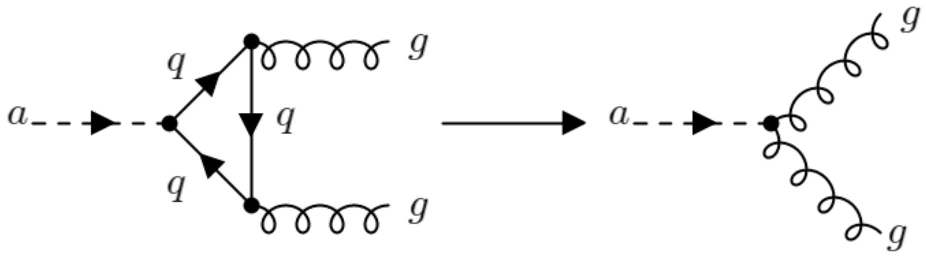
Axions as the natural answer to the dark matter question in the ΛCDM cosmological standard model

Kavermann, Michel

Axions as the natural answer to the dark matter question in the ΛCDM cosmological standard model



THESIS

submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
PHYSICS

Author :	Michel Kavermann
Student ID :	s3363511
Supervisor :	Dr. Subodh P. Patil
Second corrector :	Dr. Matthieu Schaller

Leiden, The Netherlands, February 17, 2023

Axions as the natural answer to the dark matter question in the ΛCDM cosmological standard model

Michel Kavermann
Lorentz Institute for Theoretical Physics, Leiden University
Niels Bohrweg 2, 2333 CA Leiden, The Netherlands

February 17, 2023

Abstract

The well-established ΛCDM cosmological standard model faces severe challenges, from which one is the question of what the nature of DM is. It was realized in the mid seventies that the non-trivial vacuum structure gives rise to a pseudo-Goldstone boson, the axion, that in fact is capable of solving the DM question very naturally if one finds out the actual axion mass, m_a, that is the only free parameter of the axion theory, that can only be constrained in way to still leave a vast amount of orders of magnitude to search for it. In this work, I will give a review of the fundamental theoretical insights that led to the theoretical discovery of the axion, originating from the strong-CP problem in QCD, and how it fits in the ΛCDM model, i.e. by investigating the axion as an observer field during inflation and by discussing different production mechanisms that could have led to sufficiently large axion populations in the early Universe. Afterwards, I will present the gauge-invariant linear cosmological perturbation theory, apply it to axions and in the end, briefly touch upon the non-linear regime, which is governed by the Schrödinger-Poisson equation. The work done here is fully devoted to build a strong groundwork for further investigations with numerical simulations that are part of the corresponding follow-up work.

To Dr. Jörg Meya - The inspirational man who showed me the beauty of physics and kindly asked me to always try to make an exciting story out of it.

Contents

Chapter 1

Introduction

The standard cosmological model, the ΛCDM model brings not only a well-described cosmology, supported by various observations of precision cosmology and numerical simulations, but also a good amount of open questions to explore. One of these questions is the search for the nature of dark matter. Observations show that the vast majority of the Universe's matter content is made up of dark matter, which should be cold, stable and weakly coupled, or, in other words, something that is currently not part of the standard model of particle physics. However, one fundamental part of the standard model of particle physics is the quantum chromodynamics, QCD, which faces the so-called strong-CP problem, whose solution gives rise to a new particle, the *QCD-axion*[1], that seems to fit the requirements for a proper DM candidate extraordinarily well. The original QCD-axion can be generalized in the context of GUTs, like string-theory, to a whole class of *axion-like particles*, *ALPs*. Additionally, ALPs could play an important role during inflation, seed the initial density fluctuations that grew into the large-scale structure we observe today and impacts the cosmological constant problem. ALPs even have the chance to give explanations for the existence of the matter-antimatter asymmetry. They can span a huge range of masses in theory, where we call ALPs with

$$10^{-33}\,\mathrm{eV} \lesssim m_a \lesssim 10^{-18}\,\mathrm{eV} \tag{1.1}$$

[1] Let me note, that the name axion was coined after an American detergent since it is imagined to clean up nearly all of the cosmological standard model's problems.[1]

ultra-light axions, *ULAs*, where the lower bound is of the order of the present-day Hubble constant

$$m_H = \frac{H_0}{h} \approx 100 \, \frac{\text{km}}{\text{s} \cdot \text{Mpc}} = 2.13 \cdot 10^{-33} \, \text{eV} \tag{1.2}$$

and the upper bound is related to the baryon Jeans scale to reflect the distinctive role of ULAs in structure formation. The results can be tested against well-established results of precision cosmology, namely the *Cosmic Microwave Background (CMB)*, *Large-Scale Structure (LSS)*, galaxy formation in the local universe and at high-redshift and the *epoch of reionization (EOR)*, to probe a vast range of masses to constrain the possible axion masses further. Throughout this work, we often will just use the name *axions*, but keep in mind that in fact we are always thinking of it in its most general sense as a light pseudoscalar field, what will make sense later.

