at present, but could possibly be some string theory embedding. As argued before, this is interesting, since this distance scale could be a defining scale for the cosmological constant problem. In fact, graviton loops are cutoff at M_* and hence, do not contribute to the cosmological constant problem, see also section 7.5.

Moreover, gravity is also modified in the infrared. At distances smaller than a cross-over scale r_c , gravity looks four dimensional, whereas at distances larger than r_c it becomes higher dimensional. The physical reason for this is that with infinite volume extra dimensions, there is no localizable zero-mode graviton. Instead, from a four-dimensional point of view, the 4D-graviton is a massive metastable state that can escape into the extra dimensions at very large distances. This can also be understood with regard to the Einstein-Hilbert terms in the action. Because the higher dimensional Planck mass is so small, the bulk Einstein-Hilbert term, becomes of comparable magnitude at very large distances, given by $r_c \sim M_P^2/M_*^3$, in case of one extra dimension. Gravity thus becomes weaker at large distances, which could possibly explain the observed acceleration.

The cosmological constant problem is addressed by noting that the full Einstein equations admit a flat space solution on the brane, despite the fact that the brane tension, or 4D cosmological constant, is non-zero. In this solution, vacuum energy will mainly curve the higher dimensions.

Embedding our spacetime in infinite volume extra dimensions thus has several advantages. If they are compactified, one would get a theory approaching GR in the IR, facing Weinberg's no-go theorem again. With infinite volume extra dimensions on the other hand, GR is not only modified in the UV, but also in the IR. This changes not only early cosmology, but also late-time cosmology, with perhaps a possibility of explaining the accelerated expansion without a cosmological constant. The fundamental, higher dimensional Planck scale, M_* , can be much smaller in this scenario than in ordinary models with extra dimensions. For finite volume extra dimensions, like standard Kaluza-Klein compactifications, or Randall-Sundrum warping (see section (7.2)): $M_P^2 = M_*^{2+N} V_N$, with V_N the volume of the extra dimensions, and N the number of extra dimensions. However, this relation no longer holds when $V_N \rightarrow \infty$. Therefore, M_* can be much smaller than a TeV, making gravity in the bulk much stronger.

Details on how these infinite volume models circumvent Weinberg's no-go theorem can be found in [301].

Description of the Model

The low-energy effective action is written:

$$S = M_*^{2+N} \int d^4x d^N y \sqrt{-G} \mathcal{R} + \int d^4x \sqrt{-g} (\mathcal{E}_4 + M_P^2 R + \mathcal{L}_{SM}) + S_{GH}, \quad (7.16)$$

where M_*^{2+N} is the $(4+N)$ -dimensional Planck mass, the scale of the higher dimensional theory, G_{AB} the $(4+N)$ -dimensional metric, y are the ‘perpendicular’ coordinates, and $\mathcal{E}_4 = M_{Pl}^2 \Lambda$, the brane tension, or 4D cosmological constant. Thus the first term is the bulk Einstein-Hilbert action for $(4+N)$ -dimensional gravity and the $M_P^2 R$ term is the induced 4D-Einstein-Hilbert action. So there are two free parameters: M_* and \mathcal{E} . M_* is assumed to be very small, making gravity in the extra dimensions much stronger than in our 4D world.

Furthermore, S_{GH} is a Gibbons-Hawking surface term. The Einstein equations follow from varying the Einstein-Hilbert action with a cosmological constant. Since the Ricci scalar involves second derivatives, if one considers a compact manifold with boundary ∂M , and allows variations of the metric that vanish on ∂M , but whose normal derivatives do not, one must add a surface term. This term is called the Gibbons-Hawking term and can be written as:

$$S_{GH} = \frac{1}{8\pi G} \int_{\partial M} d^3x \sqrt{h} K, \quad (7.17)$$

where K is the trace of the extrinsic curvature K_{ij} of the boundary three-surface, and h is the determinant of the metric induced on the three-surface.

The 4D-Planck mass in this setup is a derived quantity, which can be seen as follows [303]. Consider the following low energy effective action in 4 dimensions:

$$S_G = \int d^4x \sqrt{-g} \left(M_*^2 R + \sum_{n=1}^{\infty} C_n \frac{R^{n+1}}{M_*^{2(n-1)}} \right), \quad (7.18)$$

with a derivative expansion assumed in the second term and assume M_* is much smaller than M_P . At distances $r \gg M_*^{-1}$, the Einstein-Hilbert term dominates and $G_* \sim M_*^{-2} \gg G_N$, with G_N normal Newton’s constant. At shorter distances, we have to take into account the infinite series of terms in the sum in (7.18), and effective field theory breaks down.

However, the Einstein-Hilbert term is unstable under quantum corrections; it gets renormalized by matter loops, similar to the Higgs mass, when this gravity action is coupled to matter. These are loops with external gravitons with momentum $p < M_*$, since only in this region this theory of gravity makes sense. This generates an extra term in the action. With a cutoff on the standard model M_{SM} at GUT-scale for example, one obtains:

$$\Delta S_{induced} = M_{ind}^2 \int d^4x \sqrt{-g} (R + \mathcal{O} [R^2/M_{SM}^2]) \quad (7.19)$$

where the value of M_{ind} is given by [304, 305, 306]:

$$M_{ind}^2 = \frac{i}{96} \int d^4x x^2 \langle T_{SM}(x) T_{SM}(0) \rangle, \quad (7.20)$$

with on the right-hand-side the vacuum value of the time-ordered product of the trace of the energy-momentum tensor $T(x)$. The total action $S_G + \Delta S_{induced} + S_{SM}$ becomes:

$$S = \int d^4x \sqrt{-g} \left\{ (M_*^2 + M_{ind}^2) R + \sum_{n=1}^{\infty} C_n \frac{R^{n+1}}{M_*^{2(n-1)}} + \mathcal{L}_{SM} \right\} \quad (7.21)$$

where $M_{ind}^2 + M_*^2 = M_P^2$. Note that the higher derivative terms are still there, but suppressed by M_* . The four dimensional M_P thus is a derived quantity, its value is determined by the UV cutoff of the Standard Model, and by the content and dynamics of the Standard Model. Scalars and fermions of the Standard Model contribute to M_{ind} with a positive sign, and gauge fields with negative sign.

The same mechanism is used in this higher dimensional model. Since Standard Model particles can only propagate on the brane, and not in the bulk, any bulk loop gets cut-off at the scale M_* . In this way, through the induced Einstein-Hilbert term on the brane, gravity on the brane is *shielded* from the very strong gravity in the bulk. Submillimeter test of the $1/r^2$ -law of gravity indicate that $M_* \geq 10^{-3}$ eV.

Recovery of 4D-Gravity on the Brane

Gravity on the brane can be recovered either by making a decomposition into Kaluza-Klein modes, or by considering the 4D graviton as a resonance, a metastable state with a mass given by $m_g \sim M_*^3/M_{Pl}^2$.

Einstein's equation from (7.16) becomes (up to two derivatives):

$$M_*^{2+N} \left(\mathcal{R}_{AB} - \frac{1}{2} G_{AB} \mathcal{R} \right) + \delta^{(N)} M_P^2 \left(R - \frac{1}{2} g_{\mu\nu} R \right) \delta_A^\mu \delta_B^\nu = \mathcal{E}_4 \delta^{(N)}(y) g_{\mu\nu} \delta_A^\mu \delta_B^\nu. \quad (7.22)$$

The higher dimensional graviton can be expanded in 4D Kaluza-Klein modes as follows:

$$h_{\mu\nu}(x, y_n) = \int d^N m \epsilon_{\mu\nu}^m(x) \sigma_m(y_n), \quad (7.23)$$

where $\epsilon_{\mu\nu}^m(x)$ are 4D spin-2 fields with mass m and $\sigma_m(y_n)$ are their wavefunction profiles in the extra dimensions. Each of these modes gives rise to a Yukawa-type gravitational potential, the coupling-strength to brane sources of which are determined by the value of σ_m at the position of the brane, say $y = 0$:

$$V(r) \propto \frac{1}{M_*^{2+N}} \int_0^\infty dm m^{N-1} |\sigma_m(0)|^2 \frac{e^{-rm}}{r}. \quad (7.24)$$

However, in this scenario there is a cut-off of this integral; modes with $m > 1/r_c$ have suppressed wavefunctions, where r_c is some cross-over scale, given by $r_c = M_{Pl}^2/M_*^3 \sim H_0^{-1}$. For $r \ll r_c$ the gravitational potential is $1/r$, dominated by the induced 4D kinetic term, and for $r \gg r_c$ it turns to $1/r^2$, in case of one extra dimension. In ordinary extra dimensional gravity, all $|\sigma_m(0)| = 1$, here however:

$$|\sigma_m(0)| = \frac{4}{4 + m^2 r_c^2}, \quad (7.25)$$

which decreases for $m \gg r_c^{-1}$. Therefore, in case of one extra dimension, the gravitational potential interpolates between the 4D and 5D regimes at r_c . Below r_c , almost normal 4D gravity is recovered, while at larger scales it is effectively 5-dimensional and thus weaker. This could cause the universe's acceleration.

To derive (7.25), consider first the scalar part of the action (7.16) for one extra dimension [307]. The potential can be written:

$$V(r) = \int dt G_R(x, y = 0; 0, 0), \quad (7.26)$$

where G_R is the retarded Green's function, which after Fourier transformation becomes:

$$G_R(x, y, 0, 0) \equiv \int \frac{d^4 p}{(2\pi)^4} e^{ipx} \tilde{G}_R(p, y). \quad (7.27)$$

The equation for this Green's function in Euclidean momentum space is:

$$(M_*^3(p^2 - \partial_y^2) + M_P^2 p^2 \delta(y)) \tilde{G}_R(p, y) = \delta(y) \quad (7.28)$$

with p^2 the square of Euclidean 4-momentum. The solution, with appropriate boundary conditions becomes:

$$\tilde{G}_R(p, y) = \frac{1}{M_P^2 p^2 + 2M_*^3 p} e^{-p|y|}. \quad (7.29)$$

With proper normalization, the potential (7.26) takes the form:

$$V(r) = -\frac{1}{8\pi^2 M_P^2 r} \left\{ \sin\left(\frac{r}{r_c}\right) \text{Ci}\left(\frac{r}{r_c}\right) + \frac{1}{2} \cos\left(\frac{r}{r_c}\right) \left[\pi - 2\text{Si}\left(\frac{r}{r_c}\right) \right] \right\}, \quad (7.30)$$

where $\text{Ci}(z) \equiv \gamma + \ln(z) + \int_0^z (\cos(t) - 1) dt/t$, $\text{Si}(z) \equiv \int_0^z \sin(t) dt/t$, γ is the Euler-Mascheroni constant, and r_c is again the critical distance, defined as $r_c \equiv M_P^2/2M_*^3$. This again shows the cross-over behavior of r_c : Below r_c , the model is 4-dimensional, while at distances larger than r_c it is 5-dimensional.

The form of (7.25) can also be derived from consideration of the Green's function for the state (7.23). Imposing that the states $\epsilon_{\mu\nu}^m(x)$ are orthogonal to each other, we can write for the position space Green's function:

$$\begin{aligned} G(x - x', 0)_{\mu\nu, \gamma\delta} &= \langle h_{\mu\nu}(x, 0) h_{\gamma\delta}(x', 0) \rangle \\ &= \int d^N m |\sigma_m(0)|^2 \langle \epsilon_{\mu\nu}^m(x) \epsilon_{\gamma\delta}^m(x') \rangle, \end{aligned} \quad (7.31)$$

which after a Fourier transformation becomes the momentum space Green's function G . For its scalar part we have:

$$G(p, 0) = \int dm m^{N-1} \frac{|\sigma_m(0)|^2}{m^2 + p^2}, \quad (7.32)$$

which is the Källén-Lehman spectral representation of the Green's function:

$$G(p, 0) = \int ds \frac{\rho(s)}{s + p^2}, \quad \text{and} \quad \rho(s) = \frac{1}{2} s^{\frac{N-2}{2}} |\sigma_{\sqrt{s}}(0)|^2, \quad (7.33)$$

where $s \equiv m^2$.

Using again the equations of motion in Euclidean momentum space, the propagator becomes:

$$G(p, 0) = \frac{1}{M_P^2 p^2 + M_*^{2+N} D^{-1}(p, 0)}, \quad (7.34)$$

where $D^{-1}(p, 0)$ is the inverse Green's function of the bulk theory with no brane. For $N > 1$ and $p \gg r_c^{-1}$, the propagator becomes:

$$G(p, 0) \simeq \frac{1}{M_P^2 p^2}, \quad (7.35)$$

which is just the propagator of a massless four-dimensional graviton. In case of one extra dimension:

$$G(p, 0) = \frac{1}{M_P^2 p^2 + 2M_*^3 p}, \quad (7.36)$$

yielding the above result for (7.25).

The full tensorial structure is a bit more involved. We take again as a starting point the field equations, now with a brane energy-momentum tensor and expand for small metric fluctuations around flat empty space:

$$g_{AB} = \eta_{AB} + h_{AB}. \quad (7.37)$$

Now we have to pick a gauge, for example harmonic gauge in the bulk:

$$\partial^A h_{AB} = \frac{1}{2} \partial_B h_A^A, \quad (7.38)$$

where $h_{\mu 5} = 0$ and:

$$\square^{(5)} h_5^5 = \square^{(5)} h_\mu^\mu, \quad (7.39)$$

with $\square^{(5)}$ the 5D d'Alembertian. The $\mu\nu$ -component of the field equations then becomes after rearranging some terms [302]:

$$\left[\frac{1}{r_c} (\nabla^2 - \partial_y^2) + \delta(y) \nabla^2 \right] h_{\mu\nu} = -\frac{1}{M_P^2} \left\{ T_{\mu\nu} - \frac{1}{3} \eta_{\mu\nu} T_\alpha^\alpha \right\} \delta(y) + \delta(y) \partial_\mu \partial_\nu h_\alpha^\alpha \quad (7.40)$$

The second term on the r.h.s. vanishes when contracted with a conserved energy-momentum tensor, and hence is unimportant at the level of one graviton exchange. The potential $h_{\mu\nu}$ becomes:

$$\tilde{h}_{\mu\nu}(p, x^5 = 0) = \frac{8\pi}{M_P^2 p^2 + 2M_*^3 p} \left[\tilde{T}_{\mu\nu}(p) - \frac{1}{3} \eta_{\mu\nu} \tilde{T}_\alpha^\alpha(p) \right], \quad (7.41)$$

where the massive graviton behavior can be seen from the coefficient $1/3$, see (7.3.1).

Possible Solution to the Cosmological Constant Problem?

The question now is whether there exist solutions such that the 4D induced metric on the brane is forced to be flat: $g_{\mu\nu} = \eta_{\mu\nu}$. Einstein's equation from (7.16) reads:

$$M_*^{2+N} \left(\mathcal{R}_{AB} - \frac{1}{2} G_{AB} \mathcal{R} \right) + \delta^{(N)} M_P^2 \left(R - \frac{1}{2} g_{\mu\nu} R \right) \delta_A^\mu \delta_B^\nu = \mathcal{E}_4 \delta^{(N)}(y) g_{\mu\nu} \delta_A^\mu \delta_B^\nu. \quad (7.42)$$

In case of one extra dimension it is not possible to generate a viable dynamics with a flat 4D metric [302]. For $N = 2$ an analytic solution has been obtained which generates a flat 4D Minkowski metric [293]. The 4D brane tension is spent on creating a deficit angle in the bulk. The derivation of the occurrence of this deficit angle was given in section (7.2.1). As also argued there, one has to fine-tune this tension in order not to generate a deficit angle larger than 2π , since in this case $M_* \ll \text{TeV}$. So also the $N = 2$ model does not work.

For $N \geq 2$, solutions of the theory can be parameterized as:

$$ds^2 = A^2(y)g_{\mu\nu}(x)dx^\mu dx^\nu - B^2(y)dy^2 - C^2(y)y^2 d\Omega_{N-1}^2, \quad (7.43)$$

where $y \equiv \sqrt{y_1^2 + \dots + y_n^2}$ and the functions A, B, C depend on \mathcal{E}_4 and M_* :

$$A, B, C = \left(1 - \left(\frac{y_g}{y}\right)^{N-2}\right)^{\alpha, \beta, \gamma}, \quad (7.44)$$

where α, β, γ correspond to A, B, C respectively, and depend on dimensionality, and y_g is the gravitational radius of the brane:

$$y_g \sim M_*^{-1} \left(\frac{\mathcal{E}_4}{M_*^4}\right)^{\frac{1}{N-2}} \quad \text{for} \quad N \neq 2. \quad (7.45)$$

For $N > 2$ consistent solutions of the form (7.43),(7.44), do exist with a flat 4D metric, without the problem of having a too large deficit angle. Their interpretation however, is rather complicated because of the appearance of a naked singularity in the bulk at $y = y_g$. Spacetime in $4 + N$ dimensions looks like $\mathfrak{R}_4 \times S_{N-1} \times R_+$, where \mathfrak{R}_4 denotes flat spacetime on the brane, and $S_{N-1} \times R_+$ are Schwarzschild solutions in the extra dimensions. The Einstein equations cannot be satisfied at this singularity.

The final physical result is argued to be:

$$H \sim y_g^{-1} \sim M_* \left(\frac{M_*^4}{\mathcal{E}_4}\right)^{\frac{1}{N-2}}. \quad (7.46)$$

According to the 4D result, $N = 0$, the expansion rate grows as \mathcal{E}_4 increases, but for $N > 2$ the acceleration rate H decreases as \mathcal{E}_4 increases. In this sense, vacuum energy can still be very large, it just gravitates very little; 4D vacuum energy is supposed to curve mostly the extra dimensions.

This scenario has been criticized for different reasons. One immediate call for concern is the appearance of a naked singularity in the bulk. The situation can be compared with the perhaps more familiar cosmic strings. As we discussed in the previous section, a setup with a 3-brane embedded in two extra dimensions is rather similar to an ordinary cosmic string in the usual four dimensions. In both cases, the tension of the brane/string creates a deficit angle in the space around it. For more general embeddings, i.e. more than two extra dimensions, higher dimensional cosmic string examples provide good analogies.

Interestingly, in the case of local and global strings in four and six dimensions [308] solutions are known similar to the $N > 2$ DGP-model, which also suffer from a singularity. In that case, it was shown that the singularity is smoothed out, if the worldvolume

is assumed to be inflating, rather than static [309, 310, 311]; the singularity is replaced by a horizon. The situation in the $N > 2$ DGP-model is similar. In this case, the requirement of a flat-space solution on the brane is too strict.

The inflating solutions are indeed conjectured to be the only non-singular solutions [312, 301]. However, the exact analytic expressions for them have not yet been found, which makes it hard to make definite statements about them. In any case, the line element becomes:

$$ds^2 = -n^2(t, y)dt^2 + a_2^2(t, y)\gamma_{ij}(x)dx^i dx^j + b^2(t, y)dy^2. \quad (7.47)$$

The inflating brane metric is argued to generate a horizon exactly at y_g , if one would neglect the Einstein-Hilbert term on the brane, maintaining the desirable solution (7.46).

However, including the effect of the induced Einstein-Hilbert term, amounts to making a shift $\mathcal{E}_4 \rightarrow \mathcal{E}_4 - M_P^2 H^2$ [301], resulting in an inflation rate $H^2 \sim \mathcal{E}_4/M_P^2$, the same as in ordinary cosmology, and opposite to the result in 7.46. For this inflating solution there is no upper limit to the four-dimensional cosmological constant. The four-dimensional vacuum energy again curves strongly the four-dimensional worldvolume. Therefore we conclude that this scenario as it stands, does not solve the cosmological constant problem. In generating a well-behaved solution, i.e. one without naked singularities, immediately the cosmological constant problem returns.

Besides, even if the flat space solution were a genuine physical solution, this would not be enough. One would have to find a dynamical reason why the flat space solution would actually be preferred over the inflating ones. Such an argument does not (yet) exist.

Moreover, since gravity has essentially become massive in this scenario, the graviton has five degrees of freedom, and especially the extra scalar degree of freedom, often leads to deviations of GR at small scales. It indeed has been shown that the DGP-model suffers from this strong-coupling problem. We discuss this in more detail in the next section.

7.3.1 Massive Gravitons

A much studied approach to change general relativity in the infrared, which is not simply a variation of a scalar-tensor theory, is to allow for tiny masses for gravitons, like the Fierz-Pauli theory of massive gravity [313] and the example above³. The Fierz-Pauli (FP) mass term is the only possibility in four dimensions that has no ghosts in the linear theory. The mass term takes the form:

$$S_{PF} \propto M_P^2 m_g^2 \int d^4x \left(h_{\mu\nu}^2 - \frac{1}{2} (h^\mu{}_\mu)^2 \right), \quad (7.48)$$

where $h_{\mu\nu} \equiv g_{\mu\nu} - \eta_{\mu\nu}$. With this definition, the mass term is regarded as an exact term, not as a leading term in a small h expansion. Note in passing that due to mass terms, gravitons might become unstable and could possibly decay into lighter particles, for example photons. If so, gravity no longer obeys the standard inverse-square law, but becomes weaker at large scales, possibly leading to accelerated cosmic expansion.

³In [314] experimental bounds on graviton masses are discussed.

In general, the extra degrees of freedom, extra polarizations of a massive graviton, could also become noticeable at much shorter distances, putting severe constraints on such scenarios. In the UV the new scalar degrees of freedom may become strongly coupled, where the effective theory breaks down and the physics becomes sensitive to the unknown UV-completion of the theory.

Note that this is a general phenomenon. If we modify a non-Abelian gauge theory in the IR, by providing a mass m for the gauge boson, this introduces a new degree of freedom, the longitudinal component of the gauge field, which becomes strongly coupled in the UV at a scale m/g , determined by the scale of the IR-modification and g , the coupling constant. Only in a full Higgs mechanism, the couplings will remain weak, since this provides a kinetic term for the extra scalar degree of freedom.

A severe obstacle massive gravity theories have to overcome is something known as the Van Dam, Veltman, Zakharov, or (vDVZ), discontinuity [315, 316], which precisely is a manifestation of the strong coupling problem. The gravitational potential, represented by $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu}$, generated by a static source $T_{\mu\nu}$, like the sun, is given by:

$$\begin{aligned} h_{\mu\nu}^{massive}(q^2) &= -\frac{8\pi G}{q^2 + m^2} \left(T_{\mu\nu} - \frac{1}{3} \eta_{\mu\nu} T^\alpha_\alpha \right) \\ h_{\mu\nu}^{massless}(q^2) &= -\frac{8\pi G}{q^2} \left(T_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} T^\alpha_\alpha \right) \end{aligned} \tag{7.49}$$

in the massive and massless case. vDVZ argued that in the massive case, even with extremely small graviton mass, the bending of light rays passing near the sun would be too far off from experimental results, that the mass of the graviton has to be exactly equal to zero. The physical reason indeed being, that even in the limit where the mass of the graviton goes to zero, there is an additional scalar attraction, which distinguishes the theory from Einstein's GR. However, Vainshtein proposed that this discontinuity is just an artifact of using a linear approximation [317] and claimed that in the full nonlinear theory, the discontinuity no longer persists. The perturbative expansion in G breaks down. However, other problems also arise. Minkowski space does not seem to be stable and ghosts appear at the linear level [318, 319].

Since there exists no fully non-linear, generally covariant theory for massive gravity, Vainshtein calculated the non-linear completion of only the first linearized term, which becomes just the Einstein tensor, and the field equation reads:

$$G_{\mu\nu} + m_g^2 (h_{\mu\nu} - \eta_{\mu\nu} h^\alpha_\alpha) = 8\pi G T_{\mu\nu}. \tag{7.50}$$

Solutions to this equation are fully continuous and reproduce results of massless Einstein gravity, in the zero mass limit. However, deviations of standard Einstein gravity, become important at distance scales shorter than the Compton wavelength of the graviton, as one might have expected. The critical distance r_* is:

$$r_* = \frac{(m_g r_g)^{1/5}}{m_g} \tag{7.51}$$

where $r_g \equiv 2GM$ is the gravitational radius of the source of mass M .

However, this theory is still unsatisfactory, as it is not generally covariant. Higher dimensional models could possibly circumvent these problems and allow for graviton

masses. When the extra dimensions are compactified, as in the conventional Kaluza-Klein mechanisms, then from a four-dimensional perspective, there is one massless spin-2 state and a tower of massive gravitons. In the linearized approximation, these massive states have the FP-form. In this case a consistent theory is obtained as a result of an infinite number of 4-dimensional reparametrization invariances. Truncating this tower of massive states to any finite order leads to an explicit breakdown of these gauge invariances and again to inconsistencies [320, 321].

In the DGP model, the extra dimensions are infinitely large, and in the literature, there is an ongoing discussion whether this model is experimentally viable and capable of avoiding the massive gravity difficulties, see [322, 323, 324, 325, 326, 327] for criticism. It appears that indeed also this model suffers from strong interactions at short distances due to the scalar polarization of the massive graviton, that can be understood in terms of a propagating ghosts-like degree of freedom.

The IR deviations of GR take place at distances set by $r_c \equiv M_P^2/M_*^3$, in case of one extra dimension. The brane-to-brane propagator in DGP-gravity is very similar to that of massive gravity. The one-graviton exchange approximation breaks down at distances $R_* \sim (R_s r_c^2)^{1/3}$ [322], called the Vainshtein scale, with R_s the Schwarzschild radius of the source. R_* is very large for astrophysical sources, and it has been shown explicitly that at shorter distances the full non-linear solution approaches that of GR [328, 319, 329], which suggests that the DGP model may describe our universe. For distances larger than R_* gravity deviates significantly from GR, yet for smaller distances it should yield (approximately) the same results. Yet, the large distance at which deviations from GR are felt, indicates that the DGP-model has hidden strong interaction scales [322]. Their effects become important at much smaller distance scales, given by:

$$r_{crit} = \left(\frac{r_c^2}{M_P} \right)^{1/3}, \quad (7.52)$$

which can be as small as a 1000 km, for $r_c \sim H \sim 10^{28}$ cm. These strong interactions can be traced back to the appearance of a longitudinal Goldstone mode, that only obtains a kinetic term through mixing with the transverse polarization states of the graviton. This is in line with our general argument at the beginning of this section, that the absence of kinetic terms tends to result in strong coupling regimes. This mode, as observed by [322], can be interpreted as a brane bending mode, in the sense that it cannot be removed by a gauge transformation, without shifting the boundary. In other words, this mode excites the extrinsic curvature of the boundary, giving its shape as seen by a higher-dimensional observer.