This work is the first of two joint research projects, where in this work we are going to focus on building a strong framework, in which the natural existence of axions guides us to more and more fit the idea in the general, well-established ΛCDM cosmology. We split the investigations in two parts. The first part, i.e. chapter 2, is fully devoted to a motivation for and derivation of the existence of the axion as a pseudo-Goldstone boson that acquires its mass fully non-perturbative, that is an observer field during inflation and that, in general, can be produced via different production mechanisms. We will derive the fundamental properties that all ALPs have in common, consider the implications on what is going to happen if the PQ-symmetry of the axions is (un)broken during inflation and decide, based on calculations, which production mechanisms gives rise to a proper, reasonable ALP population in the early Universe. In the second part, i.e. chapter 3, we discuss the dynamics of hydrodynamical matter and scalar matter field perturbations, respectively, and even lift the latter to a quantum mechanical description that will retroactively give the initial conditions for the classical treatment. For this purpose, we will follow the recommendable review article [2] in which starting from GR, the equations are derived in a gauge-invariant fashion. We then follow the other recommendable review article [1] for the applications on the axion field, where we will arrive at the point, that the linear theory we have build up so far fails and we need non-linear theory in the form of the Schrödinger-Poisson equation, that gives a good ending point for this and an even better starting point for the follow-up research based on the here presented theoretical groundwork.

If not mentioned otherwise, we work in natural units, i.e. $c = \hbar = k_B = 1$, and try to convert everything in units of eV, $M_\odot$, pc or K. For Fourier-transformations we usually transform the coordinates x to k, where the 2π's are placed below the dk's. We use the reduced Planck mass $m_{\mathrm{pl}} \equiv M_{\mathrm{pl}} = (8\pi G)^{-1/2}$, where G is Newton's gravitational constant and finally, in flat space, we will work with the mostly positive metric signature, i.e. $\eta_{\mu\nu} = (-1, +1, +1, +1)$. New words that are defined or further explained are written in *italic* letters.

Chapter 2

Theoretical description of axions

In this chapter we are going to introduce the idea of axions as the natural consequence of QCD itself in section 2.1 by describing first the strong-CP problem and how it is solved directly afterwards. We are then in a position to distinguish different axion models, where we focus on axion models inside the QCD-axion class and use the results to give basic properties to axion-like particles in general. Then, we consider the axion in the context of inflation in section 2.2 and distinguish between the axion as a spectator field or as the field acutally driving inflation. In the end, we are interested in how cosmic axion populations could have been produced in the early Universe in section 2.3. Here, we will cover the thermal and non-thermal production in detail and take a quick glimpse into the production as the decay product of a heavier parent particle and of a toplogical defect, respectively.

2.1 QCD axions and axion-like particles

2.1.1 The strong-CP problem

In the 1970's an important problem arose in QCD, namely the *strong-CP problem*, which we would like to explore in the following. For a brief review of some important parts of QCD, most importantly the discussion of the flavour- and chiral-symmetry, respectively, I refer to appendix B.

We start by noting that the QCD-Lagrangian (B.12) is apparently invariant under axial transformations (B.15) if and only if the quark masses vanish. The corresponding Noether-current, $j_5^\mu(x)$, is then given by (B.27)

$$q(x) \rightarrow q'(x) = e^{-i\alpha\gamma^5} q(x) \tag{2.1}$$

as was shown in appendix B. For non-vanishing quark masses we now get

$$\partial_\mu j_5^\mu = 2i\bar{q}M\gamma^5 q, \tag{2.2}$$

where M is the quark mass matrix and since $M \not\equiv 0$, the current is not conserved and thus, the axial $\mathrm{U}(1)_A$ symmetry is not an approximate symmetry[3]. Adler[4], Bardeen[5], Bell and Jackiw[6] investigated this problem by considering one-loop-Feynman diagrams to analyze the divergence of (2.2), which basically connect two gluon-fields with quarks going in a triangle-loop[7]. We do not go deeper into the math here, lucky for us, Adler, Bardeen, Bell and Jackiw already did the job in the above referenced papers. The main outcome of interest for us is the chiral anomaly[1]

$$\partial_\mu j_5^\mu = -\frac{N_f g_s^2}{64\pi^2}\varepsilon_{\mu\nu\rho\sigma} G^{\mu\nu,a} G^{\rho\sigma,a} = -\frac{N_f g_s^2}{32\pi^2} G^{\mu\nu,a} \tilde{G}^a_{\mu\nu} \neq 0, \tag{2.3}$$

where we introduced the dual gluon field strength tensor $\tilde{G}$ analog to (C.24). This anomaly should introduce a new term to the QCD Lagrangian since the action is affected by

$$\delta S = \alpha \int d^4x \partial_\mu j_5^\mu = \alpha \frac{N_f g_s^2}{32\pi^2} \int d^4x G^{\mu\nu,a} \tilde{G}^a_{\mu\nu}. \tag{2.4}$$