Strong coupling, resulting in noticeable deviations of GR, at distances smaller than 10^3 km, is of course disastrous. This result, however, also argued for in [323, 330, 325], depends on the UV-regulator used, since it depends on the UV properties at the boundary. Since the UV-completion of the DGP-model is unknown, basically no one knows what happens exactly and one might hope that the UV-completion will miraculously cure the model.

However, to make things even worse, Luty et al [322] also pointed out a classical instability in the DGP-model, resulting from negative energy solutions. A classical solution on the brane that satisfies the dominant energy condition ($T_{\mu\nu}$ on the brane with $p = w\rho$, satisfying $-1 \leq w \leq 1$) has negative energy for a higher dimensional

observer. The line element is:

$$ds^2 = -f^2(r)dt^2 + \frac{1}{f^2(r)}dr^2 + r^2d\Omega_3^2, \quad (7.53)$$

for one extra dimension, with:

$$f^2(r) = 1 - \frac{R_S^2}{r^2}. \quad (7.54)$$

This is of the form of a five-dimensional black hole metric, where DGP allows the values:

$$0 > R_S^2 > -r_{crit}^2. \quad (7.55)$$

These negative solutions thus arise exactly when the DGP model becomes strongly interacting, which is another indication that the model is out of control at distances smaller than this critical distance. Interestingly, Porrati (The ‘P’ in the DGP-model), even wonders whether these obstacles indicate a no-go for IR modifications of GR [331].

The Schwarzschild solutions in the DGP model are also heavily debated and it is not yet clear what the correct way is to calculate these, and whether they will eventually lead to consistent phenomenological behavior. For a recent study and references, see [332].

In [333, 334] it is argued that the vDVZ discontinuity can be consistently circumvented in (A)dS backgrounds at the classical level. Quantum corrections however, will induce the discontinuity again [335, 336]. Furthermore, note that in supergravity theories there is a similar discontinuity for a spin-3/2 gravitino field coupled to a conserved fermionic source [337] in flat space. Also in this case, it is argued that the discontinuity does not arise in AdS space, in two particular limits where the gravitino mass $m \rightarrow 0$ and $m \rightarrow \sqrt{-\Lambda/3}$ [338].

To summarize, as in the case of massive gravity, the DGP-model appears to deviate too strongly from GR at intermediate distances. The possible appearance of the vDVZ discontinuity is a hot topic in braneworld scenarios, and for a good reason. If indeed the discontinuity cannot be circumvented in these models, this would provide a powerful argument against many such approaches. However, here and with the appearance of singularities in the bulk, the UV-regulation is very important. Since the UV-completion, a full string theory embedding, is not known at present, one might therefore hope that all the drawbacks that we encountered, will be resolved in the final formulation.

7.3.2 Non-Zero Brane Width

The DGP model of section (7.3) has also been criticized on different grounds. If one allows for a non-zero brane width (so-called fat branes) to regularize the graviton propagator, the flat space graviton propagator exhibits new poles, that correspond to very light tachyons with negative residues [339, 340, 341, 342].

This is however, a UV regularization dependent phenomenon, and it is not known at present, whether these poles persist in for example a string theory embedding [302].

7.3.3 Non-local Gravity

From a 4D-perspective, this approach can also be viewed as to make the effective Newton’s constant frequency and wavelength dependent, in such a way that for sources

that are uniform in space and time it is tiny [343]:

$$M_{Pl}^2 (1 + \mathcal{F}(L^2 \nabla^2)) G_{\mu\nu} = T_{\mu\nu}. \quad (7.56)$$

Here $\mathcal{F}(L^2 \nabla^2)$ is a filter function:

$$\begin{aligned} \mathcal{F}(\alpha) &\rightarrow 0 & \text{for } \alpha \gg 1 \\ \mathcal{F}(\alpha) &\gg 1 & \text{for } \alpha \ll 1 \end{aligned} \quad (7.57)$$

L is a distance scale at which deviations from general relativity are to be expected and $\nabla^2 \equiv \nabla_\mu \nabla^\mu$ denotes the covariant d'Alembertian. Thus (7.56) can be viewed as Einstein's equation with $(8\pi G_N^{eff})^{-1} = M_{Pl}^2(1 + \mathcal{F})$. It is argued that for vacuum energy $\mathcal{F}(0)$ is large enough, such that it will barely gravitate:

$$M_P^2 (1 + \mathcal{F}(0)) G_{\mu\nu} = (M_P^2 + \bar{M}^2) G_{\mu\nu}, \quad \text{and} \quad R = -\frac{4\mathcal{E}_4}{M_P^2 + \bar{M}^2}. \quad (7.58)$$

To reproduce the observed acceleration a value \bar{M} is needed $\bar{M} \sim 10^{48}$ GeV for a vacuum energy density of TeV level, and a $\bar{M} \sim 10^{80}$ GeV for \mathcal{E}_4 of Planck mass value, which is about equal to the mass of the universe.

In terms of the graviton propagator, it gets an extra factor $(1 + \mathcal{F}(k^2 L^2))^{-1}$ and therefore goes to zero when $\mathcal{F}(0) \rightarrow \infty$, instead of generating a tadpole.

In the limit $L \rightarrow \infty$ one arrives at:

$$M_{Pl}^2 G_{\mu\nu} - \frac{1}{4} \bar{M}^2 g_{\mu\nu} \bar{R} = T_{\mu\nu}, \quad (7.59)$$

just the zero mode part of $G_{\mu\nu}$, which is proportional to $g_{\mu\nu}$, where

$$\bar{R} \equiv \frac{\int d^4x \sqrt{g} R}{\int d^4x \sqrt{g}} \quad (7.60)$$

\bar{R} thus is the spacetime averaged Ricci curvature, which vanishes for all localized solutions, such as stars, black holes and also for FRW models. For de Sitter space however, $\bar{R} \neq 0$, but a constant and equal to $\bar{R} = R_\infty$, with R_∞ the asymptotic de Sitter curvature.

At the price of losing 4D-locality and causality, the new averaged term is both non-local and acausal, a model is constructed in which a huge vacuum energy does not lead to an unacceptably large curvature. That is at least the idea. The Planck scale is made enormous for Fourier modes with a wavelength larger than a size L . It is however argued that the acausality has no other observable effect. Moreover, it has been claimed that non-locality should be an essential element in any modification of GR in the infrared that intends to solve the cosmological constant problem [298]. The argument is that it takes local, causal physics a time $1/L$ to respond to modifications at scale $L \sim 10^{28}$ cm, and thus in particular to sources which have characteristic wavelength larger than H_0^{-1} , "such as vacuum energy" [43]. Note that in Feynman diagram language, the only contributions to the renormalization of the cosmological constant come from diagrams with infinite wavelength gravitons, $k = 0$, on the external legs.

The non-localities in this case appear in the four dimensional truncation of the $4 + N$ -dimensional theory of section (7.3). There is an infinite number of degrees of freedom

below any non-zero energy scale. Therefore, in order to rewrite the model as an effective four dimensional field theory, and infinite number of degrees of freedom have to be integrated out. This results in the appearance of non-local interactions, despite the fact that the full theory is local.

Another idea based on a model of non-local quantum gravity and field theory due to Moffat [344, 345], also suppresses the coupling of gravity to the vacuum energy density and also leads to a violation of the Weak Equivalence Principle.

7.3.4 A 5D DGP Brane-World Model

To finish our discussion of the DGP-proposal, we will briefly discuss a 5D model for dark energy. This model thus has no intention to solve the (old) cosmological constant problem, and, moreover, it suffers from the strong coupling problems discussed in section (7.3.1). However, it is interesting in the sense that an accelerated expansion can be obtained, although there is no cosmological constant or other spurious form of dark energy. This model has been extensively discussed in the literature. The so-called ‘self-accelerating phase’ the model exhibits, was first noted by Deffayet et al. in [346, 347], and detailed discussions followed in for example [327, 348, 349].

The important result is that the Friedmann equations obtain an additional term:

$$H^2 \pm \frac{H}{r_c} = \frac{8\pi G}{3}\rho, \quad (7.61)$$

with $H = \dot{a}/a$, as usual. The second term on the left is the new DGP generated term, with an ambiguous sign. As we discussed in section (7.3), in the DGP model, gravity looks higher dimensional past some critical cross-over scale, taken to be of order of the current Hubble radius. This is reflected here; at high Hubble scales, the ordinary Friedmann equation is a very good approximation, which is however substantially altered when $H(t)$ approaches r_c .

The upper sign in (7.61) gives a transition between a phase $H^2 \sim \rho$ and a phase $H^2 \sim \rho^2$. The lower sign is more interesting, and this phase is dubbed the ‘self-accelerating’ phase. With gravity becoming weaker at large distances, gravitational attraction can be overcome, leading to an asymptotic de Sitter phase. The suggestion is that this modified gravitational behavior could possibly explain the current accelerated expansion of the universe, without the need of introducing dark energy. Instead of the ‘old’ cosmological constant problem, one could hope to solve the new cosmological constant problem.

In [348, 350] however, details of this cosmological model have been compared with WMAP data. It is concluded that it gives a ‘significantly worse fit to the supernova data and the distance to the last-scattering surface’ than the ordinary Λ CDM model. Most likely therefore, the DGP-model, interesting as it is, does not solve either of the cosmological constant problems. It especially shows how hard it is to consistently modify GR, without destroying its well tested results.

7.4 Ghost Condensation or Gravitational Higgs Mechanism

One could also think of infrared modifications of gravity, as a result of interactions with a ‘ghost condensate’. This also leads, among other things, to a mass for the graviton, see [351].

Assume that a scalar field ϕ develops a time dependent vacuum expectation value, such that:

$$\langle \phi \rangle = M^2 t, \quad \rightarrow \quad \phi = M^2 t + \pi, \quad (7.62)$$

where π is the scalar excitation on this new background. So the ϕ -field is changing with a constant velocity. Assume furthermore that it obeys a shift symmetry $\phi \rightarrow \phi + a$ so that it is derivatively coupled, and that its kinetic term enters with the wrong sign in the Lagrangian:

$$\mathcal{L}_\phi = -\frac{1}{2} \partial^\mu \phi \partial_\mu \phi + \dots \quad (7.63)$$

The consequence of this wrong sign is that the usual background with $\langle \phi \rangle = 0$ is unstable and that after vacuum decay, the resulting background will break Lorentz invariance spontaneously.

The low energy effective action for the π has the form:

$$S \sim \int d^4 x \left[\frac{1}{2} \dot{\pi}^2 - \frac{1}{2M^2} (\nabla^2 \pi)^2 + \dots \right], \quad (7.64)$$

so that the π ’s have a low energy dispersion relation like:

$$\omega^2 \sim \frac{k^4}{M^2} \quad (7.65)$$

instead of the ordinary $\omega^2 \sim k^2$ relation for light excitations. Time-translational invariance is broken, because $\langle \phi \rangle = M^2 t$ and as a consequence there are two types of energy, a ‘particle physics’ and a ‘gravitational’ energy which are not the same. The particle physics energy takes the form:

$$\mathcal{E}_{pp} \sim \frac{1}{2} \dot{\pi}^2 + \frac{(\nabla^2 \pi)^2}{2M^2} + \dots, \quad (7.66)$$

whereas the gravitational energy is:

$$\mathcal{E}_{grav} = T_{00} \sim M^2 \dot{\pi} + \dots \quad (7.67)$$

Although time-translation- and shift-symmetry are broken in the background, a diagonal combination is left unbroken and generates new ‘time’ translations. The Noether charge associated with this unbroken symmetry is the conserved particle physics energy. The energy that couples to gravity is associated with the broken time translation symmetry. Since this energy begins at linear order in $\dot{\pi}$, lumps of π can either gravitate or anti-gravitate, depending on the sign of $\dot{\pi}$! The π thus maximally violate the equivalence principle.

If the standard model fields would couple directly to the condensate there would be a splitting between particle and anti-particle dispersion relations, and a new spin-dependent inverse-square force, mediated by π exchange, which results from the dispersion relation (7.65). In the non-relativistic limit:

$$\Delta \mathcal{L} \sim \frac{1}{F} \mathbf{S} \cdot \nabla \pi, \quad (7.68)$$

where F is some normalization constant. Because of the k^4 dispersion relation, the potential between two sources with spin \mathbf{S}_1 and spin \mathbf{S}_2 , will be proportional to $1/r$:

$$V \sim \frac{M^4}{\tilde{M}^2 F^2} \frac{\mathbf{S}_1 \cdot \mathbf{S}_2 - 3(\mathbf{S}_1 \cdot \hat{r})}{r}, \quad (7.69)$$

when using only static sources, ignoring retardation effects.

Moreover, not only Lorentz invariance, but also CPT is broken if the standard model fields would couple directly to the condensate. With ψ standard model Dirac fields, the leading derivative coupling is of the form:

$$\Delta\mathcal{L} = \sum_{\psi} \frac{c_{\psi}}{F} \bar{\psi} \bar{\sigma}^{\mu} \psi \partial_{\mu} \phi. \quad (7.70)$$

As noted in [351], field redefinitions $\psi \rightarrow e^{ic_{\psi}\phi/F} \psi$ may remove these couplings, but only if such a $U(1)$ symmetry is not broken by mass terms or other couplings in the Lagrangian. If the fermion field ψ has a Dirac mass term $m_D \psi \psi^c$, then the vector couplings, for which $c_{\psi} + c_{\psi^c} = 0$, still can be removed, but the axial couplings remain:

$$\Delta\mathcal{L} \sim \frac{1}{F} \bar{\Psi} \gamma^{\mu} \gamma^5 \Psi \partial_{\mu} \phi. \quad (7.71)$$

After expanding $\phi = M^2 t + \pi$ this becomes:

$$\Delta\mathcal{L} \sim \mu \bar{\Psi} \gamma^0 \gamma^5 \Psi + \frac{1}{F} \bar{\Psi} \gamma^{\mu} \gamma^5 \Psi \partial_{\mu} \pi, \quad (7.72)$$

with $\mu = M^2/F$. This first term violates both Lorentz invariance and CPT, leading to different dispersion relations for particles and their anti-particles. A bound on μ is obtained by considering the earth to be moving with respect to spatially isotropic condensate background. The induced Lorentz and CPT violating mass term then looks like:

$$\mu \bar{\Psi} \vec{\gamma} \gamma^5 \Psi \cdot \mathbf{v}_{earth}, \quad (7.73)$$

which in the non-relativistic limit gives rise to an interaction Hamiltonian:

$$\mu \mathbf{S} \cdot \mathbf{v}_{earth}. \quad (7.74)$$

The experimental limit on μ for coupling to electrons is $\mu \leq 10^{-25}$ GeV [352] assuming $|\mathbf{v}_{earth}| \sim 10^{-3}$. For other limits on CPT and Lorentz invariance, see [353, 354, 355].

If there is no direct coupling, the SM fields would still interact with the ghost sector through gravity. Interestingly, IR modifications of general relativity could be seen at relatively short distances, but only after a certain (long) period of time! Depending on the mass M and the expectation value of ϕ , deviations of Newtonian gravity could be seen at distances 1000 km, but only after a time $t_c \sim H_0^{-1}$ where H_0 is the Hubble constant. More general, the distance scale at which deviations from the Newtonian potential are predicted is $r_c \sim M_{Pl}/M^2$ and their time scale is $t_c \sim M_{Pl}^2/M^3$.

To see the IR modifications to GR explicitly, let us consider the effective gravitational potential felt by a test mass outside a source $\rho_m(r, t) = \delta^3(r) \theta(t)$, i.e. a source that turns on at time $t = 0$. This potential is given by:

$$\Phi(r, t) = -\frac{G}{r} [1 + I(r, t)], \quad (7.75)$$

where $I(r, t)$ is a spatial Fourier integral over momenta k , evaluated using an expansion around flat space; a bare cosmological constant is set to zero.

$$I(r, t) = \frac{2}{\pi} \left\{ \int_0^1 du \frac{\sin(uR)}{(u^3 - u)} \left(1 - \cosh(Tu\sqrt{1 - u^2}) \right) + \int_1^\infty du \frac{\sin(uR)}{(u^3 - u)} \left(1 - \cos(Tu\sqrt{u^2 - 1}) \right) \right\}. \quad (7.76)$$

Here $u = k/m$, $R = mr$, $T = \alpha M^3/2M_{Pl}^2$, where $m \equiv M^2/\sqrt{2}M_{Pl}$ and α is a coefficient of order 1. For late times, $t \gtrsim t_c$, or $T \gtrsim 1$, the first integrand will dominate and $I(r, t)$ can be well approximated by:

$$I(r, t) \simeq \frac{2}{\sqrt{\pi T}} \exp\left(-\frac{R^2}{8T} + \frac{T}{2}\right) \sin\left(\frac{R}{\sqrt{2}}\right). \quad (7.77)$$

For $R \ll T$, there is indeed an oscillatory behavior for the gravitational potential, growing exponentially as $\exp(T/2)$, while for $R \gg T$ the modification vanishes.

More general gravitational effects have been studied in [356], where moving sources were considered, and in [357] where inflation was studied in this context. Moreover, the quantum stability of the condensate was studied in [358].

This highly speculative scenario opens up a new way of looking at the cosmological constant problem, especially because of the distinction between particle physics energy, \mathcal{E}_{pp} and gravitational energy, \mathcal{E}_{grav} . It has to be developed further to obtain a better judgement.

7.5 Fat Gravitons

A proposal involving a sub-millimeter breakdown of the point-particle approximation for gravitons has been put forward by Sundrum [359]. In standard perturbative gravity, diagrams with external gravitons and SM-particles in loops (see figure 7.1) give a con-

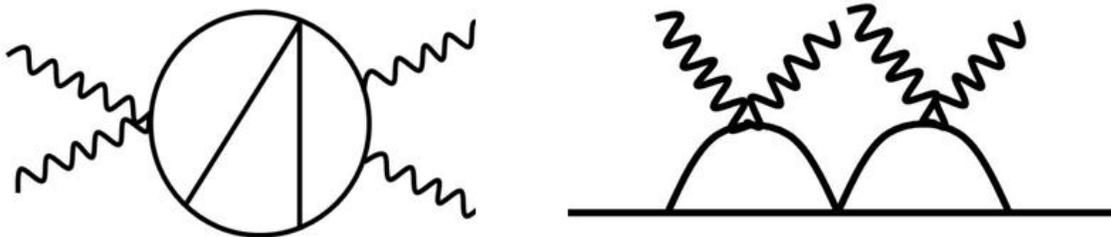


Figure 7.1: On the left-hand-side, a typical Standard Model contribution to $\Gamma_{eff}[g_{\mu\nu}]$. On the right, soft gravitons coupled to loop-correction to SM self-energy. Wiggly lines are gravitons and smooth lines are SM particles.

tribution to the effective CC of which the dominant part diverges as Λ_{UV}^4 where Λ_{UV} is some ultraviolet cutoff. This leads to the enormous discrepancy with experimental results for any reasonable value of Λ_{UV} . However, one might wonder what the risks are when throwing away these diagrams from the effective theory $\Gamma_{eff}[g_{\mu\nu}]$, when $|k^2|$, the momentum of the external gravitons, is larger than some low energy cutoff. Properties

at stake are: Unitarity, General Coordinate Invariance (GCI) and locality. In standard effective theory, diagrams where soft gravitons give corrections to the SM self energy diagrams (the right one in figure 7.1), are crucial in maintaining the equivalence principle between inertial and gravitational masses. However, in a theory with locality, the value associated to the diagram on the left, should follow unambiguously from that of the diagram on the right. Thus given locality of the couplings of the point particles in the diagrams, we cannot throw the first diagram away and keep the other. Therefore, it seems progress can be made by considering a graviton as an extended object. Note that exactly the same arguments hold for the diagrams with just one external graviton line i.e. a normal tadpole diagram, with matter particles running around in the loop, and diagrams with a standard model contribution to the self energy, with and without an external graviton leg. These latter two diagrams are also corrections to (first) the gravitational and (second) inertial mass. Again, the contribution from the loop integral in the tadpole diagram can be suppressed, if there is a loop momentum-dependent form factor to the vertex. This would also suppress the correction to the gravitational mass in the second diagram, but obviously not the third diagram, since it does not contain an external graviton, thus violating the equivalence principle. Only for an extended graviton with some non-zero size, only the first diagram can be suppressed.

Define the graviton size:

$$l_{grav} \equiv \frac{1}{\Lambda_{grav}}. \quad (7.78)$$

Such a “fat graviton” does not have to couple with point-like locality to SM loops, but with locality up to l_{grav} . Thus a fat graviton can distinguish between the two types of diagrams, possibly suppressing the first while retaining the second.

The value of the CC based on usual power counting would then be:

$$\Lambda_{eff} \sim \mathcal{O}(\Lambda_{grav}^4/16\pi^2). \quad (7.79)$$

Comparing with the observational value this gives a bound on the graviton size of:

$$l_{grav} > 20 \text{ microns} \quad (7.80)$$

which would indicate a short-distance modification of Newton’s law below 20 microns. This is however not enough to suppress standard model contributions to the cosmological constant. The same two-point diagrams also renormalize Newton’s constant, sending it to zero; the Planck mass becomes enormous. A new model by the same author has been proposed to take into account also these effects, see section 4.9.1.

But there are other difficulties as well. Purely gravitational loops will still contribute too much to the cosmological constant, and tree level effects, such as a shift in vacuum energy during phase transitions, are not accounted for either. Besides, such a fat graviton would still couple normally to particles with a mass $m < \Lambda_{grav}$, in particular massless particles. And, even more seriously, unitarity is lost when matter couples to gravity above this cutoff at 10^{-3} eV. This loss of unitarity is Planck suppressed, but is nevertheless a severe problem for this approach.

On a more positive note, the idea that gravity shuts off completely below 10^{-3} eV is a very interesting idea. The cosmological constant problem could be solved if one were to find a mechanism showing that flat spacetime is a preferred frame at distances $l < 0.1$ mm. The model of Sundrum is an approach in this direction, and one of very

few models in which gravity becomes weaker at shorter distances. Moreover, another obvious advantage is that it can at least be falsified by submillimeter experiments of the gravitational $1/r^2$ law.

7.6 Composite Graviton as Goldstone boson

Another approach is to consider the possibility that the graviton appears as a composite Goldstone boson. There exists a theorem by Weinberg and Witten, [360], stating that a Lorentz invariant theory, with a Lorentz covariant energy-momentum tensor does not admit a composite graviton. In fact, this theorem poses restrictions on the presence of any massless spin-1 and spin-2 particle. It states that such a massless particle cannot exist if it is charged under the current generated by the conserved vector or conserved symmetric 2nd rank tensor. The photon can be massless, since it is not charged under this current (no electric charge), the gluon can be massless because the current is not Lorentz-covariant due to an additional gauge freedom; the same reason for the ‘normal’ graviton to evade this Weinberg-Witten no-go theorem.

It is therefore natural to try a mechanism where the graviton appears as a Goldstone boson associated with the spontaneous breaking of Lorentz invariance. Being a Goldstone boson, the graviton would not develop a potential, and hence the normal cosmological constant problem is absent, see for example [361, 362]. We will briefly review the latter proposal.

The effective action is written:

$$\mathcal{L} = N (A^2 \sqrt{-g} R - A^4 V(h) + \text{higher derivatives}) + \mathcal{L}_{matter} + \mathcal{O}(N^0), \quad (7.81)$$

where N is some large number $N \sim 10^4$, A is a cutoff, and h is defined by:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \quad (7.82)$$

The observed Newton’s constant is $G_N \sim 1/(NA^2)$ and general covariance is only violated by the potential.

The point now is to consider potentials leading to a vev for $h_{\mu\nu}$. Lorentz transformations can bring $\langle h_{\mu\nu} \rangle$ to diagonal form and when the $h_{\mu\mu}$ are all non-zero and distinct, the Lorentz group will be completely broken. This leads to six Goldstone bosons, the off-diagonal components of $h_{\mu\nu}$. Two of these are identified to form the ‘ordinary’ graviton.

It is clear that in such a setting, one has to rethink the cosmological constant problem, since general covariance is violated. Moreover, since the graviton is identified as a Goldstone boson, it would not develop a potential.

However, so much has been ruined, that indeed the traces of broken Lorentz invariance cannot be erased. Moreover, de Sitter and anti de Sitter spacetime are still allowed solutions [362].

Violations of Lorentz invariance, or perhaps even violations of general covariance, are very interesting routes to explore with regard to the cosmological constant problem [363, 364, 365, 366]. Indeed, if flat spacetime would in some sense be a preferred frame on short distances, say at a tenth of a millimeter, without violating bounds on GR at larger distance scales, this could be considered a solution. In my opinion this could

very well be the underlying reason why we observe such a small cosmological constant. With Minkowski as a preferred frame at these rather large submillimeter distances, one could boldly suggest that in fact gravity does not exist at smaller distances. This would solve the cosmological constant problem. Note that it is phenomenologically opposite to scenarios with extra dimensions where gravity becomes stronger at smaller distances. Such a construction is however far from trivial, because one would have to deal with violations of Lorentz invariance, general covariance, and perhaps locality.

With a preferred frame, a priori Lorentz invariance and general covariance both would be broken, but rotational invariance would still be an exact symmetry. Mathematically, such a construction could be built using a unit future timelike vector field u^a . This generally leads to many possible observable effects, such as modified dispersion relations for matter fields. Another important issue is that general covariance in Einstein's equation ensures stress-energy conservation, which is not to be violated. One therefore often looks for a mechanism where the preferred frame arises, while preserving general covariance. This requires the preferred frame to be dynamical. Moreover, to prevent the theory from instabilities, the preferred frame would have to arise from local conditions [364]. The construction can be setup as a vector-tensor theory of gravitation, but with an additional constraint since the vector field is a purely unit timelike vector field. This has for example the benefit, that ghosts may not appear, even though a vector field is introduced without gauge invariance. Unfortunately however, it turns out to be very non-trivial to construct stable, viable theories of gravity this way. The PPN-parameters give strict constraints and often not only Lorentz invariance is broken, but also C , CP and CPT [363]. Moreover, introducing just any preferred frame leads to a renormalization of Newton's constant, through the 'aether' energy-momentum tensor, but has a priori no direct effect on the cosmological constant. In fact, many models break Lorentz invariance by giving a vacuum expectation value to a vector field, which in a sense only has a negative effect on the cosmological constant problem.

See [367] for a review and [368] for a very broad discussion of alternative theories of gravity including their experimental consequences.

Violations of Lorentz invariance have been studied in both loop quantum gravity [369] and string theory [370]. The missing GZK-cut-off could be an experimental indication in the same direction [371]. For modern tests of violations of Lorentz invariance, see [372].

7.6.1 Summary

Since General Relativity has only been thoroughly tested on solar system distance scales it is a very legitimate idea to consider corrections to GR at galactic and/or cosmological distance scales. However, often these models are not so harmless as supposed to be: changing the laws of gravity also at shorter scales, or leading to violations of locality. The scenarios described in this section do not directly solve the cosmological constant problem, but offer new ways of looking at it.

On the more positive side, many theories that predict modifications of GR in the IR, reproduce Einstein gravity at smaller distances, but up to some small corrections. These corrections are discussed in [373, 374, 350] and could be potentially observable at solar system distance scales. At the linearized level gravity is of the scalar-tensor type,

because the graviton has an extra polarization that also couples to conserved energy-momentum sources. If these models are correct, an anomalous perihelion precession of the planets is expected to be observed in the near future.

Most likely, the DGP-model is not such a candidate. To solve the old cosmological constant problem and arrive at a flat brane metric solution, at least three extra dimensions are necessary. However, with a flat brane metric, singularities in the bulk appear unavoidable, unless there are miraculous cancellations once the full UV completion is known. The model also suffers from hidden strong interaction scales, which lead to modifications of GR at distances smaller than 1000 km, for a cross-over distance to the higher-dimensional behavior at the current Hubble length.

Besides, submillimeter experiments of Newtonian gravity set ever more stringent bounds on both extra dimensional approaches and composite graviton scenarios. It would be very exciting to see a deviation of Newtonian gravity at short distances. Especially a weakening of gravity at these distances would be very welcome with regard to the cosmological constant problem.

On the other hand, observing no change at all, will seriously discourage the hopes that such a mechanism might help in solving the cosmological constant problem. The critical distance for this problem is roughly a tenth of a millimeter and that scale is practically within reach for experimental investigations.

8 | Type IV: Statistical Approaches

In this chapter approaches to the cosmological constant problem are studied, in which quantum effects at the beginning of the universe play a dominant role. The first two sections deal with so-called quantum cosmological scenarios based on the Wheeler-DeWitt equation [151, 375], and the third section describes the anthropic principle. Common in all three approaches is that the value of the cosmological constant is distributed according to some governing equation and that a probability for $\Lambda = 0$ is calculated.

If the cosmological constant could a priori have any value, appearing for example as a constant of integration as in section (3.2.1), or would become a dynamical variable by means of some other mechanism, then in quantum cosmology the state vector of the universe would be a superposition of states with different values of Λ_{eff} . The path integral would include all, or some range of values of this effective cosmological constant. The observable value of the CC in this framework is not a fundamental parameter. Different universes with different values of Λ_{eff} contribute to the path integral. The probability \mathcal{P} of observing a given field configuration will be proportional to $\mathcal{P} \propto \exp(-S(\Lambda_{eff}))$ in which Λ_{eff} is promoted to be a quantum number.

The wave function of the universe [376] would have to satisfy an equation, analogously to the Klein-Gordon equation, called the Wheeler-DeWitt equation. One arrives at this equation, after canonically quantizing the gravitational Hamiltonian. This Hamiltonian formulation is not covariant, but requires a 3+1 split of the metric, called the Arnowitt, Deser, Misner, or (ADM), decomposition of the spacetime metric in terms of a lapse function N , a shift vector N_i and a spatial metric h_{ij} , see [214, 377]. In terms of these quantities, the proper length $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ becomes:

$$ds^2 = -(Ndt)^2 + h_{ij}(N^i dt + dx^i)(N^j dt + dx^j). \quad (8.1)$$

Spacetime is decomposed into three-dimensional hypersurfaces Σ_t ; the phase space consists of 3-metrics h_{ij} and various matter fields ϕ on Σ_t , together with their conjugate momenta.

The Einstein-Hilbert action is assumed to be the correct action also at high energies and, in terms of the lapse function N and shift vector N_i , can be written as:

$$\mathcal{L}(N, N_i, h_{ij}) = -\sqrt{h}N (K^2 - K_{ij}K^{ij} - {}^3R) / 16\pi G, \quad (8.2)$$

here h_{ij} is the induced spatial metric, and K_{ij} the extrinsic curvature and K its trace. K_{ij} is defined as:

$$K_{ij} \equiv \frac{1}{2N} \left(N_{i|k} + N_{k|i} - \frac{\partial h_{ij}}{\partial t} \right), \quad (8.3)$$

where the subscript $|j$ denotes covariant differentiation with respect to the spatial metric h_{ij} . Since the Lagrangian (8.2) contains no time derivatives of N or N_i , their

conjugate momenta vanish:

$$\pi \equiv \frac{\delta \mathcal{L}}{\delta \dot{N}} = 0, \quad \pi^i \equiv \frac{\delta \mathcal{L}}{\delta \dot{N}_i} = 0, \quad (8.4)$$

these are called ‘primary’ constraints. Furthermore, the momentum conjugate to h_{ij} is:

$$\pi^{ij} \equiv \frac{\delta \mathcal{L}}{\delta \dot{h}_{ij}} = \sqrt{h} (K^{ij} - h^{ij} K) / 16\pi G. \quad (8.5)$$

The Hamiltonian then follows in the usual way:

$$\begin{aligned} H &= \int d^3x \left(\pi^{ij} \dot{h}_{ij} + \pi^i \dot{N}_i + \pi \dot{N} - \mathcal{L} \right) \\ &= \int d^3x (N \mathcal{H}_G + N_i \mathcal{H}^i). \end{aligned} \quad (8.6)$$

\mathcal{H}_G is defined as:

$$\begin{aligned} \mathcal{H}_G &\equiv \sqrt{h} (K_{ij} K^{ij} - K^2 - {}^3R) / 16\pi G \\ &= \frac{16\pi G}{2\sqrt{h}} (h_{ik} h_{jl} + h_{il} h_{jk} - h_{ij} h_{kl}) \pi^{ij} \pi^{kl} - \sqrt{h} {}^3R / 16\pi G \\ &\equiv 16\pi G G_{ijkl} \pi^{ij} \pi^{kl} - \sqrt{h} {}^3R / 16\pi G \end{aligned} \quad (8.7)$$

For a closed FRW-model, these are related as follows:

$$R = K^2 - K_{ij} K^{ij} - {}^3R \quad (8.8)$$

Furthermore,

$$\mathcal{H}^i \equiv -2\pi^{ij} / 16\pi G. \quad (8.9)$$

\mathcal{H}_G and \mathcal{H}^i have to satisfy the ‘secondary’ constraint $\mathcal{H}_G = \mathcal{H}^i = 0$, which for \mathcal{H}_G can be written as:

$$\mathcal{H}_G = 16\pi G G_{ijkl} \pi^{ij} \pi^{kl} - \sqrt{h} {}^3R / 16\pi G = 0 \quad (8.10)$$

with the definition of G_{ijkl} implicitly given in (8.7).

So far this is just the classical treatment. Quantization is obtained by first identifying the Hamiltonian constraint as the zero energy Schrodinger equation:

$$\mathcal{H}_G(\pi_{ij}, h_{ij}) \Psi[h_{ij}] = 0 \quad (8.11)$$

where the state vector Ψ is the wave function of the universe, and secondly by replacing the momenta π in the normal canonical quantization procedure by:

$$\pi^{ij} \rightarrow -i \left(\frac{1}{16\pi G} \right)^{3/2} \frac{\delta}{\delta h_{ij}}. \quad (8.12)$$

This gives the Wheeler-DeWitt equation:

$$\left[\frac{G_{ijkl}}{(16\pi G)^2} \frac{\delta}{\delta h_{ij}} \frac{\delta}{\delta h_{kl}} + \frac{\sqrt{h}({}^3R - 2\Lambda)}{16\pi G} \right] \Psi[h_{ij}] = 0 \quad (8.13)$$

One of the most difficult features of this equation is that the wave function of the universe is independent of time. Ψ depends only on h_{ij} and the matter content. This makes the interpretation of this wave function very complicated.

Probabilities are calculated, using the fact that the WKB approximation for tunneling is proportional to $\exp(-S_E)$, with S_E the Euclidean action. In this context, it gives the amplitude for the entire universe to tunnel from an in-state, to an out-state. The WKB approximation is justified, since we will consider de Sitter spaces, which are large relative to the Planck scale. It is assumed that in this case, short distance quantum gravity effects can be neglected.

The appropriate boundary condition is a matter of debate. Hartle and Hawking advocate the ‘no-boundary’ boundary condition, which amounts to take:

$$\Psi_0[h_{ij}] \propto \int \mathcal{D}g \exp[-S_E] \quad (8.14)$$

where the path integral over geometries extends over all compact Euclidean 4-geometries, which have no boundary. $S_E(g)$ is the Euclidean action associated with the manifold.

The ‘quantum cosmology’ one obtains this way faces many unresolved issues. We have, for example, used a Euclidean action throughout, which is clearly unsuitable for gravity, considering that the Euclidean gravitational action is unbounded from below. Besides, when using a sum-over-histories approach, it is assumed that the universe is finite and closed, since the relevant integrals are undefined in an open universe. Next, the spatial topology of the universe is assumed fixed (note that the spatial topology is undetermined in general relativity). Moreover, there is an issue in interpretation. What does it mean to have a wave function of the universe? And, last but not least, the role of time, is not exactly understood. Curiously, most of these “difficulties” do not show up in the perturbative treatment of quantum gravity.

8.1 Hawking Statistics

Eleven dimensional supergravity contains a three-form gauge field, with a four-form field strength $F_{\mu\nu\rho\sigma} = \partial_{[\mu}A_{\nu\rho\sigma]}$ [378]. When reduced to four dimensions, this gives a contribution to the cosmological constant [379, 380, 381, 382, 383]. Hawking [384] used such a three-form gauge field to argue that the wave function of the universe is peaked at zero cosmological constant. It is the first appearance of the idea that the CC could be fixed by the shape of the wave function of the universe.

The three-form field $A_{\mu\nu\lambda}$ is subject to gauge transformations:

$$A_{\mu\nu\rho} \rightarrow A_{\mu\nu\rho} + \nabla_{[\mu}C_{\nu\rho]}, \quad (8.15)$$

which leaves invariant the fields $F_{\mu\nu\rho\sigma} = \nabla_{[\mu}A_{\nu\rho\sigma]}$. This field would contribute an extra term to the action:

$$I = -\frac{1}{16\pi G} \int d^4x \sqrt{-g} (R + 2\Lambda_B) - \frac{1}{48} \int d^4x \sqrt{-g} F_{\mu\nu\rho\sigma} F^{\mu\nu\rho\sigma}. \quad (8.16)$$

The field equation for $F^{\mu\nu\rho\sigma}$ is:

$$D_\mu F^{\mu\nu\rho\sigma} = 0 \quad (8.17)$$

which in four dimensions is just a constant:

$$\sqrt{-g}F^{\mu\nu\rho\sigma} = \omega\epsilon^{\mu\nu\rho\sigma} \quad (8.18)$$

Such a field F has no dynamics, but the F^2 term in the action behaves like an effective cosmological constant term, whose value is determined by the unknown parameter ω , which takes on some arbitrary spacetime independent value. If we substitute the solution (8.18) back into the Einstein equation, we find:

$$T^{\mu\nu} = \frac{1}{6} \left(F^{\mu\alpha\beta\gamma} F_{\alpha\beta\gamma}^{\nu} - \frac{1}{8} g^{\mu\nu} F^{\alpha\beta\gamma\delta} F_{\alpha\beta\gamma\delta} \right) = \pm \frac{1}{2} \omega^2 g^{\mu\nu} \quad (8.19)$$

using that $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\rho\sigma} = \pm 4!$, where the sign depends on the metric used: in Euclidean metric $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\rho\sigma}$ is positive, whereas in Lorentzian metric it is negative. In the Euclidean action Hawking used:

$$R = -4\Lambda_{eff} = -4(\Lambda_B - 8\pi G\omega^2) \quad (8.20)$$

where Λ_B is the bare cosmological constant in Einstein's equation. It follows that:

$$S_{Hawking} = -\Lambda_{eff} \frac{V}{8\pi G}. \quad (8.21)$$

with $V = \int d^4x \sqrt{g}$ is the spacetime volume. The maximum value of this action is given when V is at its maximum, which Hawking takes to be Euclidean de Sitter space; this is just S^4 , with radius $r = (3\Lambda_{eff}^{-1})^{1/2}$ and proper circumference $2\pi r$. We have:

$$V = \frac{24\pi^2}{\Lambda_{eff}^2}, \quad \rightarrow \quad S(\Lambda) = -3\pi \frac{M_P^2}{\Lambda_{eff}} \quad (8.22)$$

and thus the probability density, which is proportional to $\exp(-S_E)$, where S_E is the Euclidean action, becomes:

$$\mathcal{P} \propto \exp \left(3\pi \frac{M_P^2}{\Lambda_{eff}} \right) \quad (8.23)$$

is peaked at $\Lambda_{eff} = 0$.

Note that we have used here that the probability is evaluated as the exponential of minus the effective action at its stationary point. The action is stationary with respect to variations in $A_{\mu\nu\lambda}$, when the covariant derivative of $F_{\mu\nu\lambda\rho}$ vanishes and stationary in $g_{\mu\nu}$ when the Einstein equations are satisfied. Eqn. (8.21) is the effective action at the stationary point. It is a good thing that we only need the effective action at its stationary point, so that we do not have to worry about the Euclidean action not being bounded from below, see for example [40].

However, Hawking's argument has been criticized, since one should not plug an ansatz for a solution back into the action, but rather vary the unconstrained action [385]. This differs a minus sign in this case, the same minus sign as going from a Lorentzian to a Euclidean metric, $\Lambda_{eff} = (\Lambda_B \pm 8\pi G\omega^2)$, but now between the coefficient of $g^{\mu\nu}$ in the Einstein equations, and the coefficient of $(8\pi G)^{-1}\sqrt{g}$ in the action. Note that a plus sign of Λ in the gravitational action, always leads to a minus sign in the Einstein field equations. The correct action becomes [385]:

$$S = (-3\Lambda_{eff} + 2\Lambda_B) \frac{-3\pi M_P^2}{\Lambda_{eff}^2} = -3\pi M_P^2 \frac{\Lambda_B - 12\pi G\omega^2}{(\Lambda_B - 4\pi G\omega^2)^2} \quad (8.24)$$

now for $\Lambda_{eff} \rightarrow 0$, the action becomes large and positive and consequently, $\Lambda_{eff} = 0$ becomes the *least* probable configuration.

Besides, in [386] it is shown that this approach has also other serious limitations. It is argued that it can only work in the ‘Landscape’ scenario that we discuss in section (8.2). The reason is that the four-form flux should be subject to Dirac quantization and the spacing in Λ then only becomes small enough with an enormous number of vacua.

8.1.1 Wormholes

In a related approach Coleman [387] argued that one did not need to introduce a 3-form gauge field, if one includes the topological effects of wormholes. This also transforms the cosmological constant into a dynamical variable. The argument assumes that on extremely small scales our universe is in contact, through wormholes, with other universes, otherwise disconnected, but governed by the same physics as ours. In addition, there are wormholes that connect our universe with itself. Both types of wormholes are assumed be very tiny, but their end points will be at different locations in the universe, and as such, can be arbitrarily far apart, connecting regions that may be causally disconnected. However, at scales larger than the wormhole size, the only effect of wormholes is to add *local* interactions, one for each type of wormhole.

The extra term in the action has the form:

$$S_{wormhole} = \sum_i (a_i + a_i^\dagger) \int d^4x \sqrt{g} e^{-S_i} K_i \quad (8.25)$$

where a_i and a_i^\dagger are the annihilation and creation operators¹ for a type i baby universe, S_i is the action of a semi-wormhole (one that terminates on a baby universe), and K_i is an infinite series of local operators, with operators of higher dimension suppressed by the wormhole size. The interaction therefore, is local at distance scales larger than the wormhole size, but non-local on the wormhole scale. Furthermore, there is an important exponential factor that suppresses the effects of all wormholes, except those of Planckian size [388, 389, 390]. This result is obtained by treating the wormholes semi-classically and in the dilute gas approximation. This dilute gas approximation, provides a way of writing the functional integral over manifolds full of baby universes and wormholes, in terms of an integral over manifolds stripped of these.

The coefficients of these interaction terms are operators $A_i = a_i + a_i^\dagger$ which only act on the variables describing the baby universes, and commute with everything else. The path integral over all 4-manifolds with given boundary conditions becomes:

$$\int [dg][d\Phi] e^{-S} = \int_{\text{No}} [dg][d\Phi] \langle B | e^{-(S+S_{wormhole})} | B \rangle \quad (8.26)$$

where No means that wormholes and baby-universe are excluded, and $|B\rangle$ is a normalized baby-universe state. This state $|B\rangle$ can always be expanded in eigenstates of the

¹We would prefer to talk about functions of fields, since we are doing path integral calculations, and because there is no clear definition of a Hilbert space, on which a and a^\dagger act. However, we will use the conventions of ref. [387].

operators $A_i = a_i + a_i^\dagger$:

$$|B\rangle = \int f_B(\alpha) \prod_i d\alpha_i |\alpha\rangle, \quad (8.27)$$

with α_i the eigenvalues of A_i :

$$(a_i + a_i^\dagger)|\alpha\rangle = \alpha_i|\alpha\rangle, \quad \text{and} \quad \langle\alpha'|\alpha\rangle = \prod_i \delta(\alpha'_i - \alpha_i), \quad (8.28)$$

and the function $f_B(\alpha)$ depends on the boundary conditions. For Hartle-Hawking boundary conditions:

$$a_i|B\rangle = 0 \quad \text{and} \quad f_B(\alpha) = \prod_i \pi^{-1/4} \exp(-\alpha_i^2/2) \quad (8.29)$$

Written in terms of A -eigenstates, the effective action becomes:

$$S_{wormhole} = \sum_i \int d^4x \sqrt{g} \alpha_i e^{-S_i} K_i. \quad (8.30)$$

According to the Copenhagen interpretation, after performing a measurement, the state vector of the universe collapses to an incoherent superposition of these $|\alpha\rangle$'s, each appearing with probability $|f_B(\alpha)|^2$. The α_i renormalize all local operators when measured at distance scales larger than the wormhole scale, i.e. for an observer who cannot detect the baby universes.

This way, the effective cosmological constant also becomes a function of the α_i , since one of these local operators is \sqrt{g} . Moreover, on scales larger than the wormhole scale, manifolds that appear disconnected will really be connected by wormholes, and therefore are to be integrated over. The Hartle-Hawking wave function of the universe Ψ_α^{HH} in the presence of wormholes can be calculated, using the fact that the no-boundary condition now also states that there are no baby universes. In terms of the α 's this reads:

$$\langle\alpha|0\rangle = e^{-\alpha^2/2}, \quad (8.31)$$

The wave function can then be written as:

$$\Psi_\alpha^{HH}(B, \alpha) = \sum e^{-S_{eff}(\alpha)} = e^{-\alpha^2/2} \psi_\alpha^{HH}(B) Z(\alpha), \quad (8.32)$$

where the sum is over all manifolds that go from no-boundary to B and $S_{eff}(\alpha) = S + S_{wormhole}(\alpha)$. This sum factorizes, since some manifolds have components that connect to B , while other components are closed, and have no boundary at all. ψ_α is therefore given by the sum over manifolds connected to B and:

$$Z(\alpha) = \sum_{CM} e^{-S_{eff}(\alpha)}, \quad (8.33)$$

a sum over closed manifolds and this $Z(\alpha)$ can be interpreted as giving the probability of finding a given value of α in the Hartle-Hawking state. This expectation value can be calculated from:

$$\langle\psi\rangle^{HH} = \frac{\int d\alpha e^{-\alpha^2/2} \langle\psi\rangle_\alpha^{HH} Z(\alpha)}{\int d\alpha e^{-\alpha^2/2} Z(\alpha)} \quad (8.34)$$

Thus we find:

$$d\mathcal{P} = e^{-\alpha^2/2} Z(\alpha) d\alpha, \quad (8.35)$$

with $d\mathcal{P}$ the probability distribution.

The sum of all vacuum-to-vacuum graphs is the exponential of the sum of connected graphs, which gives:

$$\mathcal{P} \propto \exp \left[\sum_{CCM} e^{-S_{eff}(\alpha)} \right], \quad (8.36)$$

where CCM stands for closed connected manifolds. The sum can be expressed as a background gravitational field effective action, Γ . The sum over closed connected manifolds can then be written as a sum over topologies:

$$\sum_{CCM} e^{-S_{eff}(\alpha)} = \sum_{topologies} e^{-\Gamma(g)}, \quad (8.37)$$

with g the background metric on each topology and each term on the right is again to be evaluated at its stationary point. This is progress, since the leading term in Γ for large, smooth universes is known, and is the cosmological constant term:

$$\Gamma = \Lambda(\alpha) \int d^4x \sqrt{g} + \dots, \quad (8.38)$$

$\Lambda(\alpha)$ being the fully renormalized cosmological constant. Plugging this back into (8.36) gives the final result:

$$\mathcal{P} \propto \exp \left[\exp \left(3\pi \frac{M_P^2}{\Lambda_{eff}} \right) \right], \quad (8.39)$$

and thus is even sharper peaked at $\Lambda = 0$ than in Hawking's case. For positive CC the maximum volume is taken, like in Hawking's case, the 4-sphere with $r = (3\Lambda_{eff}^{-1})^{1/2}$. Furthermore, on dimensional grounds, the higher order terms in (8.38) are neglected.

An advantage of Coleman's approach is that he is able to sidestep many technical difficulties Hawking's approach suffers from. In particular, he uses the Euclidean path integral only to calculate expectation values of some scalar field. These are independent of x , because the theory is generally covariant. It includes an average over the time in the history of the universe that the expectation value for this operator was measured. This circumvents many issues related to the notion of time in quantum gravity.

However, both Hawking's and Coleman's proposal rely strongly on using a Euclidean path integral and since it is ill-defined, it is unclear whether this is suitable for a theory of quantum gravity².

There also appears to be a more direct problem with Coleman's idea, as put forward by Fishler, Susskind and Polchinski [392, 393], also see [394, 395]. The problem is that in Coleman's scenario wormholes of every size will materialize in the vacuum with maximum kinematically allowed density, leading to a universe packed with wormholes of every size. The exponential suppression factor in (8.25) is inconsistent with the other assumptions that quantum gravity is described by a Euclidean path integral, which is dominated by large scale spherical universes connected by wormholes, where

²In [391] this is made more concrete: "Evidently, the Euclidean path integral is so ill-defined that it can be imaginatively used to prove anything."

the amplitude of a large scale universe is of order $\exp(M_P^2/\Lambda)$. In particular, taking into account the higher order terms in (8.38), leads to a violation of the dilute gas approximation, used by Coleman.

This can be seen as follows. The effect of wormholes is to renormalize couplings, so the Einstein action is written:

$$S(a) = \int d^4x \sqrt{g} (\Lambda(a) + \kappa^2(a)R + \gamma(a)R^2 + \dots) \quad (8.40)$$

where a is the size of the wormholes $\kappa^2 = 1/(8\pi G)$ and in the above action fluctuations up to distances a are integrated out. Since wormhole actions are $\propto a^2/\kappa^2$, the relative amplitudes for wormholes of sizes a and a' is:

$$\frac{\exp(-a'^2/\kappa^2)}{\exp(-a^2/\kappa^2)}, \quad (8.41)$$

so large wormholes are suppressed. Keeping also the higher order term $\propto \gamma(a)R^2$, the probability distribution $\mathcal{P}(\alpha)$ (8.39) can be written:

$$\mathcal{P}(\alpha) \propto \exp \left[\exp \left(\frac{1}{[\Lambda(a) + \alpha_1][\kappa^2(a) + \alpha_2]} + [\gamma(a) + \alpha_3] \right) \right]. \quad (8.42)$$

This shows that the probability is enhanced not only for $\Lambda(a) + \alpha_1 \rightarrow 0$, but also for $\alpha_3 \rightarrow \infty$. However, $\gamma(a)$ cannot be too large, or else unitarity would be lost and it has been assumed that it would reach its maximum value, consistent with unitarity [40]. In [392] it is shown that this requirement is not consistent with the assumption of dominance of only small wormholes. On the contrary, wormholes of all scales a will play a dominant role and strongly affect even macroscopic physics.

In conclusion, wormholes should be integrated out of the functional integral of quantum gravity. Their effect is to renormalize the values of physical constants in our universe. After integrating out the wormholes of all sizes, one should be left with a local theory. If, for some reason, it is valid to only take Planck-scale wormholes into account, this could make the wavefunction of the universe in the Euclidean formalism, peak at zero value of the cosmological constant. The next non-trivial question to answer is what the physical implications of this would be, since the formalism of a wavefunction of the universe in a Euclidean spacetime is, to say the least, not very well defined.

8.2 Anthropic Principle

The anthropic principle is a way of reasoning to better understand the circumstances of our universe. There are two different versions. The first corresponds to the trivial or weak version, which is just a tautology: intelligent observers will only experience conditions which allow for the existence of intelligent observers.

Proponents of the Strong Anthropic Principle advocate the stronger point that the physical constants and physical laws have the values they have exactly to make intelligent life possible. Most physicists and cosmologists reject this latter form. In fact it dates back to the question whether the universe has a goal or not and leads to old philosophical discussions concerning teleology. Basically the whole discussion is turned

into a debate on cause and consequence. Moreover, it is sometimes argued that the laws of nature, which are otherwise incomplete, are completed by the requirement that conditions must allow intelligent life to arise, the reason being that science (and quantum mechanics in particular) is meaningless without observers.

In the remainder of this section we will only discuss the weak version. In order for the tautology to be meaningful it is necessary that there are alternative conditions where things are different. Therefore, it is usually assumed that there is some process that produces an ensemble of a large number of universes, or different, isolated pockets of the same universe, with widely varying properties. Several inflationary scenarios [396, 397, 398, 399], quantum cosmologies, [384, 400, 401, 402, 403] and string theory [386, 404, 405, 164, 406, 407] predict different domains of the universe, or even different universes, with widely varying values for the different coupling constants. In these considerations it is assumed that there exist many discrete vacua with densely spaced vacuum energies.

Since the conditions for life to evolve as we know it are very constrained, one can use a form of the anthropic principle to select a certain state, with the right value for the cosmological constant, the fine structure constant, etc, from a huge ensemble.