Now, one can use the explicit form of G (B.10) to obtain

$$G^{\mu\nu,a} \tilde{G}^a_{\mu\nu} = \partial_\mu \left(\varepsilon^{\mu\nu\rho\sigma} A^a_\nu \left[F^a_{\rho\sigma} - \frac{g_s}{3} f^{abc} A^b_\rho A^c_\sigma \right] \right), \tag{2.5}$$

so that the integral in (2.4) seems to be just a surface integral, for which we only need a proper boundary condition for it to vanish, so that we recover the initial $\mathrm{U}(1)_A$-symmetry. One could easily say, that $A^{\mu,a} = 0$, i.e. vacuum, at spatial infinity and the integral vanishes as desired[7]. t'Hooft[8] studied this anomaly as well and discovered that the usually used vacuum in QCD is more complex than initially assumed. In fact, one can gauge-rotate the vacuum to reach a new vacuum state in which the anomaly (2.3) reappears[8]. Hence, $\mathrm{U}(1)_A$ is no true symmetry of QCD and thus there is

[1] An *anomaly* describes the phenomenon that a classical symmetry is broken at quantum level[1].

no pseudo Goldstone boson coming out of the symmetry breaking[3]. The actual vacuum state we need is the so-called *θ-vacuum*, which is discussed in length in appendix C, especially in C.2 and C.4. A major result of using the proper vacua for this problem is that the QCD-action now gets an additional term (C.3) corresponding to the Lagrangian

$$\mathcal{L} = \frac{\theta}{32\pi^2} \int d^4x Tr \left[G_{\mu\nu} \tilde{G}^{\mu\nu} \right], \tag{2.6}$$

in which the name-giving angle θ enters[7]. This term violates CP symmetry[2] and gives rise to an electric dipole moment, which is constrained to be[13]

$$\theta \cdot 10^{-16} \cdot e\,\mathrm{cm} \approx |d_n| \lesssim 3 \cdot 10^{-26} e\,\mathrm{cm}, \tag{2.7}$$

what can be rearranged to give

$$\theta \lesssim 3 \cdot 10^{-10}, \tag{2.8}$$

leaving us with a true fine-tuning problem. Note, that the introduction of a symmetry-breaking Lagrangian due to the anomaly gives rise to the picture that anomalies are explicit symmetry breaking effects. Note additionally, that the in general non-vanishing quark mass matrix complicates the situation even further because in the electroweak theory, this matrix is complex in general. To discuss physical processes, we thus need to diagonalize it first, which itself is not the big deal, but the used chiral transformation[3] to achieve this introduces an additional term

$$\theta_{\mathrm{QCD}} = \theta + \arg[\det[M]]. \tag{2.9}$$

We now need to argue, why the total angle θ_{QCD} is extremely small (2.8), so that the symmetry breaking term (2.6) basically vanishes even though there is nothing that hinders θ_{QCD} from being of order unity, in fact it could be anywhere in the interval $[0, 2\pi]$, and additionally there is no argument that forces the second term of the total angle to cancel the first term. All these problems together is what we refer to as the *strong-CP problem*.

[2]In early papers by Weinberg[9] and Wilczek[10]in 1978 they typically refer to P and T conservation, but Peccei and Quinn[11] in 1977 already adapted to the CPT-theorem and simply speak of CP conservation, since this is equivalent to T conservation[12]. We stick to the latter notation as it comes in handy and to be consistent with literature.

[3]t'Hooft computed the transition amplitude between the same vacuum states in Euclidean spacetime, i.e. with the path integral formalism we used in appendix C as well and found that the exponent of the exponential, i.e the considered Lagrangian, is proportional to certain det$[M]$-terms[8], which is why the additional angle is of the form (2.9).

2.1.2 The natural solution of the strong-CP problem

Let us now try to solve the strong-CP problem that we introduced in the previous subsection. One can think of different ways to do so, but the most natural path is by thinking of an additional chiral symmetry that must be present because it would be able to simply rotate the problematic θ-vacua away. Such a symmetry can either be achieved as we discussed in appendix B via a massless up-quark for instance, which seems rather silly since observational data clearly shows non-vanishing quark-masses in general, or via imposing a new global U(1)-symmetry on the whole standard model Lagrangian[7]. The latter one was investigated by Peccei and Quinn[11], who picked up the work of t'Hooft, who originally discussed "Euclidean-gauge solitons"[8], which correspond to the tunneling effect between different QCD-vacua[4]. He found that the amplitude for the tunneling between different vacua vanishes whereas between the same vacuum states gives rise to terms proportional to the determinant of the quark mass matrix as (2.9). The basic assumption of Peccei and Quinn is that at least one fermion of the theory acquires its mass through Yukawa-coupling[5], so that the corresponding vacuum expectation value is nonzero. Then the original Lagrangian should posses a $U(1)_{PQ}$-symmetry[6]. The idea is the following. The strong-CP problem asks us to minimize the value of $\bar{\theta}$ in order to satisfy the constraint (2.8) or even better to set $\bar{\theta} = 0$. A chiral rotation of the form (B.15), i.e. a U(1)-rotation, adds an additional term to the effective Lagrangian as we discussed in the previous subsection whilst setting up the strong-CP problem in the first place. But unlike the original rotation that gave rise to the problem, the new rotation is supposed to fix the problem. The requirement of at least one Yukawa-interaction states that we have at least one (pseudo-)scalar field ϕ in our theory,