Depending on how specific the conditions for intelligent life to form are, we can expect to find more or less bizarre looking situations of extreme fine-tuning. For example, the only reason that heavier elements are formed in the absence of stable elements with atomic weights $A = 5$ or $A = 8$, is that the process in which three ${}^4\text{He}$ nuclei build up to form ${}^{12}\text{C}$ is resonant, there is an excited energy level for the carbon nucleus that matches the typical energies of three alpha-particles in a star. Moreover, if the Higgs vev decreases by a factor of a few, the proton becomes heavier than the neutron, and hydrogen decays. If the vev increases by a factor of a few, nuclei heavier than hydrogen decay, because the neutron-proton mass difference becomes larger than the nuclear binding energy per nucleon. Insisting that carbon should form, gives an even better determination of the Higgs vev [408, 409].

In this form, just setting boundary conditions to values of physical constants, the ‘principle’³ can be useful: a theory that predicts a too rapidly decaying proton for example cannot be right, since otherwise we would not survive the ionizing particles produced by proton decay in our own bodies. Now although no one would disagree with this, it is also not very helpful: it does not explain why the proton lives so long and better experimental limits can be found using different methods. See [410] for a very general use of the anthropic principle.

8.2.1 Anthropic Prediction for the Cosmological Constant

One of the first to use anthropic arguments related to the value of the cosmological constant was Weinberg [411], see also [412, 413]. He even made the prediction in 1987 that, since the anthropic bound is just a few orders of magnitude larger than the experimental bounds, a non-zero cosmological constant would soon be discovered, which indeed happened.

³In Kolb & Turner [65] a footnote is written in which one can see a glimpse of the conflicting opinions about this approach: “It is unclear to one of the authors how a concept as lame as the “anthropic idea” was ever elevated to the status of a Principle”.

One can rather easily set anthropic bounds on the value of the cosmological constant. A large positive CC would lead very early in the evolution of the universe to an exponentially expanding de Sitter phase, which then lasts forever. If this would happen before the time of formation of galaxies, at redshift $z \sim 4$, clumps of matter would not become gravitationally bound, and galaxies, and presumably intelligent life, would not form. Therefore:

$$\Omega_\Lambda(z_{gal}) \leq \Omega_M(z_{gal}) \quad \rightarrow \quad \frac{\Omega_{\Lambda 0}}{\Omega_{M 0}} \leq a_{gal}^3 = (1 + z_{gal})^3 \sim 125. \quad (8.43)$$

This implies that the cosmological constant could have been larger than observed and still not be in conflict with galaxy formation⁴. On the other hand, a large negative cosmological constant would lead to a rapid collapse of the universe and (perhaps) a big crunch. To set this lower anthropic bound, one has to wonder how long it takes for the emergence of intelligent life. If 7 billion years is sufficient, the bound for a flat universe is $\Lambda \gtrsim -18.8 \rho_0 \sim -2 \times 10^{-28} \text{ g/cm}^3$, if 14 billion years are needed, the constraint is $\Lambda \gtrsim -4.7 \rho_0 \sim -5 \times 10^{-29} \text{ g/cm}^3$ [414].

It makes more sense however, to ask what the most likely value of the cosmological constant is, the value that would be experienced by the largest number of observers. Vilenkin's "Principle of Mediocrity" [402], stating that we should expect to find ourselves in a big bang that is typical of those in which intelligent life is possible, is often used. The probability measure for observing a value ρ_Λ , using Bayesian statistics, can be written as:

$$d\mathcal{P}(\rho_\Lambda) = N(\rho_\Lambda) \mathcal{P}_*(\rho_\Lambda) d\rho_\Lambda, \quad (8.44)$$

where $\mathcal{P}_*(\rho_\Lambda) d\rho_\Lambda$ is the a priori probability of a particular big bang having vacuum energy density between ρ_Λ and $\rho_\Lambda + d\rho_\Lambda$ and is proportional to the volume of those parts of the universe where ρ_Λ takes values in the interval $d\rho_\Lambda$. $N(\rho_\Lambda)$ is the average number of galaxies that form at a specified ρ_Λ [53], or, the average number of scientific civilizations in big bangs with energy density ρ_Λ [42], per unit volume. The quantity $N(\rho_\Lambda)$ is often assumed to be proportional to the number of baryons, that end up in galaxies.

Given a particle physics model which allows ρ_Λ to vary, and a model of inflation, one can in principle calculate $\mathcal{P}_*(\rho_\Lambda)$, see the above references for specific models and [415] for more general arguments. $\mathcal{P}_*(\rho_\Lambda) d\rho_\Lambda$ is sometimes argued to be constant [416], since $N(\rho_\Lambda)$ is only non-zero for a narrow range of values of ρ_Λ . Others point out that there may be a significant departure from a constant distribution [417]. Its value is fixed by the requirement that the total probability should be one:

$$d\mathcal{P}(\rho_\Lambda) = \frac{N(\rho_\Lambda) d\rho_\Lambda}{\int N(\rho'_\Lambda) d\rho'_\Lambda}. \quad (8.45)$$

The number $N(\rho_\Lambda)$ is usually calculated using the so-called 'spherical infall' model of Gunn and Gott [418]. Assuming a constant $\mathcal{P}_*(\rho_\Lambda)$, it is argued that the probability of a big bang with $\Omega_\Lambda \lesssim 0.7$ is roughly 10%, depending on some assumptions about the density of baryons at recombination [42, 419].

If $\tilde{\rho}_\Lambda$ is the value for which the vacuum energy density dominates at about the epoch of galaxy formation, then values $\rho_\Lambda \gg \tilde{\rho}_\Lambda$ will be rarely observed, because the density

⁴Note that in these estimates everything is held fixed, except for Ω_Λ which is allowed to vary, unless stated otherwise.

of galaxies in those universes will be very low. Values $\rho_\Lambda \ll \tilde{\rho}_\Lambda$ are also rather unlikely, because this range of values is rather small. A typical observer therefore would measure $\rho_\Lambda \sim \tilde{\rho}_\Lambda$, which is the anthropic prediction and it peaks at $\Omega_\Lambda \sim 0.9$, in agreement with the experimental value $\Omega_\Lambda \sim 0.7$ at the 2σ level [420]. It is argued that the agreement can be increased to the 1σ level, by allowing for non-zero neutrino masses [421]. Neutrino masses would slow down the growth of density fluctuations, and hence influence the value of $\tilde{\rho}_\Lambda$. The sum of the neutrino masses would have to be $m_\nu \sim 1 - 2$ eV.

However, it has been claimed that these successful predictions would not hold, when other parameters, such as the amplitude of primordial density fluctuations are also allowed to vary [422, 423]. These arguments are widely debated and no consensus has been reached [424, 425].

However, it has been very difficult to calculate the a priori distribution. The dynamics, leading to a “multiverse” in which there are different pocket universes with different values for the constants of nature, is claimed to be well understood, for example in case of eternal inflation [426, 397, 399], but the problem is that the volume of these thermalized regions with any given value of the constants is infinite. Therefore, to compare them, one has to introduce some cutoff and the results tend to be highly sensitive to the choice of cutoff procedure [427, 428, 429]. In a recent paper a different method is proposed to find this distribution [430].

It should be stressed that this approach to the cosmological constant problem is especially used within string theory, where one has stumbled upon a wide variety of possible vacuum states, rather than a unique one [386, 404, 405, 164, 406, 407, 431, 432]. By taking different combinations of extra-dimensional geometries, brane configurations, and gauge field fluxes, a wide variety of states can be constructed, with different local values of physical constants, such as the cosmological constant. These are the 3-form RR and NS fluxes that can be distributed over the 3-cycles of the Calabi Yau manifold. The number of independent fluxes therefore is related to the number of 3-cycles in the 6-dimensional Calabi Yau space, and can be several hundred. In addition, the moduli are also numerous and also in the hundreds, leading to a total number of degrees of freedom in a Calabi Yau compactification of order 1,000 or more. The number of metastable vacua for a given Calabi Yau compactification therefore could be 10^{1000} , and the spacing between the energy levels $10^{-1000} M_P^4$, of which some 10^{500} would have a vacuum energy that is anthropically allowed. The states with (nearly) vanishing vacuum energy tend to be those where one begins with a supersymmetric state with a negative vacuum energy, to which supersymmetry breaking adds just the right amount of positive vacuum energy. This picture is often referred to as the “Landscape”. The spectrum of ρ_Λ could be very dense in this ‘discretuum’ of vacua, but nearby values of ρ_Λ could correspond to very different values of string parameters. The prior distribution would then no longer be flat, and it is unclear how it should be calculated.

A review of failed attempts to apply anthropic reasoning to models with varying cosmological constant can be found in [433]. See [434] for a recent critique. Another serious criticism was given in [435], where it is argued that universes very different from our own could also lead to a small cosmological constant, long-lived stars, planets and chemistry based life, for example a cold big bang scenario. An analysis of how to make an anthropic prediction is made in [436].

Not very technical and almost foundational introductions to the anthropic principle are for example [181, 419].

8.2.2 Discrete vs. Continuous Anthropic Principle

It might be worthwhile to make a distinction between a continuous anthropic principle and a discrete version. Imagine we have a theory at our hands that describes an ensemble of universes (different possible vacuum solutions) with different discrete values for the fine structure constant:

$$\frac{1}{\alpha} = n + \mathcal{O}\left(\frac{1}{n}\right) \quad (8.46)$$

such that the terms $1/n$ are calculable. An anthropic argument could then be used to explain why we are in the universe with $n = 137$. Such a version of the anthropic principle might be easier to accept than one where all digits are supposed to be anthropically determined. Note that we are already very familiar with such use of an anthropic principle: In a finite universe, there is a finite number of planets and we live on one of the (very few?) inhabitable ones. Unfortunately, we have no theory at our hands to determine the fine structure constant this way, let alone the cosmological constant.

8.3 Summary

The statistical ideas put forward by Hawking and Coleman turned out to have serious shortcomings. Unfortunately, no consistent model evolved from their pioneering work, that could circumvent these problems.

As we have seen, this is the case with many (if not all) of the proposed solutions to the cosmological constant problem. It is therefore understandable that nowadays the majority of researchers in this field places their bets on the anthropic principle as a solution. These anthropic ideas however especially appear to highlight the problem, instead of giving an explanation.

Anthropic reasoning necessarily requires an ensemble of objects or situations in order to be meaningful. This works very well, when applied, for example, to our planet and its distance to the sun. There are many more planets, but we happen to live on one at the right distance to have a temperature that makes life as we know it possible. In general, if one starts to wonder about the size of the earth, the sun, the solar system or the galaxy, there are two ways to proceed. One is to look for a fundamental physical reason why the diameter of the earth has its particular value and not some value, say, a little bigger. The other is to realize (or assume, in case of an Old Greek scientist) that there are more objects like the earth and try to say something about the properties of these objects in general. In the history of cosmology this has turned out to be very fruitful and by now we see that the earth, our solar system and our galaxy are by no means special. The ultimate shift from the particular to the general would be made by considering multiple universes. This is the starting point of modern anthropic arguments based on some kind of mechanism. However, as a matter of principle we can only do experiments, and therefore statistics based on one universe. Other pockets

of our universe, or other universes where the cosmological constant takes on a different value, will never be accessible to experiment. Therefore, it seems very legitimate to ask whether such an ‘explanation’ can ever be falsified, let alone verified.

Besides, as we have discussed, the bounds obtained on the value of the effective cosmological constant from applying anthropic reasoning, are not very restrictive. The probability of finding oneself in a universe where $\Omega_\Lambda \simeq 0.7$ is only about 10%.

This very much debated approach offers a new line of thought, but so far, unfortunately, predictions for different constants of Nature, like the cosmological constant and the fine-structure constant, are not interrelated. We continue to search for a more satisfactory explanation.

9 | Conclusions

In this thesis we critically discussed different approaches to the cosmological constant problem. The many ways in which the problem can be phrased often blur the road to a possible solution and the wide variety of approaches makes it difficult to distinguish real progress.

So far we can only conclude that in fact none of the approaches described above is a real outstanding candidate for a solution of the ‘old’ cosmological constant problem. The most elegant solution would be a symmetry, that protects the cosmological constant. All possible candidates we can think of were treated in chapter 3 and (4). However, no symmetry, consistent with established results, was found. The symmetry analytically continuing quantum field theory and general relativity to the full complex space (chapter (4)) is interesting, but as it stands, not yet sufficient.

Back-reaction effects, as studied in chapter four are typically very weak, since they depend on quantum gravitational interactions. The model proposed by Tsamis and Woodard, chapter 6, appeared promising at first, but a closer study has revealed major obstacles.

To power their influence, almost all back-reaction approaches need an inflationary background, with a very large bare cosmological constant, which could take on its ‘natural’ value $\sim M_P^2$. Consequently, an enormously large number of e-folds is needed to build up any significant effect. This is another reflection of the fact that there are roughly 120 orders of magnitude between M_P^4 and the observed value of the vacuum energy density which have to be accounted for. Since there exists no convincing upper bound on the number of e-folds (our universe would just be enormously bigger than in usual inflationary scenarios) as discussed in chapter 4, they cannot be ruled out completely. However, we have seen that each one of them at least so far, suffers from serious drawbacks.

Perhaps more promising are approaches that suggest a modification of general relativity, in such a way that the graviton no longer universally couples to all sources. The infinite volume DGP-model of section 7.3, appeared as most serious candidate in this category. However, also this model has some serious difficulties to overcome. General relativity is a very constrained theory, and generally even infrared modifications will also be felt as strongly interacting degrees of freedom at much shorter distances.

Since the cosmological constant problem lies at the heart of a fusion between general relativity and quantum mechanics, it is reasonable to look for modifications of either one, or even both. Interpreting the data as pointing towards a very small effective cosmological constant, a modification of GR seems more plausible. As argued before in this thesis, distances of a tenth of a millimeter are very important in understanding a small cosmological constant and quantum mechanics has been tested thoroughly on much smaller distance scales.

However, since even the sometimes very drastic modifications, advocated in the proposals we discussed, do not lead to a satisfactory answer, this could imply that the ultimate theory of quantum gravity might very well be based on very different foundations than imagined so far. It certainly is conceivable that this is partly due to a misunderstanding of quantum mechanics. Perhaps ultimately the world can perfectly be described by a local, deterministic theory, describing the fundamental degrees of freedom of nature, see [437, 438, 439, 440] for interesting steps in this direction. Quantum mechanics would of course still be a perfect description at low energies, but it would be a statistical theory, not describing the ontological degrees of freedom. It is clear, that such a radical step would completely change the nature of a cosmological constant. Whether or not one should look in this direction for a solution to the cosmological constant problem, is much too early to say.

Since there is no convincing argument that naturally sets the cosmological constant to its small value, the anthropic argument of section (8.2) is considered more and more seriously. However, as stated before, we continue to search for a more satisfactory solution.

9.1 Future Evolution of the Universe

If indeed the acceleration of the universe will turn out to be the consequence of a pure non-vanishing cosmological constant, with equation of state $w = -1$, the universe will accelerate forever. But, because of the existence of an event horizon in de Sitter space, the observable universe will continue to shrink, as other sources, not gravitationally bound to our local group, will vanish beyond the horizon. Simulations based on the Λ CDM model have shown that roughly 100 billion years from now the observable universe, from our point of view, will consist of only one single galaxy, a merger product of the Milky Way and the Andromeda galaxies [441].

For variable dark energy models, there are several possibilities, even a future collapse into a Big Crunch can be part of the scenario [442, 443, 444].

A | Conventions and Definitions

Throughout this thesis, we use a metric $(-, +, +, +)$ and write Einstein's equations as:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} - \Lambda g_{\mu\nu} = -8\pi GT_{\mu\nu}. \quad (\text{A.1})$$

The minus sign with which the cosmological constant enters the Einstein equation, means that we take the action to be $\propto (R + 2\Lambda)$. Besides, we employ natural units, in which:

$$\hbar = c = k_B = 1. \quad (\text{A.2})$$

Using:

$$\begin{aligned} \hbar &= 1.054571596 \times 10^{-34} \text{ Js} \\ c &= 2.99792458 \times 10^8 \text{ m/s} \\ k_B &= 1.3806503 \times 10^{-23} \text{ J/K.} \end{aligned} \quad (\text{A.3})$$

and $E = mc^2$, only one remaining unit needs to be chosen, usually taken to be mass. The dimensions, denoted by square brackets become:

$$[\text{energy}] = [\text{temperature}] = [\text{mass}], \quad [\text{time}] = [\text{length}] = [\text{mass}]^{-1}. \quad (\text{A.4})$$

Masses are expressed in units of GeV, which can easily be converted back into SI-units, using the following expressions:

energy	1 GeV = 1.6022×10^{-10} J
temperature	1 GeV = 1.1605×10^{13} K
mass	1 GeV = 1.7827×10^{-27} kg
length	1 GeV ⁻¹ = 1.9733×10^{-16} m
time	1 GeV ⁻¹ = 6.6522×10^{-25} s

To denote the cosmological constant, we use a capital Λ , which has dimension $[\text{GeV}]^2$, and a lower case letter λ to denote vacuum energy density, which has dimension $[\text{GeV}]^4$, and is related to Λ :

$$\Lambda = 8\pi G\lambda. \quad (\text{A.5})$$

A collection of useful quantities is listed in the following table:

Gravitational constant	$G = 6.6726 \pm 0.0009 \times 10^{-11} \text{m}^3 \text{kg}^{-1} \text{s}^{-2}$
parsec	$1 \text{ pc} = 3.0856 \times 10^{16} \text{ m}$
Solar mass	$1 M_{\odot} = 1.989 \times 10^{30} \text{ kg}$
Dimensionless Hubble parameter	$h = H_0/100 \text{ kms}^{-1} \text{Mpc}^{-1}$
Density of the Universe	$\rho_0 = 1.8789 \times 10^{-26} \Omega h^2 \text{ kgm}^{-3}$ $= 2.7752 \times 10^{11} \Omega h^2 M_{\odot} \text{Mpc}^{-3}$ $= 11.26 h^2 \text{ protons}/m^3$

B Measured Values of Different Cosmological Parameters

B.1 Hubble Constant

The Hubble Space Telescope (HST) key project has determined the value of the Hubble constant to about 400 Mpc with various secondary indicators based on the primary cepheid distance. Their result, as well as the results of the secondary indicators are listed in table (B.1):

Table B.1: The Value of Hubble's Constant

HST	$H_0 = 75 \pm 10 \text{ km/s/Mpc}$
Type Ia SNe	$H_0 = 71 \pm 2 \text{ stat} \pm 6 \text{ syst km/s/Mpc}$
Tully-Fisher relation	$H_0 = 71 \pm 3 \text{ stat} \pm 7 \text{ syst km/s/Mpc}$
Surface brightness fluctuations	$H_0 = 70 \pm 5 \text{ stat} \pm 6 \text{ syst km/s/Mpc}$
Type II SNe	$H_0 = 72 \pm 9 \text{ stat} \pm 7 \text{ syst km/s/Mpc}$
Fundamental plane of elliptical galaxies	$H_0 = 82 \pm 6 \text{ stat} \pm 9 \text{ syst km/s/Mpc}$
FINAL RESULT	$H_0 = 72 \pm 8 \text{ km/s/Mpc, [24]}$

More recent (2005) data [445] from the CMB and 2dFGRS (large scale structure) give $H_0 = 74 \pm 2$.

Note that often the Hubble parameter is parameterized in terms of a dimensionless quantity h as:

$$H_0 = 100h \text{ km/sec/Mpc} \quad (\text{B.1})$$

The Sunyaev-Zel'dovich effect provides another way of determining H_0 , but it suffers from large systematic errors. Following this route, the result [446] is: $H_0 = 60 \pm 4_{-18}^{+13} \text{ km/s/Mpc}$, consistent with the HST-result.

B.2 Total Energy Density

From CMBR measurements [21, 18, 19]:

$$0.98 \lesssim \Omega_{tot} \lesssim 1.08, \quad (\text{B.2})$$

under the assumption that the Hubble parameter $h > 0.5$.

B.3 Matter in the Universe

Traditionally Ω_M is determined by “weighing” a cluster of galaxies, divide by its luminosity, and extrapolate the result to the universe as a whole. Clusters are not representative samples of the universe, but sufficiently large that such a procedure might work. Applying the virial theorem to cluster dynamics typically yielded values: $\Omega_M = 0.2 \pm 0.1$ [447, 448, 449].

Another way to determine Ω_M is through the value of baryon density, which would also include dark matter [450, 451, 452]. These measurements imply: $\Omega_M = 0.3 \pm 0.1$

Also measurements of the power spectrum of density fluctuations gives information on the amount of matter in the universe, but this information is dependent on the underlying theory, and on the specification of a number of cosmological parameters, [453, 454]. The result is identical to the precious method: $\Omega_M = 0.3 \pm 0.1$.

The total amount of baryons, again from CMBR measurements contributes about:

$$\Omega_B = 0.024 \pm 0.0012 h^{-2} \quad \rightarrow \quad \Omega_B \approx 0.04 - 0.06, \quad (\text{B.3})$$

with $h = 0.72 \pm 0.07$. This is a total amount of baryons, being luminous or not. Therefore we can conclude that most of the universe is non-baryonic.

The dark matter contribution contributes about $\Omega_{DM} \approx 0.20 - 0.35$. The need for dark matter results from a host of observations, relating large scale structure and dynamics, for a summary, see [455]. The prime candidate consists of weakly interacting massive particles, so-called ‘WIMPS’.

Combing the results for the total energy density and the matter energy density, one is led to conclude that there must be at least one other component to the energy density, contributing about 70% of the critical energy density.

B.4 Dark Energy Equation of State

An important cosmological parameter is the dark energy equation of state $w = p/\rho$, which is exactly equal to -1 for a cosmological constant. Different methods have been used to measure this parameter and they are listed in the following table (B.4):

Table B.2: The Value of the Dark Energy Equation of State Parameter

SN Ia and CMB	$-1 \leq w \leq -0.93$ [456]
SN Ia, CMB, HST, large scale structure	$w \leq -0.85$ at 1σ and $w \leq -0.72$ at 2σ [457]
X-ray clusters and SN Ia	$w = 0.95^{+0.30}_{-0.35}$ assuming flat universe [458]
WMAP	$w < -0.78$ at 2σ [18, 19]

By combining results of seven CMB experiments, data on large scale structure, Hubble parameter measurements and supernovae results, bounds found are $-1.38 \leq w \leq -0.82$ at 95 % confidence level in [459].

B.5 Summary

This brings us to the following list of values of cosmological parameters:

Table B.3: Concordance Model

$\Omega_{DE} \approx 0.7$
$\Omega_{DM} \approx 0.26$
$\Omega_B \approx 0.04$
$\Omega_R \approx 5 \times 10^{-5}$
$h \approx 0.72 \pm 8$
$-1.38 \leq w \leq -0.82$

B | Nederlandse Samenvatting

Dit proefschrift gaat over natuurkunde op zowel de allerkleinste afstanden, als op de allergrootste afstanden. Laten we met de allerkleinste schaal beginnen.

Alle materie om ons heen is opgebouwd uit atomen. Atomen bevatten een atoomkern, die bestaat uit protonen, met een positieve elektrische lading, en neutronen, die elektrisch neutraal zijn. Om deze kern beweegt een wolk negatief geladen elektronen. De elektrische kracht is aantrekkend voor tegengestelde ladingen en afstotend voor gelijke ladingen. Dat de positief geladen deeltjes in de kern toch bijeen worden gehouden, is te danken aan de zogeheten sterke kernkracht. De protonen en neutronen zijn elk opgebouwd uit drie zogeheten quarks, en deze quarks worden bijeengehouden door uitwisseling van gluonen onder de sterke kernkracht.

Naast de sterke kernkracht, kennen we ook de zwakke kernkracht, welke o.a. verantwoordelijk is voor radioactief verval van sommige atomen. De elektromagnetische kracht beschrijft naast elektriciteit en magnetisme ook de verschillende verschijningsvormen van elektromagnetische straling. Het zichtbare licht is een van de vormen van elektromagnetische straling. De langere golflengtes noemen we infrarood licht en radiogolven, de kortere golflengtes ultraviolet licht, röntgenstralen en gammastralen.

De fysica van de afgelopen eeuw heeft ons geleerd dat deze drie krachten geformuleerd kunnen worden op dezelfde wiskundige basis. Die basis wordt gegeven door quantum veldentheorie, een combinatie van quantum mechanica en speciale relativiteitstheorie. De specifieke quantum veldentheorie die deze drie krachten beschrijft staat bekend onder de weinig inspirerende naam: het standaard model. Het standaard model beschrijft deze drie krachten tot op afstanden nog veel kleiner dan een atoomkern. Dit model is in de afgelopen 30 jaar zeer nauwkeurig getest in enorm grote deeltjesversnellers zoals in Geneve, Zwitserland, bij het Europese versnellerinstituut CERN. De voorspellingen van het standaard model komen vaak tot tien cijfers achter de komma nauwkeurig overeen met de waarnemingen, een verbluffende precisie!

Naast deze drie krachten is er nog een fundamentele natuurkracht, de zwaartekracht, of gravitatie. De werking van de zwaartekracht wordt het best verklaard door de algemene relativiteitstheorie die Einstein in 1915 heeft geformuleerd. Deze theorie beschrijft gravitatie op een totaal andere manier dan de oude Newtoniaanse gravitatiewet, en leidt daardoor niet alleen tot kleine correcties op Newtoniaanse resultaten, maar doet ook echt nieuwe voorspellingen. En van die nieuwe voorspellingen is dat gravitatie in wezen niets anders is dan de kromming door energie, bijvoorbeeld een zwaar object als de zon, van ruimtetijd. De theorie voorspelt dat deze kromming ook gevoeld wordt door licht, dat afgebogen wordt wanneer het langs een massief object schijnt (je zou je dit enigszins voor kunnen stellen als een zware bal die een gespannen zeil kromt. Als je nu bijvoorbeeld een knikker over het zeil langs de bal laat rollen, dan zal de baan van die knikker afgebogen worden). De theorie doet een zeer precieze voorspelling voor

deze afbuiging, welke al snel na de formulering van de algemene relativiteitstheorie kon worden gecontroleerd bij de eclips van 1922. Een eerste schitterende bevestiging van Einsteins theorie volgde. In de periode daarna werd de theorie verder ontrafeld en haar voorspellingen beter getoetst. De experimentele status van Einsteins algemene relativiteitstheorie is nu praktisch dezelfde als die van het standaard model: uitgebreid getest en in uitstekende staat bevonden. Al zijn er wel kanttekeningen te plaatsen.

De belangrijkste is dat algemene relativiteitstheorie alleen echt goed getoetst is op afstanden groter dan ongeveer een millimeter. De reden daarvoor is dat gravitatie een extreem zwakke kracht is, vergeleken met bijvoorbeeld de elektromagnetische kracht. Elektromagnetische invloeden zijn vele malen sterker op kleine afstanden en dat maakt het erg moeilijk om goede experimenten te doen aan alleen de zwaartekracht. Pas op extreem kleine afstanden van 10^{-33} cm ($10^{-1} = 0,1$; $10^{-2} = 0,01$ etc.) worden gravitatie effecten weer sterk en is de zwaartekracht zelfs de dominante kracht. Dit zijn echter zulke kleine afstanden dat die tot op heden (en in de verre toekomst) niet experimenteel toegankelijk zijn.