$$\mathcal{L} \subset \bar{\psi} \left[g_Y \phi \frac{1}{2}(1+\gamma^5) + g_Y^* \phi^* \frac{1}{2}(1-\gamma^5) \right] \psi, \tag{2.10}$$

where g_Y is the Yukawa-coupling-constant, with nonzero expectation value $\langle \phi \rangle = \lambda e^{i\beta}$. The task is now to minimize the potential $V(\phi)$, which apparently depends on the field itself but also on other parameters like the angle

[4]Later we see that t'Hooft was already describing the so-called instantons.

[5]Recall, that we speak of couplings between two fermions and one scalar as *Yukawa-couplings* or *Yukawa-interactions*. Instead of the scalar one can consider a pseudoscalar by multiplying the scalar field with γ^5 as usual[3].

[6]The subscript PQ refers to Peccei-Quinn as we will later on always speak of the *PQ-symmetry*. Of course, in their original work they just spoke of a U(1)-symmetry.

$\bar{\theta} = \theta + \beta$, for which one finds $\beta = 0$ to be the minimum of $V(\phi)$. The corresponding fermion mass term then reads

$$\lambda \bar{\psi} \left[g_Y e^{i\beta} \frac{1}{2}(1+\gamma^5) + g_Y^* e^{-i\beta} \frac{1}{2}(1-\gamma^5) \right] \psi. \tag{2.11}$$

In order to make this mass term real again, we simply perform the new $\mathrm{U}(1)_{PQ}$-rotation $\exp\{i\gamma^5\theta\}$ for $\beta = -\theta$, what then gives

$$\bar{\theta} = \theta + \beta = \theta - \theta = 0 \tag{2.12}$$

as the total angle appearing in the Lagrangian (2.9). For each fermion configuration with different β_i it is possible to redo the calculation. The condition $\beta = -\theta$ then alters to

$$\arg \left[\prod_i (g_{Y,i} \exp\{i\beta_i\}) \exp\{i\theta\} \right] = 0, \tag{2.13}$$

so that each fermion mass needs to be made real by a $\mathrm{U}(1)_{PQ}$-transformation, which then results in

$$\sum_i \beta_i = -\theta, \tag{2.14}$$

so that the U(1)-symmetry dynamically sets $\bar{\theta}$ to zero and hence, solves the strong-CP problem and restores CP-invariance[11]. From now on, we will denote by *PQ-symmetry* the $\mathrm{U}(1)_{PQ}$-symmetry in honor of this breakthrough and to be consistent with the literature. One should highlight at this point, that the PQ-solution of the strong-CP problem is in fact a very natural feature of QCD since the only assumption made is, that at least one fermion is acquiring mass through Yukawa-coupling which originally was designed to describe strong interactions. Note, that above, we found that ϕ actually is a pseudoparticle due to the γ^5 factor. Shortly after presenting the PQ-solution, Peccei and Quinn extended their discussion to the inclusion of electroweak interactions and found that the solution is still valid and a natural feature of QCD[14].

Now, that we have seen how to solve the strong-CP problem, we would like to go back to what has started the necessity of this discussion in first place, namely the *instantons* that correspond to the transition between θ-vacuum states, first described by t'Hooft[8]. We have discussed the most important properties of θ-vacua and instantons in appendix C. However, the transition between vacua is a non-perturbative effect, which is crucial for instanton effects as we will see later. The idea is that since we already saw in the PQ-solution, there is a new pseudoscalar field ϕ that dynamically sets $\theta_{\text{QCD}} = 0$, i.e. it couples to the $G\tilde{G}$-term in (2.6), so one can simply take

$$\theta_{\text{QCD}} = C\frac{\phi}{f_a}, \tag{2.15}$$

where the constant of proportionality, C, describes the color anomaly, that we describe late, ϕ is the canonically normalized *axion* field and f_a is the *axion decay constant*. Finally, we arrived at the axion entering our theory. Note, that shortly after Pecceis and Quinns publication, Weinberg[9] and Wilczek[10] already theorized the axion-implication of the PQ-solution. Since QCD naturally introduces the axion to us, we would like to call it the *QCD-axion* to distinguish it from other axions we discuss in the next subsection.

In order to keep the property of setting $\theta_{\text{QCD}} = 0$ dynamically we demand that ϕ has a shift-symmetry, $\phi \to \phi + \text{const.}$, and only derivatives of the axion field appear in the action. The shift-symmetry is a crucial property of the instanton, which is protected to all orders in perturbation theory since instanton-effects are purely non-perturbative and all possible quantum corrections will be suppressed by powers of f_a[1]. This ensures that contributions to θ_{QCD} can be absorbed by the axion field via the shift-symmetry, so that the action and the potential induced by the instanton-effects solely depend on the overall axion field. Let us make this clear by an example. The vacuum energy E_{vac} depends on θ_{QCD}

$$E_{\text{vac}} \sim \cos\theta_{\text{QCD}} \sim \theta_{\text{QCD}}^2, \tag{2.16}$$

which was shown by t'Hooft[8] or see (C.5) and the derivation of that property in appendix C.4. Since the θ-vacua are topologically distinct, the transition between different vacua is forbidden[8], which leads to the so-called *superselection rule* for θ-vacua

$$\langle\theta|\text{anything}|\theta'\rangle = \delta_{\theta\theta'} \tag{2.17}$$

and the fact that such a process cannot minimize the vacuum energy. After introducing the axion field, the vacuum energy is now

$$E_{\text{vac}}\left(\theta_{\text{QCD}} + N_{\text{DW}}\frac{\phi}{f_a}\right), \tag{2.18}$$

but due to the shift-symmetry we can absorb θ_{QCD} in the field to get

$$E_{\text{vac}}\left(N_{\text{DW}}\frac{\phi}{f_a}\right) \tag{2.19}$$

and since ϕ is a dynamical field, the vacuum energy can be minimized[1].

That the shift-symmetry is protected from quantum corrections can be rewritten as the fact that quantum effects break violate the classical symmetry, which is just the definition of an anomaly. Say, we call Q_{PQ} the PQ-charge, so that a PQ-rotation is given by

$$x_i \to x_i' = \exp\left\{iQ_{PQ,i}\frac{\phi}{f_a}\right\}x_i, \tag{2.20}$$

where x_i is a field with PQ-charge $Q_{PQ,i}$, then the *color anomaly*, C, is given by

$$C\delta_{ab} = 2Tr\left[Q_{PQ}T_aT_b\right], \tag{2.21}$$

where the trace goes over all fermions in the theory and $T_{a,b}$ are the generators of the SU(3)-representation of the fermions (see (B.7) and the text thereafter). The color anomaly sets the number of vacua that ϕ has in the range $[0, 2\pi f_a]$ and according to the shift-symmetry of ϕ we get $\phi \to \phi + 2\pi f_a$ and since ϕ is an angular variable we have $C \in \mathbb{Z}$[1]. Srednicki[15] showed that this is in fact always achievable. Due to this property of the color anomaly we rename it for later purposes already to being the *domain wall number*[1], $N_{DW} \equiv C$.

From [11] we already know that there are interaction terms between the axion- and quark-fields, so that we are able to compute a mass for the axion, m_a, since after QCD-confinement at $T \sim \Lambda_{\text{QCD}}$, we can effectively replace the $\bar{q}q$-terms by their vacuum expectation values $\langle\bar{q}q\rangle$. By simply assuming f_a to be large, so that m_a is going to be small, we can consider only up- and down-quarks and note, that under this assumption

$$\cos\left(N_{\text{DW}}\frac{\phi}{f_a}\right) \sim \frac{(m_u+m_d)\langle\bar{q}q\rangle}{f_a^2} = m_\pi^2\frac{f_\pi^2}{f_a^2}, \tag{2.22}$$

where m_π is the pion mass and f_π is the pion decay constant, holds. In the end we get the renormalized axion mass

$$m_a = \frac{m_\pi^2 f_\pi^2}{\left(\frac{f_a}{N_{\text{DW}}}\right)^2}\frac{m_u m_d}{(m_u+m_d)^2}\left[1+\frac{m_\pi^2}{m_\eta^2}\left[-1+\mathcal{O}\left(1-\frac{m_\pi}{m_\eta}\right)\right]\right]. \tag{2.23}$$

Note, that if $m_\pi = m_\eta$ the quantum effects cancel and the axion would be massless, which is why one is led to say that the instantons, whose non-perturbative effects achieve $m_\pi < m_\eta$, give mass to the axions for $T < \Lambda_{\text{QCD}}$[1]. Wilczek mentioned that the axion seems to acquire mass through processes in which two instantons interact with each other, where one instanton splits in a pair of left- and right-handed up-quarks whilst the other instanton splits in a pair of left- and right-handed down-quarks, so that the four quarks can effectively interact at a four fermion-vertex[10]. By plugging in numbers for the known meson masses and f_π we get to first order[1]

$$m_{\text{a,QCD}} \approx 6 \cdot 10^{-6}\,\text{eV} \cdot \left(\frac{10^{12}\,\text{GeV}}{\frac{f_a}{N_{\text{DW}}}} \right). \tag{2.24}$$