Op kosmologische afstandsschalen is gravitatie ook de belangrijkste kracht. Sterren en planeten zijn elektrisch neutraal, dus hun onderlinge elektromagnetische wisselwerking is minimaal. De dynamica van het heelal wordt dan ook bepaald door gravitatie en de algemene relativiteitstheorie is dus ook hier de aangewezen theorie. In de jaren '20 (en grotendeels '30) van de afgelopen eeuw, was het standaard idee dat het heelal statisch is, d.w.z. op grote afstandsschaal onveranderlijk in de tijd. Einstein was dan ook op zoek naar een statische oplossing van zijn vergelijkingen. Echter, een dergelijke oplossing is niet mogelijk wanneer het heelal slechts bestaat uit materie. Einstein realiseerde zich echter dat zijn vergelijkingen een extra term toestaan, genaamd de "kosmologische constante", en dat met deze term een statisch heelal wellicht wel mogelijk was. Indien deze term klein genoeg gekozen zou worden, zou deze term zich alleen op heel grote afstanden doen laten gelden en zouden de waarneembare effecten ervan tot dan toe volledig aan de aandacht ontsnapt kunnen zijn.

Het bleek echter dat een statisch heelal nog steeds niet mogelijk is, en bovendien werd langzaam maar duidelijk dat het heelal ook helemaal niet statisch is, maar expandeert. Daarmee was voor Einstein de noodzaak voor de invoering van de kosmologische constante weg. Geruchten gaan zelfs dat Einstein de invoering van deze extra term later "de grootste blunder" van zijn leven noemde, waarbij hij misschien doelde op het missen van de voorspelling dat het heelal uitdijt.

Hoe dan ook, de vergelijkingen van de algemene relativiteitstheorie zijn zeer strikt en een kosmologische constante term is de enige soort term die men wiskundig kan toestaan. Sterker, er is in feite geen enkele reden waarom je deze term niet zou moeten toevoegen en Einsteins toevoeging was dan ook zeker geen blunder. De kosmologische constante negeren, impliceert dat je deze parameter gelijk nul kiest, en daar is geen goede reden voor, zoals we zullen zien. De fysische interpretatie van een kosmologische constante, is dat deze term een maat geeft voor de kromming van de lege ruimte. Zoals we al zeiden is een van de belangrijkste veranderingen van de algemene relativiteitstheorie, ten opzichte van Newtoniaanse gravitatie, dat in Einsteins theorie gravitatie eigenlijk niets anders is dan de kromming van ruimte en tijd. In principe komt deze kromming tot stand door materie, zware objecten als sterren en planeten vooral, die de ruimtetijd krommen en de zwaartekracht genereren. De kromming van ruimtetijd door de aarde genereert de zwaartekracht, die we zo kennen uit het dagelijkse leven. De

kosmologische constante geeft een extra intrinsieke kromming aan de structuur van de ruimtetijd, zelfs als er helemaal geen materie is, waarbij de grootte van de kosmologische constante de mate van kromming aangeeft.

Binnen de algemene relativiteitstheorie is er geen enkele reden waarom deze kromming een bepaalde waarde zou moeten hebben, en niet een andere. De kosmologische constante zou nul kunnen zijn, maar ook bijvoorbeeld een heel erg groot negatief getal, of een heel erg groot positief getal. Het komt er dan simpelweg op neer om uit de waarnemingen af te leiden wat de waarde van deze parameter is, net als bij veel andere natuurconstanten, zoals Newtons constante, die de sterkte van de zwaartekracht aangeeft, de lading van het elektron, of de constante van Planck.

De situatie verandert echter, en wordt pas echt interessant, als we gaan kijken naar de combinatie van gravitatie en het standaard model. Een kosmologische constante kan namelijk ook opgevat worden als een maat voor de energie-dichtheid van het vacuüm. En daar waar de kosmologische constante binnen de algemene relativiteitstheorie totaal onbepaald is, lijkt het standaard model wel degelijk een voorspelling te doen voor de waarde van deze energie-dichtheid. Dit geeft een andere reden waarom we niet vrij zijn om zomaar aan te nemen dat de kosmologische constante gelijk nul is.

Binnen het standaard model, dus zonder gravitatie, is de absolute waarde van de energiedichtheid van het vacuüm overigens niet meetbaar. Dit komt omdat wanneer we energieën van deeltjes meten, we altijd een relatieve meting doen: de energie van het ene deeltje wordt vergeleken met de energie van een ander deeltje. Bij beide deeltjes eenzelfde willekeurige constante optellen, heeft dan geen enkel effect; het verschil blijft altijd hetzelfde. Dit is een groot verschil met gravitatie. Gravitatie, volgens de algemene relativiteitstheorie, voelt iedere vorm van energie op dezelfde manier, en voelt dus ook de energie van het vacuüm.

Klassiek (niet-quantum mechanisch) zou je geneigd zijn te denken dat de energie van het vacuüm precies nul is. Als immers alle materie wordt weggezogen, alle velden worden weggehaald, blijft er niets over, en is de energie precies gelijk aan nul. Deze redenering is echter niet correct. In de moderne natuurkunde heeft het vacuüm daarvoor een te ingewikkelde structuur. Quantum mechanica heeft ons geleerd dat ook onder die omstandigheden er nog steeds paren van deeltjes en anti-deeltjes geproduceerd worden, en die representeren een bepaalde energie¹. Dit proces van creatie en annihilatie gebeurt continu, en dit geeft een constante energie aan het vacuüm. Een andere manier om de energie van een systeem te minimaliseren is het af te koelen, totdat alle deeltjes stilstaan. Binnen de klassieke theorie is dit mogelijk en resulteert dit in een energie die praktisch nul is. Quantum mechanica echter heeft duidelijk gemaakt dat dit in de praktijk nooit gaat, deeltjes komen nooit helemaal stil te staan, maar zullen altijd in beweging blijven, hetgeen resulteert in een vacuüm energie die dus zeker niet nul is. Daarnaast zijn er ook nog allerlei andere bijdragen aan de energie-dichtheid van het vacuüm binnen het standaard model, bijvoorbeeld gerelateerd aan fase-overgangen.

Als je nu probeert uit te rekenen binnen dit standaard model wat de waarde van deze energie-dichtheid is, dan lijkt dit ‘oneindig’ te geven! We weten echter dat het

¹Anti-deeltjes hebben dezelfde massa als deeltjes, maar verder zijn al hun ‘quantum getallen’, zoals de elektrische lading, precies tegengesteld. Bij de productie van een deeltje en een anti-deeltje blijft dus elektrische lading behouden. Omdat deze deeltjes wel een bepaalde energie hebben, kunnen ze slechts heel kort bestaan. Dit is een van de consequenties van het onzekerheidsprincipe van Heisenberg en een belangrijk resultaat binnen de quantum mechanica.

standaard model een beperkt geldigheidsgebied heeft. Bij heel grote energieën, ofwel extreem kleine afstanden, geeft dit model niet langer een correcte beschrijving van de natuur. Dit komt omdat, zoals we zeiden, op extreem kleine afstanden gravitatie een zeer sterke kracht wordt, en niet langer buiten beschouwing kan worden gelaten. Een theorie van quantum gravitatie, die quantum mechanica combineert met gravitatie lijkt hier nodig te zijn. Als we hier rekening mee houden, dan komt de waarde voor de energie-dichtheid niet uit op oneindig, maar op ongeveer 10^{76}GeV^4 , dus een 1 met 76 nullen, nog steeds een erg groot getal! En maar liefst een factor 10^{120} groter dan de experimentele limiet. Een GeV is een grootheid voor de energie. Ter vergelijking, in de beste deeltjesversnellers op aarde zijn de hoogst haalbare energieën ongeveer $10^3 = 1000 \text{ GeV}$.

Deze energiedichtheid van het vacuüm plus de eventueel aanwezige kromming van de lege ruimte, vormen samen wat we de ‘effectieve kosmologische constante’ noemen. Alleen deze effectieve kosmologische constante is waarneembaar. Deze leidt tot een extra kromming van ruimtetijd, en afhankelijk van de hoeveelheid materie in het heelal tot een inkrimping (als de waarde negatief is) of een uitdijning (bij positieve waarde) van het heelal.

De expansie van het heelal is waar te nemen als ‘roodverschuiving’. Roodverschuiving treedt op doordat bijvoorbeeld elektromagnetische golven, licht in het bijzonder, in een uitdijend heelal een grotere afstand moeten afleggen. Beschouw licht dat van de zon naar ons toe schijnt. Als de afstand tussen de zon en de aarde steeds groter zou worden, terwijl het licht onderweg is, dan verliezen de lichtgolven steeds meer energie. Omdat licht altijd met de lichtsnelheid beweegt, kan het niet langzamer gaan bewegen. Wat er gebeurt, is dat de frequentie van het licht, het aantal trillingen van de lichtgolven per seconde, steeds lager wordt. Als gevolg daarvan wordt de golflengte steeds groter. Blauw licht heeft een kortere golflengte en een hogere frequentie dan rood licht, dus het blauwe licht komt er steeds roder uit te zien, vandaar de naam. Overigens gebeurt hetzelfde met alle golflengtes van het spectrum, ook rood licht ‘roodverschuift’ alleen wordt het dan niet roder, maar achtereenvolgens infrarood licht, microgolven of zelfs radiogolven.

Op een heel intuïtieve manier kan nu een gevoel worden gekregen voor de mate van expansie van het heelal, en dus voor de waarde van de effectieve kosmologische constante. De waarde van deze parameter vertelt je namelijk op welke afstandsschaal je deze roodverschuiving zult waarnemen. De kosmologische constant heeft dimensie $1/\text{afstand}^2$. Dit betekent dat als de waarde $1/(1 \text{ meter})^2$ zou zijn, je zou merken dat objecten op 1 meter afstand van je ogen al zover roodverschoven zouden zijn, dat je ze al helemaal niet meer zou zien. Met andere woorden, het heelal zou zo snel uitdijen dat licht dat op een meter afstand wordt uitgezonden, je ogen nooit meer zou kunnen bereiken². In de praktijk kunnen we natuurlijk zonder problemen met het blote oog enkele kilometers ver zien, zonder dat objecten een andere kleur hebben dan als we ze van dichtbij bekijken. We zien zelfs de maan als een wit baken aan de hemel staan, zonder enige roodverschuiving, hetgeen een sterke limiet geeft op de waarde van de kosmologische constante. Deze parameter is zelfs zo klein dat je het effect ervan alleen kan merken door heel grote afstanden te bestuderen met behulp van sterke telescopen³.

²Wat preciezer gezegd, we bevinden ons dan in een de Sitter heelal, met een straal van een meter, en dus liggen objecten op een meter afstand net op de horizon.

³Binnen ons eigen sterrenstelsel is overigens ook een roodverschuiving waar te nemen, maar deze is een gevolg

In 1998 hebben astronomen met behulp van nieuwe technieken voor het eerst echt precieze metingen kunnen doen om de mate van expansie van het heelal te achterhalen. Hieruit bleek, dat het heelal de laatste ongeveer vijf miljard jaar versneld uitdijt. Een resultaat dat maar weinigen hadden verwacht. Omdat een kosmologische constante zo klein moet zijn, nam men aan dat deze waarschijnlijk exact nul zou zijn. Bovendien, gravitatie is een aantrekkende kracht, dus je zou naïef verwachten dat omdat sterrenstelsels elkaar aantrekken het heelal of met constante snelheid expandeert, of dat de expansie tot een stop komt, en dat het heelal weer begint te krimpen. Een versnelde expansie duidt op een zekere vorm van ‘donkere energie’. De experimenten lijken erop te wijzen dat deze donkere energie precies een kosmologische constante representeert, al hoeft dit niet het geval te zijn⁴.

De geringe versnelling van de expansie betekent dat de effectieve kosmologische constante een hele kleine positieve waarde heeft, ongeveer $1/(10^{28} \text{ cm})^2$. Omdat het heelal echter ontzettend groot en leeg is, is deze heel kleine waarde voor de energiedichtheid van het vacuüm wel ongeveer 70% van de totale energiedichtheid van het heelal! Gewone materie draagt slechts 4% bij, donkere materie, een andere wat mysterieuze component, ongeveer 26%. Interessant is verder dat deze waarde van de kosmologische constante, deze vacuüm energie-dichtheid, aan lijkt te geven dat er iets bijzonders gebeurt op afstanden kleiner dan ongeveer een tiende van een millimeter, precies de kritieke afstand tot waarop de algemene relativiteittheorie is getest.

We meten dus een hele kleine effectieve kosmologische constante. Deze effectieve kosmologische constante is de som van de kromming van de lege ruimte, plus de energiedichtheid van het vacuüm zoals de quantum mechanica voorschrijft. Deze laatste bijdrage is enorm groot, terwijl we geen idee hebben wat de waarde van de eerste component is. De grote vraag is nu, waarom is die effectieve kosmologische constante zo verschrikkelijk klein? Waarom tellen al de verschillende bijdragen aan de energiedichtheid van het vacuüm zo bij elkaar op dat ze, samen met de kromming van de lege ruimte precies deze heel kleine waarde geven?

De meest eenvoudige houding is om simpelweg de mate van kromming van de lege ruimte zodanig te kiezen, dat de som uitkomt op precies de waarde van de effectieve kosmologische constante die we waarnemen. Dit is echter nogal kunstmatig. De kromming van de lege ruimte moeten we dan nauwkeurig kiezen tot op 120 decimalen achter de komma. Praktisch levert dit geen bezwaren op, en de theorie beschrijft dan prima de waarnemingen. Maar we hebben dan eigenlijk geen idee wat we aan het doen zijn. Het kan zijn dat er geen betere verklaring is, maar we hebben goede redenen aan te nemen dat er meer aan de hand is, dus we proberen een betere oplossing te vinden.

Verschillende andere suggesties om dit probleem op te lossen zijn naar voren gebracht en passeren in dit proefschrift de revue. De meest elegante oplossing zou een symmetrie zijn die bepaalt dat de kosmologische constante gelijk nul is. Zeker voor 1998 leek dit de meest waarschijnlijke oplossing, omdat men toen voetstoots aannam dat de kosmologische constante inderdaad precies nul zou zijn. Tegenwoordig ligt het iets

van het Doppler effect, hetzelfde effect als de veranderende toonhoogte van bijvoorbeeld een langrijdende politie-auto met sirene. Sommige objecten komen soms wat naar ons toe, anderen gaan wat van ons af. Het netto-effect daarvan middelt uit naar nul. Pas wanneer we objecten bestuderen die buiten ons eigen sterrenstelsel staan kunnen we de effecten van de versnelde expansie van het heelal waarnemen en meten hoe snel het heelal uitdijt.

⁴Andere mogelijkheden zijn bijvoorbeeld een scalair veld met een hele kleine massa en specifieke potentiaal met $w \approx -1$ (‘quintessence’), of een exotische gravitationele term.

subtieler, hoewel de waarnemingen nog niet precies genoeg zijn om inderdaad 100 % duidelijkheid te geven dat de versnelde expansie inderdaad gegenereerd wordt door een kosmologische constante. Hoe dan ook, het ontdekken van een extra symmetrie in de natuurwetten wanneer de waarde van de kosmologische constante nul is, zou enorme vooruitgang betekenen. Helaas is tot op heden een dergelijke symmetrie niet gevonden. In hoofdstuk drie van dit proefschrift besteden we uitgebreid aandacht aan dit type oplossing, en geven we argumenten waarom de voorgestelde ideeën niet werken.

In het hoofdstuk daarna behandelen we een eigen idee dat gebaseerd is op een uitgebreidere symmetrie van transformaties naar de imaginaire ruimte. Dit is een zeer interessant idee, en zou de kosmologische constante op nul kunnen zetten. Helaas zitten er nog wel een paar nadelen aan dit idee, zoals bijvoorbeeld problemen met interacties en massa's van fermionen.

In hoofdstuk vijf wordt een ander type oplossing geanalyseerd. Bij deze voorstellen gaat men er vanuit dat de kosmologische constante in het vroege heelal een zeer grote waarde had. We nemen nu een hele kleine kosmologische constante waar als gevolg van quantum mechanische screening effecten tijdens de evolutie van het heelal. Dit scenario is vaak gebaseerd op een soort terugkoppelings mechanisme, in de zin dat de screening effecten gevoed worden door de kosmologische constante en derhalve doorwerken totdat de kosmologische constante nul is. Dit is in principe een zeer aantrekkelijk idee, er is nauwelijks nieuwe speculatieve fysica voor nodig. De kosmologische constante is dan simpelweg klein, omdat het heelal erg oud is. Echter, deze screening effecten zijn over het algemeen uiterst minimaal en kunnen nauwelijks enig soelaas bieden. Quantum effecten spelen een zeer belangrijke rol op kleine afstandsschalen, kleiner dan een millimeter ruwweg, maar hun invloed op heel grote afstanden op de kosmologische constante is veel te klein om het probleem op te lossen.

Het daaropvolgende hoofdstuk is gewijd aan een specifiek geval van zo'n terugkoppelingsmechanisme en vinden hier veel van de problemen die standaard zijn in een dergelijke opzet.

In hoofdstuk zeven wordt een ander zeer natuurlijk idee uitgewerkt. Zoals gezegd is binnen het standaard model de waarde van de vacuüm energie onmeetbaar en totaal irrelevant. Het kosmologische constante probleem ontstaat pas wanneer we gaan kijken naar de theorie voor gravitatie. Wellicht dus dat de algemene relativiteitstheorie correcties nodig heeft. Het zou kunnen dat vacuüm energie helemaal geen gravitationeel effect heeft, of misschien leven we in meer dimensies dan de drie ruimtelijke en een tijddimensie die we om ons heen zien, en kromt vacuüm energie alleen deze extra dimensies. Zeer interessante, concrete scenario's zijn bedacht, gebaseerd op deze ideeën. Hoe aardig het echter ook klinkt, correcties aanbrengen op algemene relativiteitstheorie is beslist geen sinecure. De theorie steekt wiskundig zo strikt in elkaar, dat er weinig mogelijkheden zijn tot aanpassingen. Nog belangrijker, de theorie werkt fantastisch goed op afstanden vanaf ongeveer een millimeter, tot afstanden zeker zo groot als ons zonnestelsel. Pas als we op afstanden kleiner dan een millimeter of op nog grotere afstanden gaan kijken, zouden we correcties mogen toestaan. Dit geeft grote moeilijkheden voor alle soorten correcties die tot nu toe gesuggereerd zijn. Over het algemeen is het zo dat je de theorie niet op heel grote afstanden kan wijzigen, zodanig dat op kleinere afstanden de goede punten van de theorie overeind blijven. We staan uitgebreid stil bij het zogenoemde 'DGP-model', genoemd naar de bedenkers ervan, en zien hier duidelijk waar de tekortkomingen liggen. Veel scenario's in dit hoofdstuk geven een

‘oplossing’ van het kosmologische constante probleem, met als vervelende bijwerking dat ons zonnestelsel niet meer stabiel is, en de planeten allang alle kanten op hadden moeten vliegen. Je belandt van de wal in de sloot.

In het laatste hoofdstuk voor de conclusies worden statistische benaderingen bestudeerd. In deze verklaringen gaat men ervan uit dat er in feite een heleboel heelallen bestaan, waar wij nooit mee in contact zouden kunnen komen, maar die allemaal een andere waarde voor de kosmologische constante hebben. Deze ideeën maken gebruik van het feit dat voor slechts een beperkt gebied van waarden voor de kosmologische constante leven kan ontstaan. In heelallen met een kosmologische constante die ruwweg 10% groter of kleiner is dan de waarde in ons heelal, zullen zich nooit sterrenstelsels en planeten kunnen ontwikkelen zoals wij die kennen. Het heelal zou daar simpelweg te snel voor expanderen, of te snel weer voor in elkaar klappen in een ‘Big Crunch’. De waarde van de kosmologische constante die wij waarnemen is zo klein, juist omdat we er zijn, en metingen kunnen doen. Het grote bezwaar tegen dit soort verklaringen is dat we slechts in een heelal leven, en nooit de mogelijkheid hebben om andere heelallen waar te nemen. Een statistische verklaring voor de waarde van de kosmologische constante in ons heelal, lijkt dan ook nooit falsifieerbaar, laat staan verifieerbaar. We kunnen op geen enkele manier toetsen of deze verklaring al dan niet correct is. Men kan zich daarom zelfs afvragen of dit type verklaring werkelijk wetenschappelijke uitspraken doet.

Toch heeft dit idee, dat bekend staat als het ‘antropisch principe’, een grote aanhang binnen de wetenschappelijke wereld gekregen. Een belangrijke reden hiervoor is dat verschillende theorieën, bijvoorbeeld over inflatie, maar ook stringtheorie, aanlopen tegen de situatie dat er geen unieke oplossingen voor een heelal bestaan van hun vergelijkingen. Het lijkt dan logisch dit te interpreteren in de zin dat alle oplossingen gerealiseerd worden, met een ‘multiversum’, in tegenstelling tot een universum, als gevolg. Wij zullen als we waarnemingen doen altijd merken dat we ons in een heelal bevinden met een hele kleine kosmologische constante.

Wij blijven echter zoeken naar een meer bevredigender verklaring en zien dit als weer een ander teken hoe belangrijk het is te weten waarom die effectieve kosmologische constante zo klein is. Belangrijke nieuwe wendingen zouden kunnen ontstaan wanneer in de nieuwe grote deeltjesversneller van CERN in Geneve een belangrijke sector van het standaard model eindelijk experimenteel kan worden onderzocht. Dit betreft namelijk precies een belangrijke aanwijzing voor de vacuüm structuur van het standaard model, welke gecentreerd is rondom het zogenaamde Higgs deeltje.

Daarnaast worden de metingen aan de zwaartekracht op afstanden kleiner dan een millimeter steeds nauwkeuriger, wat verrassingen zou kunnen brengen. Indien afwijkingen van de standaard theorie gevonden zouden worden, dan zou dit grote implicaties voor het kosmologische constante probleem kunnen hebben.

Een derde lichtpuntje tenslotte, is dat ook de kosmologische waarnemingen nog steeds beter worden. Het belangrijke nieuws dat we hiervan nodig hebben is antwoord op de vraag of de versnelde expansie inderdaad door een kosmologische constante wordt gegenereerd, of toch door een andere vorm van donkere energie. Indien dit laatste het geval is, betekent dit dat de kosmologische constante toch identiek nul is, en dan lijkt een symmetrie als verklaring de meest voor de hand liggende. Die moet dan alleen nog wel ontdekt worden...

B | Dankwoord

Verschillende mensen hebben op directe of indirecte wijze bijgedragen aan de totstandkoming van dit proefschrift. In de eerste plaats wil ik mijn promotor, Gerard 't Hooft, bedanken voor de mij geboden mogelijkheid een promotie onderzoek te verrichten op het raakvlak van elementaire deeltjesfysica en kosmologie, de twee gebieden die mij van jongs af aan zeer geboeid hebben. Ik heb erg veel geleerd van je kritische houding en geweldige intuïtie. Waar vele anderen enthousiast kunnen zijn voor een of meer van de vele scenario's die in dit proefschrift besproken zijn, wist jij vaak al heel snel de zwakke plekken en zelfs onmogelijkheden van de voorgestelde suggesties bloot te leggen. Dit heeft er mede voor gezorgd dat je me niet zo heel vaak 'uit een moeras hoefde te trekken'. Voor praktisch al mijn vragen kon ik bij je binnen lopen, waarop je vaak een krijtje pakte en het bord binnen een mum van tijd vol schreef. Van deze discussies heb ik enorm veel opgestoken en ze hebben grote bewondering opgewekt voor je manier van werken.

Daarnaast speciale dank aan mijn paranimfen Reinier Warntjes en Björn Garbrecht voor hun interesse en grote enthousiasme voor het 'paranimf-schap'. De vele stapavonden en soms wat bizarre 'Reinier-muziek' hebben zeker hun sporen achterlaten in dit werk, echter uitsluitend positieve aspecten. Very pleasant were also the many discussions and beers I have had with Björn Garbrecht and I am happy that our friendship survived beyond the COSMO-04 conference.

Mijn onderzoek isoleerde me enigszins van de andere theoretici binnen het Spinoza Instituut, maar het was desondanks een stimulerende omgeving om onderzoek te doen. In particular I would like to thank my office mates through the years: Vladimir Shevchenko, Giovanni Arcioni (Avanti Savoia!), Emiliano Imeroni, Fernando Alday and Pedro Machado. They all caused a lot of distraction, but always made it a very fun place to work.

Speciale dank ook voor Marijn Davidse voor de hulp bij het ontdekken van L^AT_EX-trucs voor de layout en bij de gemeenschappelijke besommeringen bij het beeindigen van een promotieonderzoek.

Daarnaast wil ik graag de oud-collega's van de Wetenschapswinkel bedanken, in het bijzonder de coordinator Patricia Huisman-Kleinheerenbrink. Ik heb altijd met veel plezier geprobeerd conclusies van rapporten onderuit te halen en denk dat dit ook een goede basis was voor mijn promotieonderzoek. De sfeer op de Winkel maakte het een zeer gezellige en boeiende tijd waarin ik veel heb geleerd.

Verder zou ik graag naast Reinier en Björn, ook een aantal andere goede vrienden, Michel Overman (ik had nooit kunnen denken dat Egypte nog zoveel inspiratie zou herbergen voor zo'n proefschrift), Patrick van Eck, Anthony ter Horst, Robert Zuidbroek, Patrick van Loeff, en Dennis van Mil specifiek willen bedanken voor de vele uren ontspanning, in de kroeg en anderszins, die onontbeerlijk zijn voor het doen van een

promotieonderzoek. Daarnaast dank aan Martin Bootsma en Robbert van Vossen voor de vele mogelijkheden mijn agressie af te reageren op de schaakstukken. Ook Johan Noldus wil ik hiervoor danken, al was het schaken en de rode wijn vooral decor voor onze natuurkunde discussies. Het was ‘zeer plezant’ te spreken met iemand met dezelfde ideeën over het doen van natuurkunde en over de knelpunten in de huidige inzichten.

Voor een andere uitlaatklep wil ik graag Harm Jan Verweij en Thom Zwagers bedanken. Het snookeren was de ideale manier om mijn nieuwe theorieën over de zwaartekracht in de praktijk te testen. Helaas echter verlopen experimenten nooit geheel zonder kleerscheuren.