Further, for $T < \Lambda_{\text{QCD}}$ these instanton effects break the shift symmetry of the axion explicitly to a discrete symmetry,

$$\phi \to \phi + 2\pi \frac{f_a}{N_{\text{DW}}}, \tag{2.25}$$

which agrees to our earlier findings that the color anomaly is an integer. This symmetry breaking implies the QCD-axion potential[7]

$$V(\phi) = m_u \Lambda_{\text{QCD}}^3 \left[1 - \cos\left(N_{\text{DW}} \frac{\phi}{f_a} \right) \right] \tag{2.26}$$

induced by the instantons[1], where the cosine potential comes from E_{vac}, see appendix C.4, and is already shifted, so that the potential is minimized at $\phi = 0$ as we need to solve the strong-CP problem. If one considers SU(2) in the electroweak theory (short: EWT), one finds electroweak instantons as well since the weak force breaks the CP-symmetry and the electroweak instantons also lead to a shift of the minimum of the axion potential, which has to be corrected for by the PQ-symmetry[14]. As we have shown in detail in appendix C, the instanton action for a gauge group G with coupling constant g_i is (C.4)

$$S_{\text{inst.}} = \frac{8\pi^2}{g_i^2}, \tag{2.27}$$

with which one can set the prefactor of the axion potential since (C.5)

$$V_i(\theta) \sim \cos(\theta) \exp\{-S_{\text{inst.}}(g_i)\} \tag{2.28}$$

[7]Note, that for a generic ALP the potential reads $V(\phi) = \Lambda_a^4 \left[1 \pm \cos\left(N_{\text{DW}} \frac{\phi}{f_a} \right) \right]$, where it is typical to choose the minus sign in order to set $\phi = 0$ as the potential minimum[1].

holds. By comparing the coupling constants of QCD and the EWT, one finds immediately that, for instance, the W-boson potential only weakly breaks CP-invariance compared to QCD, so we can safely neglect the EW-effects for our considerations[16]. For further discussions of the implications of instantons I like to refer to the early papers of Peccei and Quinn[17] and t'Hooft[18] that reviewed this topic in detail. For the general notion of instantons and their basic properties that are relevant for this work, I refer to appendix C.

Let us highlight the link (2.24) between the axion-mass, m_a, and the axion decay constant, f_a. If f_a is sufficiently large, the axion mass is sufficiently small or, in other words, the QCD-axion is extremely light and stable. Thus, the QCD-axion fulfills the first properties we discussed in appendix A in order to be an appropriate DM candidate. Note, that the axion comes out naturally of QCD with the possibility of having suitable properties. So even though the standard model of particle physics is not offering us the correct DM particle directly, there is a chance that it is just hidden under the cover of QCD.

2.1.3 QCD-axion models and ALPs in general

The QCD-axion we introduced in the previous subsection can be described in different ways. In fact, string theory teaches us that there is a whole class of particles, *Axion-Like Particles* or *ALPs* for short, that fulfill the requirements to be called an axion. Through different theoretical descriptions one gets different implications of the behavior of the considered ALP, which we will discuss in the subsequent sections. Thus, a careful distinguishement between the most-relevant ALPs is necessary. Now, we would like to describe three differen QCD-axions. We start by describing the PQWW-axion as the natural continuation of the historical story we told in the previous subsection, even though it will turn out that the PQWW-axion is ruled out by experiment. We then continue with KSVZ- and the DFSZ-axion, which will be important throughout the rest of this work. If not stated otherwise, I will follow closely the presentation of Marsh[1].

Let us start with the **P**eccei-**Q**uinn-**W**einberg-**W**ilczek-axion. A single additional complex scalar field φ is introduced to the standard model as a second Higgs doublet, so that one Higgs field gives mass to the u-type quarks, namely the up-, charm- and top-quark, whereas the second Higgs field gives mass to the d-type quarks, namely the down-, strange- and bottom-quark. This fixes the representation of φ in the SU(2)⊗U(1) symmetry group and the whole Lagrangian is then taken to be invariant under a global PQ-symmetry, which shift the angular part of φ by a constant, as we discussed in the previous subsection. The PQ-field couples, as discussed, via Yukawa-interactions to the standard model particles to give mass to the fermions as in the usual Higgs mechanism with the potential

$$V(\varphi) = \lambda \left(|\varphi|^2 - \frac{f_a^2}{2} \right)^2 . \tag{2.29}$$

The potential takes the vacuum expectation value $\langle \varphi \rangle = f_a / \sqrt{2}$ at the EW phase transition, which fixes the symmetry breaking scale $f_a \approx 250\,\text{GeV}$. After the symmetry breaking, there are four real, electromagnetic (short: EM) neutral scalars left, i.e. one that gives mass to the Z-boson, one is the standard model Higgs boson, one heavy radial φ field and one angular φ field. The angular degree of freedom appears as $\langle \varphi \rangle e^{i\phi/f_a}$ after canonically normalizing the kinetic term and ϕ manifests itself as the Goldstone boson of the spontaneously broken PQ-symmetry. Due to the global PQ-invariance, the PQ-charges of the fermions are fixed. By expanding in powers of f_a^{-1} one get quark couplings of the form

$$m_q \frac{\phi}{f_a} i\bar{q}\gamma^5 q \tag{2.30}$$

and the chiral anomaly induces couplings to the gauge bosons via fermion loops, which is a result of effective field theory since the fermion loops are integrated out for low-energy processes. We will come back to this with somewhat more detail at the KSVZ-axion. However, the couplings to the gauge bosons are of the form $\frac{\phi}{f_a} G\tilde{G}$ for the strong interaction and likewise $\frac{\phi}{f_a} F\tilde{F}$ for the EM interaction. Obviously, we seek for the gluon-term in order to solve the strong-CP problem. Note, that all axion couplings are suppressed by the symmetry-breaking scale f_a, which in the PQWW-model is fixed to the EW vacuum expectation value, which is too small, i.e. the couplings are too large and the PQWW-axion is thus excluded by experiment, see for instance the constraints on the couplings coming from collider experiments presented in[19].

Let us now turn to the **K**im-**S**hifman-**V**ainshtein-**Z**akharov-axion. A heavy quark doublet, Q_L, Q_R ∈SU(3), is introduced. The PQ scalar field field φ has charge 2 under chiral rotations and is now a standard model singlet, which interacts with heavy quarks via the PQ-invariant Yukawa-Lagrangian

$$\mathcal{L}_Y = -\lambda_Q \varphi \bar{Q}_L Q_R + \text{h.c.}, \tag{2.31}$$

which provides the quark masses, where λ_Q is a free parameter. Now, as in the PQWW-model, there is a global PQ-symmetry, which is spontaneously broken and produces the potential (2.29). At the classical level, the Lagrangian is still unaffected by the chiral rotations and φ is thus, not yet coupled to the standard model. However, at the quantum level, chiral rotations on Q affect the gluon-term via the chiral anomaly, in which the KSVZ model sets $N_{DW} = 1$. At low energies, after PQ-symmetry breaking, φ can be replaced by its vacuum expectation value and the Q-fields obtain a large mass

$$m_Q \sim \lambda_Q f_a, \tag{2.32}$$

since f_a is thought to be very large. Let us now come back shortly to the basic idea of effective field theory, that at low energies q we can replace a fundamental by an effective action by integrating out fields with masses bigger than the considered energy, $m > q$. A typical example is the muon decay in the EWT. The action for this process contains a term

$$i g_2 W_\mu \bar{l}_i \gamma^\mu \nu_i + \text{h.c.}, \tag{2.33}$$

where g_2 is the EW coupling, l_i is the charged lepton field, ν_i are the neutrinos and W_μ is a charged W-boson with $m_W = 80.4\,\text{GeV}$. At small momentum transfer, we have $q^2 \ll m_W^2$ and the original W-boson propagator, which is proportional to $(q^2 + m_W^2)^{-1}$, can be replaced by an effective four-fermion interaction, which is proportional g_2^2/m_W^2. Hence, after introducing the Fermi interaction constant

$$G_F = \frac{\sqrt{2}}{8}\frac{g_2^2}{m_W^2} \tag{2.34}$$

one obtains the effective action, that now contains a term

$$G_F(\bar{e}\nu_e)(\bar{\nu}_\mu \mu) + \text{h.c}, \tag{2.35}$$

that describes the mentioned four-fermion interaction. Likewise we can consider a KSVZ axion that acts with two gluons, see the left-hand side of figure 2.1. The virtual quarks then induce an effective action between the axion field ϕ and the two gluon fields g at loop-level, so that at low momentum-transfer, i.e. for $q^2 \ll m_Q^2$, the heavy quarks can be integrated out and the effective action now contains a term

$$\frac{\phi}{32\pi^2 f_a} G\tilde{G}. \tag{2.36}$$

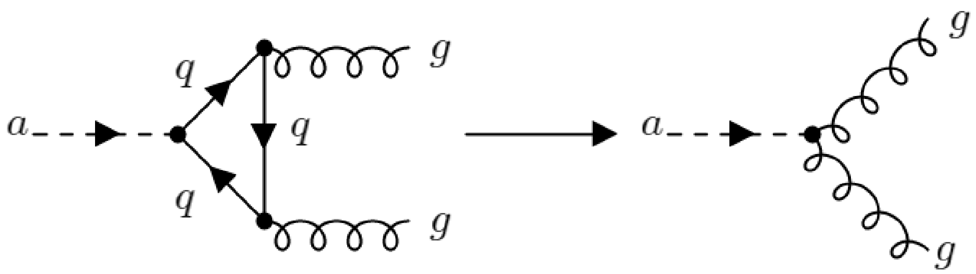


Figure 2.1: *Left:Fundamental interaction between an axion, a with two gluons, g, via a massive quark-loop, where from the fundamental Lagrangian one can use the vertices aqq and gqq. We assume that all quarks in the loop have the momentum p^{μ}. Right: Effective interaction between an axion, a with two gluons, g. The quark-loop was integrated out, so that one is left with an effective interaction vertex agg in the effective Lagrangian. Note, that the vertex is giving a factor f_a^{-1}[1].*

Note, that for the KSVZ-axion we have already set $N_{\text{DW}} = 1$. Hence, the effective coupling, see the right-hand side of fig.2.1, is f_a^{-1}. Back to our description of the KSVZ-axion, we note that the induced topological term, i.e. the interaction between the gluons and the axion, is the only additional term to the standard model Lagrangian and that there are no unsuppresed tree-level[8] couplings to standard model matter fields. There is an axion-photon coupling that gives an EM anomaly that depends on the EM-charges of the quark-fields. Later we will see how other couplings can be induced by loops and mixing.