Tenslotte wil ik graag mijn ouders bedanken voor hun onvoorwaardelijke steun en interesse. Speciaal ook dank aan mijn vader voor zijn interesse in sterrenkunde en kosmologie, waardoor ook bij mij al jong het heilige vuur werd aangewakkerd.

B | Curriculum Vitae

The author was born in Almelo, The Netherlands, on January 21, 1977. After graduating from the Pius X College in Almelo in 1995, he started his physics studies at Utrecht University. He spent the '98-'99 academic year as an exchange student at the University of California, Berkeley. After returning to The Netherlands, he wrote his Master's thesis on supersymmetry under supervision of Prof. Dr. G. 't Hooft. Graduation followed in 2000.

After obtaining another Master's degree, this time in Philosophy with a thesis on Foundations of Quantum Field Theories, under supervision of Prof. D.G.B.J. Dieks, he became a Ph.D. student, again under supervision of Prof. Dr. G. 't Hooft, at Utrecht University, in 2001. During the past years, in addition to conducting research, the author also has assisted in teaching several courses. He has attended several conferences and has presented his work in Erice (It), Barcelona (Sp), Ambleside (UK), Toronto (Ca), Heidelberg (Ge), Bonn (Ge) and different places in The Netherlands. Besides, interviews regarding his research were published in the 'U-blad' (University newspaper) and in NRC-Handelsblad (interview together with Prof. Dr. G. 't Hooft).

Bibliografie

- [1] Y. B. Zeldovich, “Cosmological Constant and Elementary Particles,” *JETP Lett.* **6** (1967) 316.
- [2] Y. B. Zeldovich *Soviet Physics Uspekhi* **11** (1968) 381.
- [3] M. Veltman *Phys. Rev. Lett.* **34** (1975) 777.
- [4] W. Pauli, *Die allgemeinen Prinzipien der Wellmechanik*, vol. Handbuch der Physik, Vol. XXIV (1933). Springer-Verlag, 1990. New edition by N. Straumann.
- [5] N. Straumann, “The history of the cosmological constant problem,” [gr-qc/0208027](#).
- [6] R. L. Jaffe, “The Casimir effect and the quantum vacuum,” [hep-th/0503158](#).
- [7] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, “Observational evidence from supernovae for an accelerating universe and a cosmological constant,” *Astron. J.* **116** (1998) 1009–1038, [astro-ph/9805201](#).
- [8] **Supernova Search Team** Collaboration, P. M. Garnavich *et al.*, “Supernova limits on the cosmic equation of state,” *Astrophys. J.* **509** (1998) 74–79, [astro-ph/9806396](#).
- [9] A. V. Filippenko and A. G. Riess, “Results from the high-z supernova search team,” *Phys. Rept.* **307** (1998) 31–44, [astro-ph/9807008](#).
- [10] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et al.*, “Measurements of omega and lambda from 42 high-redshift supernovae,” *Astrophys. J.* **517** (1999) 565–586, [astro-ph/9812133](#).
- [11] **Supernova cosmology project** Collaboration, S. Perlmutter *et al.*, “Cosmology from Type Ia Supernovae,” *Bull. Am. Astron. Soc.* **29** (1997) 1351, [astro-ph/9812473](#).
- [12] A. G. Riess *et al.*, “Tests of the accelerating universe with near-infrared observations of a high-redshift type Ia supernova,” *Astrophys. J.* **536** (2000) 62, [astro-ph/0001384](#).
- [13] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, “The farthest known supernova: support for an accelerating universe and a glimpse of the epoch of deceleration,” *Astrophys. J.* **560** (2001) 49–71, [astro-ph/0104455](#).
- [14] **Supernova Search Team** Collaboration, J. L. Tonry *et al.*, “Cosmological results from high-z supernovae,” *Astrophys. J.* **594** (2003) 1–24, [astro-ph/0305008](#).
- [15] **The Supernova Cosmology Project** Collaboration, R. A. Knop *et al.*, “New constraints on Ω_M , Ω_Λ , and w from an independent set of eleven high-redshift supernovae Observed with HST,” *Astrophys. J.* **598** (2003) 102, [astro-ph/0309368](#).
- [16] B. J. Barris *et al.*, “23 high redshift supernovae from the Ifa deep survey: doubling the SN sample at $z_i 0.7$,” *Astrophys. J.* **602** (2004) 571–594, [astro-ph/0310843](#).
- [17] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, “Type Ia supernova discoveries at $z_i 1$ from the Hubble space telescope: evidence for past deceleration and constraints on dark energy evolution,” *Astrophys. J.* **607** (2004) 665–687, [astro-ph/0402512](#).
- [18] C. L. Bennett *et al.*, “First year Wilkinson microwave anisotropy probe (WMAP) observations: preliminary maps and basic results,” *Astrophys. J. Suppl.* **148** (2003) 1, [astro-ph/0302207](#).
- [19] **WMAP** Collaboration, D. N. Spergel *et al.*, “First year Wilkinson microwave anisotropy probe (WMAP) observations: determination of cosmological parameters,” *Astrophys. J. Suppl.* **148** (2003) 175, [astro-ph/0302209](#).

- [20] **Boomerang** Collaboration, C. B. Netterfield *et al.*, “A measurement by BOOMERANG of multiple peaks in the angular power spectrum of the cosmic microwave background,” *Astrophys. J.* **571** (2002) 604–614, [astro-ph/0104460](#).
- [21] **Boomerang** Collaboration, P. de Bernardis *et al.*, “A flat universe from high-resolution maps of the cosmic microwave background radiation,” *Nature* **404** (2000) 955–959, [astro-ph/0004404](#).
- [22] **SDSS** Collaboration, M. Tegmark *et al.*, “The 3D power spectrum of galaxies from the SDSS,” *Astrophys. J.* **606** (2004) 702–740, [astro-ph/0310725](#).
- [23] N. Afshordi, Y. Loh, and M. A. Strauss, “Cross-correlation of the cosmic microwave background with the 2MASS galaxy survey: signatures of dark energy, hot gas, and point sources,” *Phys. Rev.* **D69** (2004) 083524, [astro-ph/0308260](#).
- [24] W. L. Freedman *et al.*, “Final results from the Hubble space telescope key project to measure the Hubble constant,” *Astrophys. J.* **553** (2001) 47–72, [astro-ph/0012376](#).
- [25] B. S. DeWitt, “Quantum field theory in curved space-time,” *Phys. Rept.* **19** (1975) 295–357.
- [26] E. K. Akhmedov, “Vacuum energy and relativistic invariance,” [hep-th/0204048](#).
- [27] G. Ossola and A. Sirlin, “Considerations concerning the contributions of fundamental particles to the vacuum energy density,” *Eur. Phys. J.* **C31** (2003) 165–175, [hep-ph/0305050](#).
- [28] M. J. G. Veltman, “QUANTUM THEORY OF GRAVITATION,”. In *Les Houches 1975, Proceedings, Methods In Field Theory*, Amsterdam 1976, 265-327.
- [29] R. M. Wald, “Quantum field theory in curved space-time,”. Lecture given at Les Houches Summer School on Gravitation and Quantizations, Session 57, Les Houches, France, 5 Jul - 1 Aug 1992.
- [30] N. Birrel and P. Davies, *Quantum fields in curved space*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1982.
- [31] R. M. Wald, “Quantum field theory in curved space-time and black hole thermodynamics,”. Chicago, USA: Univ. Pr. (1994) 205 p.
- [32] D. Sorkin, “Forks in the road on the way to quantum gravity,” [gr-qc/9706002](#).
- [33] S. W. Hawking and G. Ellis, “The large-scale structure of spacetime,”. Cambridge, Uk: Univ. Pr. 1973.
- [34] E. Witten, “Quantum gravity in de Sitter space,” [hep-th/0106109](#).
- [35] N. Goheer, M. Kleban, and L. Susskind, “The trouble with de Sitter space,” *JHEP* **07** (2003) 056, [hep-th/0212209](#).
- [36] D. Klemm and L. Vanzo, “Aspects of quantum gravity in de Sitter spaces,” *JCAP* **0411** (2004) 006, [hep-th/0407255](#).
- [37] A. Strominger, “Inflation and the dS/CFT correspondence,” *JHEP* **11** (2001) 049, [hep-th/0110087](#).
- [38] T. Padmanabhan, “Cosmological constant: The weight of the vacuum,” *Phys. Rept.* **380** (2003) 235–320, [hep-th/0212290](#).
- [39] T. Banks, “Cosmological breaking of supersymmetry or little Lambda goes back to the future. II,” [hep-th/0007146](#).
- [40] S. Weinberg *Rev. Mod. Phys.* **61** (1989) 1.
- [41] S. Weinberg, “Theories of the cosmological constant,” [astro-ph/9610044](#).
- [42] S. Weinberg, “The cosmological constant problems,” [astro-ph/0005265](#).
- [43] G. Dvali, G. Gabadadze, and M. Shifman, “Diluting cosmological constant via large distance modification of gravity,” [hep-th/0208096](#).
- [44] S. D. H. Hsu and A. Zee, “A speculative relation between the cosmological constant and the Planck mass,” [hep-th/0406142](#).

- [45] T. Banks, “Cosmological breaking of supersymmetry or little Lambda goes back to the future. II,” [hep-th/0007146](#).
- [46] S. M. Carroll, “Why is the universe accelerating?,” *ECONF C0307282* (2003) TTH09, [astro-ph/0310342](#).
- [47] A. A. Starobinsky, “Beyond the simplest inflationary cosmological models,” *Grav. Cosmol.* **4** (1998) 88–99, [astro-ph/9811360](#).
- [48] G. ’t Hooft, “A physical interpretation of gravitational instantons,” *Nucl. Phys.* **B315** (1989) 517.
- [49] S. Nobbenhuis, “Categorizing different approaches to the cosmological constant problem,” *Found. Phys.* **36** (5) (2006) [gr-qc/0411093](#).
- [50] G. ’t Hooft and S. Nobbenhuis, “Invariance under complex transformations, and its relevance to the cosmological constant problem,” *Class. Quant. Grav.* **23** (2006) 3819–3832, [gr-qc/0602076](#).
- [51] A. D. Dolgov, “The Cosmological Constant Problem,”. Talk presented at Rencontre de Moriond Astrophysics Session, Les Arcs, France, Mar 5-12, 1989.
- [52] V. Sahni and A. A. Starobinsky, “The case for a positive cosmological lambda-term,” *Int. J. Mod. Phys.* **D9** (2000) 373–444, [astro-ph/9904398](#).
- [53] S. M. Carroll, “The cosmological constant,” *Living Rev. Rel.* **4** (2001) 1, [astro-ph/0004075](#).
- [54] P. J. E. Peebles and B. Ratra, “The cosmological constant and dark energy,” *Rev. Mod. Phys.* **75** (2003) 559–606, [astro-ph/0207347](#).
- [55] N. Straumann, “The history of the cosmological constant problem,” [gr-qc/0208027](#).
- [56] U. Ellwanger, “The cosmological constant,” [hep-ph/0203252](#).
- [57] J. Yokoyama, “Issues on the cosmological constant,” [gr-qc/0305068](#).
- [58] S. Weinberg, *Gravitation and Cosmology, principles and applications of the general theory of relativity*. John Wiley & Sons, New York, USA, 1972.
- [59] V. Slipher, “table in Eddington (1924) The Mathematical Theory of Relativity 2nd, Ed.”.
- [60] A. Friedmann, “On the curvature of space,” *Z. Phys.* **10** (1922) 377–386.
- [61] A. Friedmann, “On the possibility of a world with constant negative curvature of space,” *Z. Phys.* **21** (1924) 326–332.
- [62] A. Einstein *Sitzungsber. Preuss. Akad. Wiss* (1931) 235–237.
- [63] G. Gamow, “My world line,”. See page 44.
- [64] J. A. Peacock, “Cosmological physics,”. Cambridge, UK: Univ. Pr. (1999) 682 p.
- [65] E. Kolb and M. Turner, *The early universe*. Addison Wesley, 1990.
- [66] J. Garcia-Bellido, “Astrophysics and cosmology,” [hep-ph/0004188](#).
- [67] J. Garcia-Bellido, “Cosmology and astrophysics,” [astro-ph/0502139](#).
- [68] R. R. Caldwell, “A Phantom Menace?,” *Phys. Lett.* **B545** (2002) 23–29, [astro-ph/9908168](#).
- [69] G. W. Gibbons, “Phantom matter and the cosmological constant,” [hep-th/0302199](#).
- [70] P. Singh, M. Sami, and N. Dadhich, “Cosmological dynamics of phantom field,” *Phys. Rev.* **D68** (2003) 023522, [hep-th/0305110](#).
- [71] S. Nojiri and S. D. Odintsov, “Quantum de Sitter cosmology and phantom matter,” *Phys. Lett.* **B562** (2003) 147–152, [hep-th/0303117](#).
- [72] U. Alam and V. Sahni, “Supernova constraints on braneworld dark energy,” [astro-ph/0209443](#).

- [73] V. Sahni and Y. Shtanov, “Braneworld models of dark energy,” *JCAP* **0311** (2003) 014, [astro-ph/0202346](#).
- [74] T. Prokopec, O. Tornkvist, and R. P. Woodard, “Photon mass from inflation,” *Phys. Rev. Lett.* **89** (2002) 101301, [astro-ph/0205331](#).
- [75] W. Hu and M. J. White, “Acoustic signatures in the cosmic microwave background,” *Astrophys. J.* **471** (1996) 30–51, [astro-ph/9602019](#).
- [76] M. Tegmark, “Doppler peaks and all that: CMB anisotropies and what they can tell us,” [astro-ph/9511148](#).
- [77] Y. B. Zeldovich and R. A. Sunyaev, “The interaction of matter and radiation in a hot-model universe,” *Astrophys. Space Sci.* **4** (1969) 301.
- [78] R. A. Sunyaev and Y. B. Zeldovich, “Microwave background radiation as a probe of the contemporary structure and history of the universe,” *Ann. Rev. Astron. Astrophys.* **18** (1980) 537–560.
- [79] U. Seljak and M. Zaldarriaga, “A Line of Sight Approach to Cosmic Microwave Background Anisotropies,” *Astrophys. J.* **469** (1996) 437–444, [astro-ph/9603033](#). See also www.cmbfast.org.
- [80] D. D. Reid, D. W. Kittell, E. E. Arsznov, and G. B. Thompson, “The picture of our universe: a view from modern cosmology,” [astro-ph/0209504](#).
- [81] % tt <http://background.uchicago.edu/~whu/>.
- [82] S. Hanany *et al.*, “MAXIMA-1: A measurement of the cosmic microwave background anisotropy on angular scales of 10 arcminutes to 5 degrees,” *Astrophys. J.* **545** (2000) L5, [astro-ph/0005123](#).
- [83] <http://map.gsfc.nasa.gov/>.
- [84] D. Roy, “Basic constituents of the visible and invisible matter - a microscopic view of the universe,” [physics/0007025](#).
- [85] V. Sahni, “Dark matter and dark energy,” *Lect. Notes Phys.* **653** (2004) 141–180, [astro-ph/0403324](#).
- [86] K. Freese, “The dark side of the universe,” [astro-ph/0508279](#).
- [87] J. R. Ellis, K. A. Olive, Y. Santoso, and V. C. Spanos, “Supersymmetric dark matter in light of WMAP,” *Phys. Lett.* **B565** (2003) 176–182, [hep-ph/0303043](#).
- [88] M. Milgrom, “A Modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis,” *Astrophys. J.* **270** (1983) 365–370.
- [89] M. Milgrom, “A Modification of the Newtonian dynamics: Implications for galaxies,” *Astrophys. J.* **270** (1983) 371–389.
- [90] J. Bekenstein and M. Milgrom, “Does the missing mass problem signal the breakdown of Newtonian gravity?,” *Astrophys. J.* **286** (1984) 7–14.
- [91] J. D. Bekenstein, “Relativistic gravitation theory for the MOND paradigm,” *Phys. Rev.* **D70** (2004) 083509, [astro-ph/0403694](#).
- [92] E. Pointecouteau and J. Silk, “New constraints on MOND from galaxy clusters,” [astro-ph/0505017](#).
- [93] J. Felten and R. Isaacman *Rev. Mod. Phys.* **58** (1986) 689.
- [94] A. Blanchard, M. Douspis, M. Rowan-Robinson, and S. Sarkar, “An alternative to the cosmological ‘concordance model’,” *Astron. Astrophys.* **412** (2003) 35–44, [astro-ph/0304237](#).
- [95] M. Turatto, “Classification of supernovae,” [astro-ph/0301107](#).
- [96] R. Barbon, F. Ciatti, and L. Rosino *Astron. Astrophys.* **72** (1979) 287.
- [97] A. V. Filippenko, “Type Ia supernovae and cosmology,” [astro-ph/0410609](#).

- [98] M. M. Nieto and J. D. Anderson, “Using early data to illuminate the Pioneer anomaly,” [gr-qc/0507052](#).
- [99] J. D. Anderson, M. M. Nieto, and S. G. Turyshev, “A mission to test the pioneer anomaly,” *Int. J. Mod. Phys.* **D11** (2002) 1545–1552, [gr-qc/0205059](#).
- [100] M. M. Nieto and S. G. Turyshev, “Finding the Origin of the Pioneer Anomaly,” *Class. Quant. Grav.* **21** (2004) 4005–4024, [gr-qc/0308017](#).
- [101] S. G. Turyshev, M. Nieto, and J. D. Anderson, “Lessons Learned from the Pioneers 10/11 for a Mission to Test the Pioneer Anomaly,” [gr-qc/0409117](#).
- [102] **Pioneer** Collaboration, H. Dittus *et al.*, “A mission to explore the Pioneer anomaly,” [gr-qc/0506139](#).
- [103] M. T. Jaekel and S. Reynaud, “Gravity tests in the solar system and the Pioneer anomaly,” *Mod. Phys. Lett.* **A20** (2005) 1047, [gr-qc/0410148](#).
- [104] S. Reynaud and M. T. Jaekel, “Testing the Newton law at long distances,” *Int. J. Mod. Phys.* **A20** (2005) 2294, [gr-qc/0501038](#).
- [105] F. Munyaneza and R. D. Viollier, “Heavy neutrino dark matter in the solar system,” [astro-ph/9910566](#).
- [106] R. Foot and R. R. Volkas, “A mirror world explanation for the Pioneer spacecraft anomalies?,” *Phys. Lett.* **B517** (2001) 13–17, [hep-ph/0108051](#).
- [107] R. Foot, “Experimental implications of mirror matter-type dark matter,” *Int. J. Mod. Phys.* **A19** (2004) 3807–3818, [astro-ph/0309330](#).
- [108] M. Milgrom, “MOND–theoretical aspects,” *New Astron. Rev.* **46** (2002) 741–753, [astro-ph/0207231](#).
- [109] J. W. Moffat, “Scalar-tensor-vector gravity theory,” [gr-qc/0506021](#).
- [110] O. Bertolami and J. Paramos, “The Pioneer anomaly in a bimetric theory of gravity on the brane,” *Class. Quant. Grav.* **21** (2004) 3309–3321, [gr-qc/0310101](#).
- [111] D. Bailin and A. Love, *Supersymmetric gauge field theory and string theory*, vol. Graduate Studnet Series in Physics. Institute of Physics Publishing, Bristol, 1994.
- [112] J. A. Grifols, R. N. Mohapatra, and A. Riotto, “New astrophysical constraints on the mass of the superlight gravitino,” *Phys. Lett.* **B400** (1997) 124–128, [hep-ph/9612253](#).
- [113] A. B. Lahanas and D. V. Nanopoulos, “THE ROAD TO NO SCALE SUPERGRAVITY,” *Phys. Rept.* **145** (1987) 1.
- [114] E. Witten, “Dimensional reduction of superstring models,” *Phys. Lett.* **B155** (1985) 151.
- [115] K. Bautier, S. Deser, M. Henneaux, and D. Seminara, “No cosmological $D = 11$ supergravity,” *Phys. Lett.* **B406** (1997) 49–53, [hep-th/9704131](#).
- [116] E. Witten, “Is supersymmetry really broken?,” *Int. J. Mod. Phys.* **A10** (1995) 1247–1248, [hep-th/9409111](#). CITATION = HEP-TH 9409111;
- [117] E. Witten, “The cosmological constant from the viewpoint of string theory,” [hep-ph/0002297 v2](#).
- [118] S. Deser, R. Jackiw, and G. ’t Hooft *Ann. Phys.* **152** (1984) 220.
- [119] M. Henneaux, “Energy momentum, angular momentum, and supercharge in $2 + 1$ supergravity,” *Phys. Rev.* **D29** (1984) 2766–2768.
- [120] K. Becker, M. Becker, and A. Strominger, “Three-dimensional supergravity and the cosmological constant,” *Phys. Rev.* **D51** (1995) 6603–6607, [hep-th/9502107](#).
- [121] G. R. Dvali, “Cosmological constant and Fermi-Bose degeneracy,” [hep-th/0004057](#).
- [122] J. D. Edelstein, C. Nunez, and F. A. Schaposnik, “Bogomol’nyi bounds and killing spinors in $D=3$ supergravity,” *Phys. Lett.* **B375** (1996) 163–169, [hep-th/9512117](#).

- [123] J. D. Edelstein, “Semi-local cosmic strings and the cosmological constant problem,” *Phys. Lett.* **B390** (1997) 101–106, [hep-th/9610163](#).
- [124] R. Gregory, V. A. Rubakov, and S. M. Sibiryakov, “Opening up extra dimensions at ultra-large scales,” *Phys. Rev. Lett.* **84** (2000) 5928–5931, [hep-th/0002072](#).
- [125] C. Csaki, J. Erlich, and T. J. Hollowood, “Quasi-localization of gravity by resonant modes,” *Phys. Rev. Lett.* **84** (2000) 5932–5935, [hep-th/0002161](#).
- [126] G. R. Dvali, G. Gabadadze, and M. Porrati, “Metastable gravitons and infinite volume extra dimensions,” *Phys. Lett.* **B484** (2000) 112–118, [hep-th/0002190](#).
- [127] S. Kachru and E. Silverstein, “4d conformal theories and strings on orbifolds,” *Phys. Rev. Lett.* **80** (1998) 4855–4858, [hep-th/9802183](#).
- [128] S. Kachru, J. Kumar, and E. Silverstein, “Vacuum energy cancellation in a non-supersymmetric string,” *Phys. Rev.* **D59** (1999) 106004, [hep-th/9807076](#).
- [129] J. A. Harvey, “String duality and non-supersymmetric strings,” *Phys. Rev.* **D59** (1999) 026002, [hep-th/9807213](#).
- [130] R. Iengo and C. Zhu, “Evidence for nonvanishing cosmological constant in nonSUSY superstring models,” *JHEP* **04** (2000) 028, [hep-th/9912074](#).
- [131] R. Blumenhagen and L. Gorlich, “Orientifolds of non-supersymmetric, asymmetric orbifolds,” *Nucl. Phys.* **B551** (1999) 601–616, [hep-th/9812158](#).
- [132] C. Angelantonj, I. Antoniadis, and K. Forger, “Non-supersymmetric type I strings with zero vacuum energy,” *Nucl. Phys.* **B555** (1999) 116–134, [hep-th/9904092](#).
- [133] C. Angelantonj and I. Antoniadis, “Suppressing the cosmological constant in non-supersymmetric type I strings,” *Nucl. Phys.* **B676** (2004) 129–148, [hep-th/0307254](#).
- [134] S. Deser *Ann. Phys.* **54** (1970) 248.
- [135] P. A. M. Dirac, “Long range forces and broken symmetries,” *Proc. Roy. Soc. Lond.* **A333** (1973) 403–418.
- [136] R. Booth, “Scale invariant gravity and the quasi-static universe,” [gr-qc/0203065](#).
- [137] J. Barbour, “Scale-invariant gravity: particle dynamics,” *Class. Quant. Grav.* **20** (2003) 1543–1570, [gr-qc/0211021](#).
- [138] P. D. Mannheim, “Attractive and repulsive gravity,” *Found. Phys.* **30** (2000) 709–746, [gr-qc/0001011](#).
- [139] P. D. Mannheim, “Cosmic acceleration as the solution to the cosmological constant problem,” *Astrophys. J.* **561** (2001) 1–12, [astro-ph/9910093](#).
- [140] P. D. Mannheim, “Conformal gravity and a naturally small cosmological constant,” [astro-ph/9901219](#).
- [141] P. D. Mannheim, “Local and global gravity,” *Found. Phys.* **26** (1996) 1683–1709, [gr-qc/9611038](#).
- [142] P. D. Mannheim and D. Kazanas, “Higgs mechanism and the structure of the energy-momentum tensor in Einstein gravity and conformal gravity,” [gr-qc/9409050](#).
- [143] P. D. Mannheim and D. Kazanas, “Newtonian limit of conformal gravity and the lack of necessity of the second order Poisson equation,” *Gen. Rel. Grav.* **26** (1994) 337–361.
- [144] V. Perlick and C. Xu, “Matching exterior to interior solutions in Weyl gravity: comment on ‘exact vacuum solution to conformal Weyl gravity and galactic rotation curves’,” *Ap. J.* **449** (1995) 47–51.
- [145] J. J. van der Bij, H. van Dam, and Y. J. Ng, “Theory of gravity and the cosmological term: the little group viewpoint,” *Physica* **116A** (1982) 307.
- [146] W. G. Unruh, “A unimodular theory of canonical quantum gravity,” *Phys. Rev.* **D40** (1989) 1048.

- [147] J. Anderson and D. Finkelstein *Am. J. Phys.* **39** (1971) 901.
- [148] J. J. van der Bij, H. van Dam, and Y. J. Ng, “Theory of Gravity and the Cosmological Term: The Little Group Viewpoint,” *Physica* **116A** (1982) 307.
- [149] T. Padmanabhan, “Gravity as elasticity of spacetime: A paradigm to understand horizon thermodynamics and cosmological constant,” *Int. J. Mod. Phys.* **D13** (2004) 2293–2298, [gr-qc/0408051](#).
- [150] T. Padmanabhan, “Dark energy: the cosmological challenge of the millennium,” *Curr. Sci.* **88** (2005) 1057, [astro-ph/0411044](#).
- [151] B. S. DeWitt, “Quantum theory of gravity. I. The canonical theory,” *Phys. Rev.* **160** (1967) 1113–1148.
- [152] B. S. DeWitt, “Quantum theory of gravity. II. The manifestly covariant theory,” *Phys. Rev.* **162** (1967) 1195–1239.
- [153] G. ’t Hooft, “Quantum gravity: a fundamental problem and some radical ideas,”. Talk given at 8th Conf. on General Relativity and Gravitation, Waterloo, Canada, Aug 7-12, 1977, Published in Cargese Summer Inst.1978:0323.
- [154] P. O. Mazur and E. Mottola, “The gravitational measure, solution of the conformal factor problem and stability of the ground state of quantum gravity,” *Nucl. Phys.* **B341** (1990) 187–212.
- [155] G. ’t Hooft, “Dimensional reduction in quantum gravity,” [gr-qc/9310026](#).
- [156] L. Susskind, “The world as a hologram,” *J. Math. Phys.* **36** (1995) 6377–6396, [hep-th/9409089](#).
- [157] S. D. Thomas, “Holographic vacuum energy,” [hep-th/0010145](#).
- [158] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, “Effective field theory, black holes, and the cosmological constant,” *Phys. Rev. Lett.* **82** (1999) 4971–4974, [hep-th/9803132](#).
- [159] S. D. H. Hsu, “Entropy bounds and dark energy,” *Phys. Lett.* **B594** (2004) 13–16, [hep-th/0403052](#).
- [160] M. S. Turner, “Making sense of the new cosmology,” *Int. J. Mod. Phys.* **A17S1** (2002) 180–196, [astro-ph/0202008](#).
- [161] B. Kelleher, “Scale-invariant gravity. II: Geometrodynamics,” [gr-qc/0310109](#).
- [162] M. Li, “A model of holographic dark energy,” *Phys. Lett.* **B603** (2004) 1, [hep-th/0403127](#).
- [163] C. Balazs and I. Szapudi, “Naturalness of the vacuum energy in holographic theories,” [hep-th/0603133](#).
- [164] S. Kachru, R. Kallosh, A. Linde, and S. P. Trivedi, “De Sitter vacua in string theory,” *Phys. Rev.* **D68** (2003) 046005, [hep-th/0301240](#).
- [165] W. Fischler, “Taking De Sitter seriously,”. Talk given at ‘The role of scaling Laws in physics and biology’, Santa Fe.
- [166] R. Bousso, O. DeWolfe, and R. C. Myers, “Unbounded entropy in spacetimes with positive cosmological constant,” *Found. Phys.* **33** (2003) 297–321, [hep-th/0205080](#).
- [167] M. K. Parikh and E. P. Verlinde, “De Sitter holography with a finite number of states,” *JHEP* **01** (2005) 054, [hep-th/0410227](#).
- [168] A. Strominger, “The dS/CFT correspondence,” *JHEP* **10** (2001) 034, [hep-th/0106113](#).
- [169] M. Spradlin, A. Strominger, and A. Volovich, “Les Houches lectures on de Sitter space,” [hep-th/0110007](#).
- [170] G. E. Volovik, “Superfluid analogies of cosmological phenomena,” *Phys. Rept.* **351** (2001) 195–348, [gr-qc/0005091](#).
- [171] G. E. Volovik, “Vacuum energy and cosmological constant: View from condensed matter,” [gr-qc/0101111](#).

- [172] G. E. Volovik, “Vacuum in quantum liquids and in general relativity,” [gr-qc/0104046](#).
- [173] G. Volovik, “Effective gravity and quantum vacuum in superfluids,”. In *Novello, M. (ed.) et al.: Artificial black holes* 127- 177.
- [174] G. E. Volovik, “The universe in a helium droplet,”. Oxford, UK: Clarendon (2003) 509 p.
- [175] G. E. Volovik, “Comment on contributions of fundamental particles to the vacuum energy,” [hep-ph/0306011](#).
- [176] G. E. Volovik, “Cosmological constant and vacuum energy,” *Annalen Phys.* **14** (2005) 165–176, [gr-qc/0405012](#).
- [177] G. E. Volovik, “Vacuum energy: quantum hydrodynamics vs quantum gravity,” [gr-qc/0505104](#).
- [178] H. Kleinert and J. Zaanen, “Nematic world crystal model of gravity explaining the absence of torsion,” *Phys. Lett.* **A324** (2004) 361–365, [gr-qc/0307033](#).
- [179] I. L. Shapiro and J. Sola, “Scaling behavior of the cosmological constant: Interface between quantum field theory and cosmology,” *JHEP* **02** (2002) 006, [hep-th/0012227](#).
- [180] A. Linde, “The universe multiplication and the cosmological constant problem,” *Phys. Lett* **B200** (1988) 272.
- [181] A. Linde, “Inflation, quantum cosmology and the anthropic principle,” [hep-th/0211048 v2](#).
- [182] U. Ellwanger, “Vanishing cosmological constant via gravitational S- duality,” [hep-th/0410265](#).
- [183] H. Nishino and S. Rajpoot, “Hodge duality and cosmological constant,” [hep-th/0404088](#).
- [184] D. E. Kaplan and R. Sundrum, “A symmetry for the cosmological constant,” [hep-th/0505265](#).
- [185] C. M. Bender, “Non-Hermitian quantum field theory,” *Int. J. Mod. Phys.* **A20** (2005) 4646–4652.
- [186] C. M. Bender, H. F. Jones, and R. J. Rivers, “Dual PT-symmetric quantum field theories,” *Phys. Lett.* **B625** (2005) 333–340, [hep-th/0508105](#).
- [187] C. M. Bender and S. Boettcher, “Real Spectra in Non-Hermitian Hamiltonians Having PT Symmetry,” *Phys. Rev. Lett.* **80** (1998) 5243–5246, [physics/9712001](#).
- [188] Z. Ahmed, C. M. Bender, and M. Berry, “Reflectionless Potentials and PT Symmetry,” [quant-ph/0508117](#).
- [189] G. Bonelli and A. M. Boyarsky, “Six dimensional topological gravity and the cosmological constant problem,” *Phys. Lett.* **B490** (2000) 147–153, [hep-th/0004058](#).
- [190] R. Erdem, “A symmetry for vanishing cosmological constant in an extra dimensional toy model,” *Phys. Lett.* **B621** (2005) 11–17, [hep-th/0410063](#).
- [191] Private communication.
- [192] C. M. Bender and A. Turbiner, “Analytic continuation of Eigenvalue problems,” *Phys. Lett. A* **173** (1993) 442–446. [WU-HEP-92-13](#).
- [193] J. Cline, S. Jeon, and G. D. Moore, “The phantom menaced: Constraints on low-energy effective ghosts,” *Phys. Rev.* **D70** (2004) 043543, [hep-ph/0311312](#).
- [194] L. H. Ford, “Quantum instability of de sitter space-time,” *Phys. Rev.* **D31** (1985) 710.
- [195] L. H. Ford and L. Parker, “Quantized gravitational wave perturbations in Robertson- Walker universes,” *Phys. Rev.* **D16** (1977) 1601–1608.
- [196] L. Ford, “What does quantum field theory in curved spacetime have to say about the dark energy?,” [hep-th/0210096](#).
- [197] N. P. Myhrvold, “Runaway particle production in de Sitter space,” *Phys. Rev.* **D28** (1983) 2439.

- [198] N. P. Myhrvold, “The existence and stability of semiclassical de sitter and Anti-de Sitter space-times,” *Phys. Lett.* **B132** (1983) 308.
- [199] A. D. Dolgov, “An attempt to get rid of the cosmological constant,”. In *Cambridge 1982, Proceedings, The Very Early Universe*, 449-458.
- [200] A. D. Dolgov and M. Kawasaki, “Realistic cosmological model with dynamical cancellation of vacuum energy,” [astro-ph/0307442](#).
- [201] A. D. Dolgov and M. Kawasaki, “Stability of a cosmological model with dynamical cancellation of vacuum energy,” [astro-ph/0310822](#).
- [202] G. 't Hooft and M. J. G. Veltman, “One loop divergencies in the theory of gravitation,” *Annales Poincare Phys. Theor.* **A20** (1974) 69–94.
- [203] C. Wetterich, “Cosmology and the fate of dilatation symmetry,” *Nucl. Phys.* **B302** (1988) 668.
- [204] C. Wetterich, “The Cosmon model for an asymptotically vanishing time dependent cosmological 'constant',” *Astron. Astrophys.* **301** (1995) 321–328, [hep-th/9408025](#).
- [205] C. Wetterich, “Conformal fixed point, cosmological constant and quintessence,” *Phys. Rev. Lett.* **90** (2003) 231302, [hep-th/0210156](#).
- [206] S. Mukohyama and L. Randall, “A dynamical approach to the cosmological constant,” *Phys. Rev. Lett.* **92** (2004) 211302, [hep-th/0306108](#).
- [207] S. Mukohyama, “Gravity in the dynamical approach to the cosmological constant,” *Phys. Rev.* **D70** (2004) 063505, [hep-th/0306208](#).
- [208] V. A. Rubakov, “Relaxation of the cosmological constant at inflation?,” *Phys. Rev.* **D61** (2000) 061501, [hep-ph/9911305](#).
- [209] A. Hebecker and C. Wetterich, “Quintessential adjustment of the cosmological constant,” *Phys. Rev. Lett.* **85** (2000) 3339–3342, [hep-ph/0003287](#).
- [210] A. Hebecker, “On dynamical adjustment mechanisms for the cosmological constant,” [hep-ph/0105315](#).
- [211] E. Lifshitz, “On the Gravitational stability of the expanding universe,” *J. Phys. (USSR)* **10** (1946) 116.
- [212] R. Brandenberger, “Back reaction of cosmological perturbations and the cosmological constant problem,” [hep-th/0210165](#).
- [213] G. Geshnizjani and R. Brandenberger, “Back reaction and local cosmological expansion rate,” *Phys. Rev.* **D66** (2002) 123507, [gr-qc/0204074](#).
- [214] R. M. Wald, “General relativity,”. Chicago, Usa: Univ. Pr. (1984) 491p.
- [215] P. Martineau and R. Brandenberger, “Back-reaction: A cosmological panacea,” [astro-ph/0510523](#).
- [216] L. P. Grishchuk, “Density perturbations of quantum mechanical origin and anisotropy of the microwave background,” *Phys. Rev.* **D50** (1994) 7154–7172, [gr-qc/9405059](#).
- [217] J. Martin and D. J. Schwarz, “The influence of cosmological transitions on the evolution of density perturbations,” *Phys. Rev.* **D57** (1998) 3302–3316, [gr-qc/9704049](#).
- [218] L. P. Grishchuk, “Comment on the *Influence of cosmological transitions on the evolution of density perturbations*,” [gr-qc/9801011](#).
- [219] J. Martin and D. J. Schwarz, “Reply to: Comment on *The influence of cosmological transitions on the evolution of density perturbations*,” [gr-qc/9805069](#).
- [220] W. Unruh, “Cosmological long wavelength perturbations,” [astro-ph/9802323](#).
- [221] H. Kodama and T. Hamazaki, “Evolution of cosmological perturbations in the long wavelength limit,” *Phys. Rev.* **D57** (1998) 7177–7185, [gr-qc/9712045](#).
- [222] G. Geshnizjani and R. Brandenberger, “Back reaction of perturbations in two scalar field inflationary models,” *JCAP* **0504** (2005) 006, [hep-th/0310265](#).

- [223] R. H. Brandenberger and C. S. Lam, “Back-reaction of cosmological perturbations in the infinite wavelength approximation,” [hep-th/0407048](#).
- [224] A. Ishibashi and R. M. Wald, “Can the acceleration of our universe be explained by the effects of inhomogeneities?,” [gr-qc/0509108](#).
- [225] E. W. Kolb, S. Matarrese, A. Notari, and A. Riotto, “Cosmological influence of super-Hubble perturbations,” [astro-ph/0410541](#).
- [226] E. W. Kolb, S. Matarrese, and A. Riotto, “On cosmic acceleration without dark energy,” [astro-ph/0506534](#).
- [227] E. W. Kolb, S. Matarrese, A. Notari, and A. Riotto, “Primordial inflation explains why the universe is accelerating today,” [hep-th/0503117](#).
- [228] E. E. Flanagan, “Can superhorizon perturbations drive the acceleration of the universe?,” *Phys. Rev.* **D71** (2005) 103521, [hep-th/0503202](#).
- [229] C. M. Hirata and U. Seljak, “Can superhorizon cosmological perturbations explain the acceleration of the universe?,” *Phys. Rev.* **D72** (2005) 083501, [astro-ph/0503582](#).
- [230] M. Giovannini, “Inhomogeneous dusty universes and their deceleration,” [hep-th/0505222](#).
- [231] H. Alnes, M. Amarzguioui, and O. Gron, “Can a dust dominated universe have accelerated expansion?,” [astro-ph/0506449](#).
- [232] S. Rasanen, “Backreaction and spatial curvature in a dust universe,” *Class. Quant. Grav.* **23** (2006) 1823–1835, [astro-ph/0504005](#).
- [233] G. A. Burnett, “THE HIGH FREQUENCY LIMIT IN GENERAL RELATIVITY,” *J. Math. Phys.* **30** (1989) 90–96.
- [234] L. S. Brown, “Quantum field theory,”. Cambridge, UK: Univ. Pr. (1992) 542 p.
- [235] I. L. Shapiro, “Asymptotically finite theories and the screening of cosmological constant by quantum effects,” *Phys. Lett.* **B329** (1994) 181. Prepared for 20th Yamada International Colloquium on Group Theoretical Methods in Physics, Toyonaka, Japan, 4-9 Jul 1994.
- [236] E. Elizalde, S. D. Odintsov, and I. L. Shapiro, “Asymptotic regimes in quantum gravity at large distances and running Newtonian and cosmological constants,” *Class. Quant. Grav.* **11** (1994) 1607–1614, [hep-th/9404064](#).
- [237] I. L. Shapiro, J. Sola, and H. Stefancic, “Running G and Lambda at low energies from physics at M(X): Possible cosmological and astrophysical implications,” *JCAP* **0501** (2005) 012, [hep-ph/0410095](#).
- [238] A. Babic, B. Guberina, R. Horvat, and H. Stefancic, “Renormalization-group running cosmologies: A scale-setting procedure,” *Phys. Rev.* **D71** (2005) 124041, [astro-ph/0407572](#).
- [239] M. Reuter and H. Weyer, “Quantum gravity at astrophysical distances?,” *JCAP* **0412** (2004) 001, [hep-th/0410119](#).
- [240] C. Espana-Bonet, P. Ruiz-Lapuente, I. L. Shapiro, and J. Sola, “Testing the running of the cosmological constant with type Ia supernovae at high z,” *JCAP* **0402** (2004) 006, [hep-ph/0311171](#).
- [241] I. L. Shapiro, J. Sola, C. Espana-Bonet, and P. Ruiz-Lapuente, “Variable cosmological constant as a Planck scale effect,” *Phys. Lett.* **B574** (2003) 149–155, [astro-ph/0303306](#).
- [242] J. Sola and H. Stefancic, “Effective equation of state for dark energy: Mimicking quintessence and phantom energy through a variable Lambda,” *Phys. Lett.* **B624** (2005) 147–157, [astro-ph/0505133](#).
- [243] M. Reuter and H. Weyer, “Renormalization group improved gravitational actions: A Brans-Dicke approach,” *Phys. Rev.* **D69** (2004) 104022, [hep-th/0311196](#).
- [244] R. Jackiw, C. Nunez, and S. Y. Pi, “Quantum relaxation of the cosmological constant,” [hep-th/0502215](#).
- [245] K. G. Wilson, “Renormalization group and critical phenomena. 2. Phase space cell analysis of critical behavior,” *Phys. Rev.* **B4** (1971) 3184–3205.

- [246] A. M. Polyakov, “String theory as a universal language,” *Phys. Atom. Nucl.* **64** (2001) 540–547, [hep-th/0006132](#).
- [247] R. D. Peccei, J. Sola, and C. Wetterich, “Adjusting the cosmological constant dynamically: cosmons and a new force weaker than gravity,” *Phys. Lett.* **B195** (1987) 183.
- [248] I. Antoniadis and E. Mottola, “4-D quantum gravity in the conformal sector,” *Phys. Rev.* **D45** (1992) 2013–2025.
- [249] I. Antoniadis, P. O. Mazur, and E. Mottola, “Physical states of the quantum conformal factor,” *Phys. Rev.* **D55** (1997) 4770–4784, [hep-th/9509169](#).
- [250] I. Antoniadis, P. O. Mazur, and E. Mottola, “Conformal invariance and cosmic background radiation,” *Phys. Rev. Lett.* **79** (1997) 14–17, [astro-ph/9611208](#).
- [251] I. Antoniadis, P. O. Mazur, and E. Mottola, “Fractal geometry of quantum spacetime at large scales,” *Phys. Lett.* **B444** (1998) 284–292, [hep-th/9808070](#).
- [252] P. O. Mazur and E. Mottola, “Weyl cohomology and the effective action for conformal anomalies,” *Phys. Rev.* **D64** (2001) 104022, [hep-th/0106151](#).
- [253] S. D. Odintsov and I. L. Shapiro, “Perturbative approach to induced quantum gravity,” *Class. Quant. Grav.* **8** (1991) L57–L60.
- [254] N. C. Tsamis and R. P. Woodard, “Relaxing the cosmological constant,” *Phys. Lett.* **B301** (1993) 351–357.
- [255] N. C. Tsamis and R. P. Woodard, “Quantum gravity slows inflation,” *Nucl. Phys.* **B474** (1996) 235–248, [hep-ph/9602315](#).
- [256] N. C. Tsamis and R. P. Woodard, “The quantum gravitational back-reaction on inflation,” *Annals Phys.* **253** (1997) 1–54, [hep-ph/9602316](#).
- [257] N. C. Tsamis and R. P. Woodard, “Strong infrared effects in quantum gravity,” *Ann. Phys.* **238** (1995) 1–82.
- [258] N. C. Tsamis and R. P. Woodard, “Nonperturbative models for the quantum gravitational back-reaction on inflation,” *Annals Phys.* **267** (1998) 145–192, [hep-ph/9712331](#).
- [259] N. C. Tsamis and R. P. Woodard, “Post-inflationary dynamics,” *JCAP* **0405** (2004) 006, [hep-ph/0303175](#).
- [260] N. C. Tsamis and R. P. Woodard, “Improved estimates of cosmological perturbations,” *Phys. Rev.* **D69** (2004) 084005, [astro-ph/0307463](#).
- [261] T. R. Taylor and G. Veneziano, “QUANTUM GRAVITY AT LARGE DISTANCES AND THE COSMOLOGICAL CONSTANT,” *Nucl. Phys.* **B345** (1990) 210–230.
- [262] L. P. Grishchuk, “Amplification of gravitational waves in an isotropic universe,” *Sov. Phys. JETP* **40** (1975) 409–415.
- [263] R. P. Woodard, “A leading logarithm approximation for inflationary quantum field theory,” *Nucl. Phys. Proc. Suppl.* **148** (2005) 108–119, [astro-ph/0502556](#).
- [264] R. P. Woodard, “A Newtonian model for the quantum gravitational back-reaction on inflation,” *Nucl. Phys. Proc. Suppl.* **104** (2002) 173–176, [gr-qc/0107042](#).
- [265] L. R. Abramo and R. P. Woodard, “Back-reaction is for real,” *Phys. Rev.* **D65** (2002) 063516, [astro-ph/0109273](#).
- [266] T. Banks and W. Fischler, “An upper bound on the number of e-foldings,” [astro-ph/0307459](#).
- [267] A. A. Starobinsky and J. Yokoyama, “Equilibrium state of a selfinteracting scalar field in the De Sitter background,” *Phys. Rev.* **D50** (1994) 6357–6368, [astro-ph/9407016](#).
- [268] A. Vilenkin and L. H. Ford, “GRAVITATIONAL EFFECTS UPON COSMOLOGICAL PHASE TRANSITIONS,” *Phys. Rev.* **D26** (1982) 1231.
- [269] V. A. Rubakov and M. E. Shaposhnikov, “Extra space-time dimensions: towards a solution to the cosmological constant problem,” *Phys. Lett.* **B125** (1983) 139.

- [270] C. M. Will, “The confrontation between general relativity and experiment,” *Living Rev. Rel.* **4** (2001) 4, [gr-qc/0103036](#).
- [271] R. Durrer and P. Kocian, “Testing braneworlds with the binary pulsar,” *Class. Quant. Grav.* **21** (2004) 2127–2138, [hep-th/0305181](#).
- [272] R. Durrer, “Braneworlds,” *AIP Conf. Proc.* **782** (2005) 202–240, [hep-th/0507006](#).
- [273] D. Langlois, “Is our universe brany?,” [hep-th/0509231](#).
- [274] H. Nilles, A. Papazoglou, and G. Tasinato, “Self-tuning and its footprints,” *Nucl. Phys.* **B677** (2004) 405–429, [hep-th/0309042](#).
- [275] N. Arkani-Hamed, S. Dimopoulos, N. Kaloper, and R. Sundrum, “A small cosmological constant from a large extra dimension,” *Phys. Lett.* **B480** (2000) 193–199, [hep-th/0001197](#).
- [276] S. Kachru, M. B. Schulz, and E. Silverstein, “Self-tuning flat domain walls in 5d gravity and string theory,” *Phys. Rev.* **D62** (2000) 045021, [hep-th/0001206](#).
- [277] P. Binetruy, J. M. Cline, and C. Grojean, “Dynamical instability of brane solutions with a self-tuning cosmological constant,” *Phys. Lett.* **B489** (2000) 403–410, [hep-th/0007029](#).
- [278] C. Csaki, J. Erlich, and C. Grojean, “Gravitational Lorentz violations and adjustment of the cosmological constant in asymmetrically warped spacetimes,” *Nucl. Phys.* **B604** (2001) 312–342, [hep-th/0012143](#).
- [279] C. Csaki, J. Erlich, and C. Grojean, “Essay on gravitation: The cosmological constant problem in brane-worlds and gravitational Lorentz violations,” *Gen. Rel. Grav.* **33** (2001) 1921–1928, [gr-qc/0105114](#).
- [280] S. Forste, Z. Lalak, S. Lavignac, and H. Nilles, “The cosmological constant problem from a brane-world perspective,” *JHEP* **09** (2000) 034, [hep-th/0006139](#).
- [281] S. Forste, Z. Lalak, S. Lavignac, and H. P. Nilles, “A comment on self-tuning and vanishing cosmological constant in the brane world,” *Phys. Lett.* **B481** (2000) 360–364, [hep-th/0002164](#).
- [282] J. M. Cline and H. Firouzjahi, “Quest for a self-tuning brane-world solution to the cosmological constant problem,” *AIP Conf. Proc.* **646** (2003) 197–202, [hep-th/0207155](#).
- [283] E. Verlinde and H. Verlinde, “RG-flow, gravity and the cosmological constant,” *JHEP* **05** (2000) 034, [hep-th/9912018](#).
- [284] E. P. Verlinde and H. L. Verlinde, “RG-flow, gravity and the cosmological constant,” *JHEP* **05** (2000) 034, [hep-th/9912018](#).
- [285] C. Schmidhuber, “AdS(5) and the 4D cosmological constant,” *Nucl. Phys.* **B580** (2000) 140–146, [hep-th/9912156](#).
- [286] C. Schmidhuber, “Micrometer gravitinos and the cosmological constant,” *Nucl. Phys.* **B585** (2000) 385–394, [hep-th/0005248](#).
- [287] S. M. Carroll and M. M. Guica, “Sidestepping the cosmological constant with football-shaped extra dimensions,” [hep-th/0302067](#).
- [288] Y. Aghababaie, C. P. Burgess, S. L. Parameswaran, and F. Quevedo, “Towards a naturally small cosmological constant from branes in 6D supergravity,” *Nucl. Phys.* **B680** (2004) 389–414, [hep-th/0304256](#).
- [289] C. P. Burgess, “Towards a natural theory of dark energy: Supersymmetric large extra dimensions,” *AIP Conf. Proc.* **743** (2005) 417–449, [hep-th/0411140](#).
- [290] A. Kehagias, “A conical tear drop as a vacuum-energy drain for the solution of the cosmological constant problem,” *Phys. Lett.* **B600** (2004) 133–141, [hep-th/0406025](#).
- [291] C. P. Burgess, “Supersymmetric large extra dimensions and the cosmological constant: An update,” *Ann. Phys.* **313** (2004) 283–401, [hep-th/0402200](#).
- [292] C. P. Burgess, “Supersymmetric large extra dimensions and the cosmological constant problem,” [hep-th/0510123](#).

- [293] R. Sundrum, “Compactification for a three-brane universe,” *Phys. Rev.* **D59** (1999) 085010, [hep-ph/9807348](#).
- [294] I. Navarro, “Spheres, deficit angles and the cosmological constant,” *Class. Quant. Grav.* **20** (2003) 3603–3612, [hep-th/0305014](#).
- [295] J. Garriga and M. Porrati, “Football shaped extra dimensions and the absence of self-tuning,” *JHEP* **08** (2004) 028, [hep-th/0406158](#).
- [296] A. J. Tolley, C. P. Burgess, D. Hoover, and Y. Aghababaie, “Bulk singularities and the effective cosmological constant for higher co-dimension branes,” [hep-th/0512218](#).
- [297] I. Y. Arefeva, B. G. Dragovic, and I. V. Volovich, “The extra timelike dimensions lead to a vanishing cosmological constant,” *Phys. Lett.* **B177** (1986) 357.
- [298] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and G. Gabadadze, “Non-local modification of gravity and the cosmological constant problem,” [hep-th/0209227](#).
- [299] G. R. Dvali, G. Gabadadze, and M. Porrati, “4D gravity on a brane in 5D Minkowski space,” *Phys. Lett.* **B485** (2000) 208–214, [hep-th/0005016](#).
- [300] G. R. Dvali and G. Gabadadze, “Gravity on a brane in infinite-volume extra space,” *Phys. Rev.* **D63** (2001) 065007, [hep-th/0008054](#).
- [301] G. Dvali, G. Gabadadze, and M. Shifman, “Diluting cosmological constant in infinite volume extra dimensions,” *Phys. Rev.* **D67** (2003) 044020, [hep-th/0202174](#).
- [302] G. Gabadadze, “Looking at the cosmological constant from infinite-volume bulk,” [hep-th/0408118](#).
- [303] A. D. Sakharov, “Vacuum quantum fluctuations in curved space and the theory of gravitation,” *Sov. Phys. Dokl.* **12** (1968) 1040–1041.
- [304] A. Zee, “Calculating Newton’s gravitational constant in infrared stable Yang-Mills theories,” *Phys. Rev. Lett.* **48** (1982) 295.
- [305] S. L. Adler, “Order r vacuum action functional in scalar free unified theories with spontaneous scale breaking,” *Phys. Rev. Lett.* **44** (1980) 1567.
- [306] S. L. Adler, “A formula for the induced gravitational constant,” *Phys. Lett.* **B95** (1980) 241.
- [307] G. Gabadadze, “ICTP lectures on large extra dimensions,” [hep-ph/0308112](#).
- [308] A. G. Cohen and D. B. Kaplan, “THE EXACT METRIC ABOUT GLOBAL COSMIC STRINGS,” *Phys. Lett.* **B215** (1988) 67.
- [309] R. Gregory, “Non-singular global strings,” *Phys. Rev.* **D54** (1996) 4955–4962, [gr-qc/9606002](#).
- [310] I. Cho, “Inflation and nonsingular spacetimes of cosmic strings,” *Phys. Rev.* **D58** (1998) 103509, [gr-qc/9804086](#).
- [311] R. Gregory, “Nonsingular global string compactifications,” *Phys. Rev. Lett.* **84** (2000) 2564–2567, [hep-th/9911015](#).
- [312] A. Chodos, E. Poppitz, and D. Tsimpis, “Nonsingular deformations of singular compactifications, the cosmological constant, and the hierarchy problem,” *Class. Quant. Grav.* **17** (2000) 3865–3880, [hep-th/0006093](#).
- [313] M. Fierz and W. Pauli, “On relativistic wave equations for particles of arbitrary spin in an electromagnetic field,” *Proc. Roy. Soc. Lond.* **A173** (1939) 211–232.
- [314] C. M. Will and N. Yunes, “Testing alternative theories of gravity using LISA,” *Class. Quant. Grav.* **21** (2004) 4367, [gr-qc/0403100](#).
- [315] H. van Dam and M. J. G. Veltman, “Massive and massless Yang-Mills and gravitational fields,” *Nucl. Phys.* **B22** (1970) 397–411.
- [316] V. I. Zakharov *JETP* **12** (1970) 312.
- [317] A. I. Vainshtein, “To the problem of nonvanishing gravitation mass,” *Phys. Lett.* **B39** (1972) 393–394.

- [318] G. Gabadadze and A. Gruzinov, “Graviton mass or cosmological constant?,” [hep-th/0312074](#).
- [319] A. Gruzinov, “On the graviton mass,” *New Astron.* **10** (2005) 311–314, [astro-ph/0112246](#).
- [320] C. R. Nappi and L. Witten, “Interacting Lagrangian for massive spin two field,” *Phys. Rev.* **D40** (1989) 1095.
- [321] M. J. Duff, C. N. Pope, and K. S. Stelle, “Consistent interacting massive spin-2 requires an infinity of states,” *Phys. Lett.* **B223** (1989) 386.
- [322] M. A. Luty, M. Porrati, and R. Rattazzi, “Strong interactions and stability in the DGP model,” *JHEP* **09** (2003) 029, [hep-th/0303116](#).
- [323] V. A. Rubakov, “Strong coupling in brane-induced gravity in five dimensions,” [hep-th/0303125](#).
- [324] M. Porrati and J. W. Rombouts, “Strong coupling vs. 4-D locality in induced gravity,” *Phys. Rev.* **D69** (2004) 122003, [hep-th/0401211](#).
- [325] A. Nicolis and R. Rattazzi, “Classical and quantum consistency of the DGP model,” *JHEP* **06** (2004) 059, [hep-th/0404159](#).
- [326] C. Deffayet and J. Rombouts, “Ghosts, strong coupling and accidental symmetries in massive gravity,” *Phys. Rev.* **D72** (2005) 044003, [gr-qc/0505134](#).
- [327] A. Lue, “The phenomenology of Dvali-Gabadadze-Porrati cosmologies,” *Phys. Rept.* **423** (2006) 1–48, [astro-ph/0510068](#).
- [328] C. Deffayet, G. R. Dvali, G. Gabadadze, and A. Vainshtein, “Nonperturbative continuity in graviton mass versus perturbative discontinuity,” *Phys. Rev.* **D65** (2002) 044026, [hep-th/0106001](#).
- [329] M. Porrati, “Fully covariant van Dam-Veltman-Zakharov discontinuity, and absence thereof,” *Phys. Lett.* **B534** (2002) 209–215, [hep-th/0203014](#).
- [330] D. Gorbunov, K. Koyama, and S. Sibiryakov, “More on ghosts in DGP model,” *Phys. Rev.* **D73** (2006) 044016, [hep-th/0512097](#).
- [331] M. Porrati, “Strong interaction scale in induced gravity,” talk at Perimeter Institute, Canada, May 13, 2003.
- [332] G. Gabadadze and A. Iglesias, “Short distance non-perturbative effects of large distance modified gravity,” [hep-th/0508201](#).
- [333] I. I. Kogan, S. Mouslopoulos, and A. Papazoglou, “The $m \rightarrow 0$ limit for massive graviton in $dS(4)$ and $AdS(4)$: How to circumvent the van Dam-Veltman-Zakharov discontinuity,” *Phys. Lett.* **B503** (2001) 173–180, [hep-th/0011138](#).
- [334] M. Porrati, “No van Dam-Veltman-Zakharov discontinuity in AdS space,” *Phys. Lett.* **B498** (2001) 92–96, [hep-th/0011152](#).
- [335] F. A. Dilkes, M. J. Duff, J. T. Liu, and H. Sati, “Quantum discontinuity between zero and infinitesimal graviton mass with a Lambda term,” *Phys. Rev. Lett.* **87** (2001) 041301, [hep-th/0102093](#).
- [336] M. J. Duff, J. T. Liu, and H. Sati, “Quantum $M^2 \rightarrow 2\Lambda/3$ discontinuity for massive gravity with a Lambda term,” *Phys. Lett.* **B516** (2001) 156–160, [hep-th/0105008](#).
- [337] S. Deser, J. H. Kay, and K. S. Stelle, “HAMILTONIAN FORMULATION OF SUPERGRAVITY,” *Phys. Rev.* **D16** (1977) 2448.
- [338] S. Deser and A. Waldron, “(Dis)continuities of massless limits in spin 3/2-mediated interactions and cosmological supergravity,” *Phys. Lett.* **B501** (2001) 134–139, [hep-th/0012014](#).
- [339] U. Ellwanger, “Blown-up p-branes and the cosmological constant,” *JCAP* **0311** (2003) 013, [hep-th/0304057](#).
- [340] S. L. Dubovsky and V. A. Rubakov, “Brane-induced gravity in more than one extra dimensions: Violation of equivalence principle and ghost,” *Phys. Rev.* **D67** (2003) 104014, [hep-th/0212222](#).

- [341] C. Middleton and G. Siopsis, “Fat branes in infinite volume extra space,” [hep-th/0210033](#).
- [342] U. Ellwanger, “Brane universes and the cosmological constant,” [hep-th/0509062](#).
- [343] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and G. Gabadadze, “Non-local modification of gravity and the cosmological constant problem,” [hep-th/0209227](#).
- [344] J. W. Moffat, “The cosmological constant problem and nonlocal quantum gravity,” *AIP Conf. Proc.* **646** (2003) 130–138, [hep-th/0207198](#).
- [345] J. W. Moffat, “Cosmological constant problem,” [gr-qc/0312115](#).
- [346] P. Binetruy, C. Deffayet, and D. Langlois, “Non-conventional cosmology from a brane-universe,” *Nucl. Phys.* **B565** (2000) 269–287, [hep-th/9905012](#).
- [347] C. Deffayet, “Cosmology on a brane in Minkowski bulk,” *Phys. Lett.* **B502** (2001) 199–208, [hep-th/0010186](#).
- [348] I. Sawicki and S. M. Carroll, “Cosmological Structure Evolution and CMB Anisotropies in DGP Braneworlds,” [astro-ph/0510364](#).
- [349] R. Dick, “Standard cosmology in the DGP brane model,” *Acta Phys. Polon.* **B32** (2001) 3669–3682, [hep-th/0110162](#).
- [350] M. Ishak, A. Upadhye, and D. N. Spergel, “Is cosmic acceleration a symptom of the breakdown of general relativity?,” [astro-ph/0507184](#).
- [351] N. Arkani-Hamed, H. Cheng, M. A. Luty, and S. Mukohyama, “Ghost condensation and a consistent infrared modification of gravity,” *JHEP* **05** (2004) 074, [hep-th/0312099](#).
- [352] B. R. Heckel, E. G. Adelberger, J. H. Gundlach, M. G. Harris, and H. E. Swanson, “Torsion balance test of spin coupled forces,”. Prepared for International Conference on Orbis Scientiae 1999: Quantum Gravity, Generalized Theory of Gravitation and Superstring Theory Based Unification (28th Conference on High-Energy Physics and Cosmology Since 1964), Coral Gables, Florida, 16-19 Dec 1999.
- [353] D. F. Phillips *et al.*, “Limit on Lorentz and CPT violation of the proton using a hydrogen maser,” *Phys. Rev.* **D63** (2001) 111101, [physics/0008230](#).
- [354] R. Bluhm, “Lorentz and CPT tests in matter and antimatter,” *Nucl. Instrum. Meth.* **B221** (2004) 6–11, [hep-ph/0308281](#).
- [355] F. Cane *et al.*, “Bound on Lorentz- and CPT-violating boost effects for the neutron,” *Phys. Rev. Lett.* **93** (2004) 230801, [physics/0309070](#).
- [356] S. L. Dubovsky, “Star tracks in the ghost condensate,” *JCAP* **0407** (2004) 009, [hep-ph/0403308](#).
- [357] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, and M. Zaldarriaga, “Ghost inflation,” *JCAP* **0404** (2004) 001, [hep-th/0312100](#).
- [358] D. Krotov, C. Rebbi, V. A. Rubakov, and V. Zakharov, “Holes in the ghost condensate,” *Phys. Rev.* **D71** (2005) 045014, [hep-ph/0407081](#).
- [359] R. Sundrum, “Fat gravitons, the cosmological constant and sub-millimeter tests,” *Phys. Rev.* **D69** (2004) 044014, [hep-th/0306106](#).
- [360] S. Weinberg and E. Witten, “Limits on massless particles,” *Phys. Lett.* **B96** (1980) 59.
- [361] C. J. Isham, A. Salam, and J. A. Strathdee, “Nonlinear realizations of space-time symmetries. Scalar and tensor gravity,” *Annals Phys.* **62** (1971) 98–119.
- [362] P. Kraus and E. T. Tomboulis, “Photons and gravitons as Goldstone bosons, and the cosmological constant,” *Phys. Rev.* **D66** (2002) 045015, [hep-th/0203221](#).
- [363] A. Jenkins, “Spontaneous breaking of Lorentz invariance,” *Phys. Rev.* **D69** (2004) 105007, [hep-th/0311127](#).
- [364] T. Jacobson and D. Mattingly, “Gravity with a dynamical preferred frame,” *Phys. Rev.* **D64** (2001) 024028, [gr-qc/0007031](#).

- [365] M. L. Graesser, A. Jenkins, and M. B. Wise, “Spontaneous Lorentz violation and the long-range gravitational preferred-frame effect,” *Phys. Lett.* **B613** (2005) 5–10, [hep-th/0501223](#).
- [366] O. Bertolami and J. Paramos, “The flight of the bumblebee: Vacuum solutions of a gravity model with vector-induced spontaneous Lorentz symmetry breaking,” *Phys. Rev.* **D72** (2005) 044001, [hep-th/0504215](#).
- [367] C. Eling, T. Jacobson, and D. Mattingly, “Einstein-aether theory,” [gr-qc/0410001](#).
- [368] C. M. Will, “Theory and experiment in gravitational physics,”. Cambridge, UK: Univ. Pr. (1993) 380 p.
- [369] J. Alfaro, H. A. Morales-Tecotl, and L. F. Urrutia, “Loop quantum gravity and light propagation,” *Phys. Rev.* **D65** (2002) 103509, [hep-th/0108061](#).
- [370] V. A. Kostelecky and S. Samuel, “Spontaneous breaking of Lorentz symmetry in string theory,” *Phys. Rev.* **D39** (1989) 683.
- [371] S. R. Coleman and S. L. Glashow, “High-energy tests of Lorentz invariance,” *Phys. Rev.* **D59** (1999) 116008, [hep-ph/9812418](#).
- [372] D. Mattingly, “Modern tests of Lorentz invariance,” *Living Rev. Rel.* **8** (2005) 5, [gr-qc/0502097](#).
- [373] G. Dvali, A. Gruzinov, and M. Zaldarriaga, “The accelerated universe and the Moon,” *Phys. Rev.* **D68** (2003) 024012, [hep-ph/0212069](#).
- [374] G. Dvali and M. S. Turner, “Dark energy as a modification of the Friedmann equation,” [astro-ph/0301510](#).
- [375] J. A. Wheeler, “Batelle Rencontres,”. eds. DeWitt, C. and Wheeler, J.A.
- [376] J. B. Hartle and S. W. Hawking, “Wave function of the universe,” *Phys. Rev.* **D28** (1983) 2960–2975.
- [377] I. G. Moss, “Quantum theory, black holes and inflation,”. Appendix C. 3+1-decomposition.
- [378] A. Aurilia, H. Nicolai, and P. K. Townsend, “Hidden constants: the Theta parameter of QCD and the cosmological constant of $N = 8$ supergravity,” *Nucl. Phys.* **B176** (1980) 509.
- [379] M. J. Duff and P. van Nieuwenhuizen, “Quantum inequivalence of different field representations,” *Phys. Lett.* **B94** (1980) 179.
- [380] E. Baum *Phys. Lett.* **B133** (1983) 185.
- [381] A. Gomberoff, M. Henneaux, C. Teitelboim, and F. Wilczek, “Thermal decay of the cosmological constant into black holes,” *Phys. Rev.* **D69** (2004) 083520, [hep-th/0311011](#).
- [382] J. D. Brown and C. Teitelboim, “Neutralization of the cosmological constant by membrane creation,” *Nucl. Phys.* **B297** (1988) 787–836.
- [383] J. D. Brown and C. Teitelboim, “Dynamical neutralization of the cosmological constant,” *Phys. Lett.* **B195** (1987) 177–182.
- [384] S. Hawking, “The cosmological constant is probably zero,” *Phys. Lett* **B134** (1984) 403.
- [385] M. J. Duff, “The cosmological constant is possibly zero, but the proof is probably wrong,” *Phys. Lett.* **B226** (1989) 36.
- [386] R. Bousso and J. Polchinski, “Quantization of four-form fluxes and dynamical neutralization of the cosmological constant,” *JHEP* **06** (2000) 006, [hep-th/0004134](#).
- [387] S. Coleman, “Why there is nothing, rather than something: A theory of the cosmological constant,” *Nucl. Phys.* **B310** (1988) 643.
- [388] S. W. Hawking, “Quantum coherence down the wormhole,” *Phys. Lett.* **B195** (1987) 337.
- [389] A. Strominger, “Vacuum topology and incoherence in quantum gravity,” *Phys. Rev. Lett.* **52** (1984) 1733.

- [390] S. R. Coleman, “Black holes as red herrings: topological fluctuations and the loss of quantum coherence,” *Nucl. Phys.* **B307** (1988) 867.
- [391] W. Fischler, I. R. Klebanov, J. Polchinski, and L. Susskind, “Quantum mechanics of the googolplexus,” *Nucl. Phys.* **B327** (1989) 157.
- [392] W. Fischler and L. Susskind, “A wormhole catastrophe,” *Phys. Lett.* **B217** (1989) 48.
- [393] J. Polchinski, “Decoupling versus excluded volume or return of the giant wormholes,” *Nucl. Phys.* **B325** (1989) 619.
- [394] J. Preskill, “Wormholes in space-time and the constants of nature,” *Nucl. Phys.* **B323** (1989) 141.
- [395] S. R. Coleman and K. Lee, “Escape from the menace of the giant wormholes,” *Phys. Lett.* **B221** (1989) 242.
- [396] A. Linde, *The new inflationary universe scenario*, vol. The Very Early Universe. Cambridge University Press, Cambridge, 1983. Edited by Gibbons, G.W. and Hawking, S.W. and Siklos, S.
- [397] A. Linde, “Eternally existing selfreproducing chaotic inflationary universe,” *Phys. Lett.* **B175** (1986) 395.
- [398] A. Linde and A. Mezhlumian, “From the big bang theory to the theory of a stationary universe,” *Phys. Rev.* **D49** (1994) 1783, [gr-qc/9306035](#).
- [399] A. D. Linde, “Eternal chaotic inflation,” *Mod. Phys. Lett.* **A1** (1986) 81.
- [400] S. Coleman, “Black holes and red herrings: topological fluctuations and the loss of quantum coherence,” *Nucl. Phys.* **B307** (1988) 867.
- [401] A. Linde, *Particle physics and inflationary cosmology*. Harwood Academic Publishers, Chur, Switzerland, 1990.
- [402] A. Vilenkin, “Predictions from quantum cosmology,” *Phys. Rev. Lett.* **74** (1995) 846, [gr-qc/9406010](#).
- [403] J. Garcia-Bellido and A. Linde, “Stationarity from inflation and predictions from quantum cosmology,” *Phys. Rev.* **D51** (1995) 429, [hep-th/9408023](#).
- [404] L. Susskind, “The anthropic landscape of string theory,” [hep-th/0302219](#).
- [405] L. Susskind, “Supersymmetry breaking in the anthropic landscape,” [hep-th/0405189](#).
- [406] F. Denef and M. R. Douglas, “Distributions of flux vacua,” *JHEP* **05** (2004) 072, [hep-th/0404116](#).
- [407] B. Freivogel and L. Susskind, “A framework for the landscape,” *Phys. Rev.* **D70** (2004) 126007, [hep-th/0408133](#).
- [408] V. Agrawal, S. M. Barr, J. F. Donoghue, and D. Seckel, “The anthropic principle and the mass scale of the standard model,” *Phys. Rev.* **D57** (1998) 5480–5492, [hep-ph/9707380](#).
- [409] N. Arkani-Hamed and S. Dimopoulos, “Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC,” *JHEP* **06** (2005) 073, [hep-th/0405159](#).
- [410] <http://www.anthropic-principle.com>.
- [411] S. Weinberg, “Anthropic bound on the cosmological constant,” *Phys. Rev. Lett.* **59** (1987) 2607.
- [412] T. Banks, “T C P, Quantum gravity, the cosmological constant and all that..,” *Nucl. Phys.* **B249** (1985) 332.
- [413] J. Barrow and F. Tipler, *The anthropic cosmological principle*. Clarendon, Oxford, 1986.
- [414] R. Kallosh and A. Linde, “M-theory, cosmological constant and anthropic principle,” *Phys. Rev.* **D67** (2003) 023510, [hep-th/0208157](#).
- [415] A. Vilenkin, “Anthropic predictions: The case of the cosmological constant,” [astro-ph/0407586](#).

- [416] H. Martel, P. R. Shapiro, and S. Weinberg, “Likely values of the cosmological constant,” *Astrophys. J.* **492** (1998) 29, [astro-ph/9701099](#).
- [417] J. Garriga and A. Vilenkin, “On likely values of the cosmological constant,” *Phys. Rev.* **D61** (2000) 083502, [astro-ph/9908115](#).
- [418] J. E. Gunn and I. Gott, J. Richard, “On the infall of matter into cluster of galaxies and some effects on their evolution,” *Astrophys. J.* **176** (1972) 1–19.
- [419] S. Weinberg, “Living in the Multiverse,” [hep-th/0511037](#).
- [420] J. Garriga, A. Linde, and A. Vilenkin, “Dark energy equation of state and anthropic selection,” *Phys. Rev.* **D69** (2004) 063521, [hep-th/0310034](#).
- [421] L. Pogosian, A. Vilenkin, and M. Tegmark, “Anthropic predictions for vacuum energy and neutrino masses,” *JCAP* **0407** (2004) 005, [astro-ph/0404497](#).
- [422] T. Banks, M. Dine, and E. Gorbatov, “Is there a string theory landscape?,” *JHEP* **08** (2004) 058, [hep-th/0309170](#).
- [423] M. L. Graesser, S. D. H. Hsu, A. Jenkins, and M. B. Wise, “Anthropic distribution for cosmological constant and primordial density perturbations,” *Phys. Lett.* **B600** (2004) 15–21, [hep-th/0407174](#).
- [424] J. Garriga and A. Vilenkin, “Anthropic prediction for Lambda and the Q catastrophe,” [hep-th/0508005](#).
- [425] B. Feldstein, L. J. Hall, and T. Watari, “Density perturbations and the cosmological constant from inflationary landscapes,” [hep-th/0506235](#).
- [426] A. Vilenkin, “Quantum fluctuations in the new inflationary universe,” *Nucl. Phys.* **B226** (1983) 527.
- [427] A. D. Linde, D. A. Linde, and A. Mezhlumian, “From the big bang theory to the theory of a stationary universe,” *Phys. Rev.* **D49** (1994) 1783–1826, [gr-qc/9306035](#).
- [428] A. D. Linde, D. A. Linde, and A. Mezhlumian, “Do we live in the center of the world?,” *Phys. Lett.* **B345** (1995) 203–210, [hep-th/9411111](#).
- [429] A. D. Linde and A. Mezhlumian, “On regularization scheme dependence of predictions in inflationary cosmology,” *Phys. Rev.* **D53** (1996) 4267–4274, [gr-qc/9511058](#).
- [430] J. Garriga, D. Schwartz-Perlov, A. Vilenkin, and S. Winitzki, “Probabilities in the inflationary multiverse,” [hep-th/0509184](#).
- [431] M. R. Douglas, “The statistics of string / M theory vacua,” *JHEP* **05** (2003) 046, [hep-th/0303194](#).
- [432] S. Ashok and M. R. Douglas, “Counting flux vacua,” *JHEP* **01** (2004) 060, [hep-th/0307049](#).
- [433] J. Garriga and A. Vilenkin, “Solutions to the cosmological constant problems,” *Phys. Rev.* **D64** (2001) 023517, [hep-th/0011262](#).
- [434] L. Smolin, “Scientific alternatives to the anthropic principle,” [hep-th/0407213](#).
- [435] A. Aguirre, “The cold big-bang cosmology as a counter-example to several anthropic arguments,” *Phys. Rev.* **D64** (2001) 083508, [astro-ph/0106143](#).
- [436] A. Aguirre, “On making predictions in a multiverse: Conundrums, dangers, and coincidences,” [astro-ph/0506519](#).
- [437] G. ’t Hooft, “Determinism in free bosons,” *Int. J. Theor. Phys.* **42** (2003) 355–361, [hep-th/0104080](#).
- [438] G. ’t Hooft, “How Does God Play Dice? (Pre-)Determinism at the Planck Scale,” [hep-th/0104219](#).
- [439] G. ’t Hooft, “Quantum mechanics and determinism,” [hep-th/0105105](#).
- [440] G. ’t Hooft, “The mathematical basis for deterministic quantum mechanics,” [quant-ph/0604008](#).

- [441] K. Nagamine and A. Loeb, “Future evolution of nearby large-scale structure in a universe dominated by a cosmological constant,” *New Astron.* **8** (2003) 439–448, [astro-ph/0204249](#).
- [442] U. Alam, V. Sahni, and A. A. Starobinsky, “Is dark energy decaying?,” *JCAP* **0304** (2003) 002, [astro-ph/0302302](#).
- [443] R. Kallosh and A. Linde, “Dark energy and the fate of the universe,” *JCAP* **0302** (2003) 002, [astro-ph/0301087](#).
- [444] R. Kallosh, A. Linde, S. Prokushkin, and M. Shmakova, “Supergravity, dark energy and the fate of the universe,” *Phys. Rev.* **D66** (2002) 123503, [hep-th/0208156](#).
- [445] A. G. Sanchez *et al.*, “Cosmological parameters from CMB measurements and the final 2dFGRS power spectrum,” [astro-ph/0507583](#).
- [446] E. D. Reese *et al.*, “Determining the cosmic distance scale from interferometric measurements of the Sunyaev-Zel’dovich effect,” *Astrophys. J.* **581** (2002) 53–85, [astro-ph/0205350](#).
- [447] A. Dekel, D. Burstein, and S. D. M. White, “Measuring Omega,” [astro-ph/9611108](#).
- [448] N. A. Bahcall, X. Fan, and R. Cen, “Constraining omega with cluster evolution,” *Astrophys. J.* **485** (1997) L53, [astro-ph/9706018](#).
- [449] N. A. Bahcall and X. Fan, “The most massive distant clusters: determining omega and sigma 8,” *Astrophys. J.* **504** (1998) 1, [astro-ph/9803277](#).
- [450] S. D. M. White, J. F. Navarro, A. E. Evrard, and C. S. Frenk, “The baryon content of galaxy clusters: a challenge to cosmological orthodoxy,” *Nature* **366** (1993) 429–433.
- [451] M. Fukugita, C. J. Hogan, and P. J. E. Peebles, “The cosmic baryon budget,” *Astrophys. J.* **503** (1998) 518, [astro-ph/9712020](#).
- [452] J. J. Mohr, B. Mathiesen, and A. E. Evrard, “Properties of the Intracluster Medium in an Ensemble of Nearby Galaxy Clusters,” *Astrophys. J.* **517** (1999) 627, [astro-ph/9901281](#).
- [453] J. A. Peacock and S. J. Dodds, “Reconstructing the linear power spectrum of cosmological mass fluctuations,” *Mon. Not. Roy. Astron. Soc.* **267** (1994) 1020–1034, [astro-ph/9311057](#).
- [454] D. J. Eisenstein and W. Hu, “Power spectra for cold dark matter and its variants,” *Astrophys. J.* **511** (1997) 5, [astro-ph/9710252](#).
- [455] P. J. E. Peebles, “Testing General Relativity on the Scales of Cosmology,” [astro-ph/0410284](#).
- [456] P. S. Corasaniti and E. J. Copeland, “Constraining the quintessence equation of state with SNIa data and CMB peaks,” *Phys. Rev.* **D65** (2002) 043004, [astro-ph/0107378](#).
- [457] R. Bean, S. H. Hansen, and A. Melchiorri, “Constraining the dark universe,” *Nucl. Phys. Proc. Suppl.* **110** (2002) 167–172, [astro-ph/0201127](#).
- [458] P. Schuecker, R. R. Caldwell, H. Bohringer, C. A. Collins, and L. Guzzo, “Observational constraints on general relativistic energy conditions, cosmic matter density and dark energy from X-ray clusters of galaxies and type-Ia supernovae,” *Astron. Astrophys.* **402** (2003) 53–64, [astro-ph/0211480](#).
- [459] A. Melchiorri, L. Mersini-Houghton, C. J. Odman, and M. Trodden, “The state of the dark energy equation of state,” *Phys. Rev.* **D68** (2003) 043509, [astro-ph/0211522](#).