Next, we would like to describe the **D**ine-**F**ischler-**S**rednicki-**Z**hitnitsky-axion. It couples to the standard model via the Higgs sector and contains two Higgs doublets, H_u, H_d, just like in the PQWW-model, but φ is, unlike in the PQWW-model, a standard model singlet. Again, we impose a global PQ-symmetry, which is spontaneously broken, so that we get the potential (2.29). Now, the PQ- and the Higgs-fields interact via the scalar potential

$$V = \lambda_H \varphi^2 H_u H_d, \tag{2.37}$$

which is PQ-invariant if φ has PQ-charge $+1$ and the Higgs fields carry PQ-charge -1 each. After PQ-symmetry breaking, φ can again be replaced by its vacuum expectation value, but now the parameters in (2.29) and λ_H

[8]Recall, that connected Feynman graphs without loops are called *tree-graphs* and the corresponding order of perturbation theory is called the *tree-level*[20].

must be chosen so that for the Higgs field remains consistent with the observed standard model mass of $m_H = 125\,\text{GeV}$ and the EW vacuum expectation value

$$v_{\text{EW}} = \sqrt{\langle H_u \rangle^2 + \langle H_d \rangle^2}. \tag{2.38}$$

The Higgs field must also couple to all the standard model fermions to provide their masses through Yukawa couplings and in order for these couplings to be PQ-invariant, the standard model fermions must carry proper PQ-charges. After symmetry breaking also the Higgs fields are replaced by their vacuum expectation values, which induce axial current couplings between the axion and the standard model fermions from well-known terms like (2.30). These currents induce the coupling between the axion and the gluons again via the color anomaly. Unlike in the KSVZ-model, all standard model quarks carry PQ-charges, so that the color anomaly is $N_{\text{DW}} = 6$. Following the same steps as in the KSVZ-model we can consider processes as the one in figure 2.1, but unlike in the KSVZ-model, where we integrated out heavy quark fields, we now integrate out light quark fields giving the same interaction as before. Note, that unlike in the KSVZ-model, in the DFSZ model we find tree-level couplings between the axion and the standard model fermions.

As stated in the beginning of this subsection, there are many more descriptions of ALPs. All of them have several properties in common, which we like to summarize now. First, the classical action has a global PQ-symmetry. Second, the spontaneous symmetry breaking scale f_a leads to an angular degree of freedom, ϕ/f_a, that contains a shift-symmetry. Third, the PQ-symmetry is anomalous and thus, explicitly broken at quantum quantum effects by non-perturbative instanton-effects that protect the classical shift-symmetry. Fourth, the protected shift-symmetry, $\phi \to \phi + 2n\pi f_a$ with $n \in \mathbb{Z}$, manifests the axion as a pseudo-Golstone-boson, that obtains a periodic potential $V(\phi/f_a)$, when the non-perturbative quantum effects switch on at some scale Λ_a. The same effects induce the axion mass, which is proportional to Λ_a^2/f_a.

Let us briefly discuss the couplings of ALPs to the standard model. First of all, we define the QCD-axion to be coupled to $G\tilde{G}$ with coupling strength unity due to the replacement (2.15). Since the axion is a pseudo-Goldstone boson, the coupling to fermions must be of the form (B.27) and since is has a shift-symmetry, only derivatives of the axion field are allowed in the coupling, so in order to couple to fermions, we get terms of the form

$$\partial_\mu \left(\frac{\phi}{f_a}\right) (\bar{\psi}\gamma^\mu\gamma^5\psi). \tag{2.39}$$

An important consequence of these couplings is that the mediated force is spin-dependent and thus, the axion does not mediate long-range scalar forces between macroscopic objects[1]. One can show, that for an EM anomaly, there is also a coupling to EM

$$\phi\vec{E}\cdot\vec{B} = -\frac{\phi}{4}F_{\mu\nu}\tilde{F}^{\mu\nu}, \tag{2.40}$$

what we do not want to derive in more detail[1]. However, with these couplings at hand, the general interaction Lagrangian reads

$$\begin{aligned}\mathcal{L}_{\text{int}} = -\frac{g_{\phi\gamma\gamma}}{4}\phi F_{\mu\nu}\tilde{F}^{\mu\nu} + \frac{g_{\phi NN}}{2m_N}\partial_\mu\phi(\bar{N}\gamma^\mu\gamma^5 N) + \frac{g_{\phi ee}}{2m_e}\partial_\mu\phi(\bar{e}\gamma^\mu\gamma^5 e) \\ \frac{g_{\phi NN\gamma}}{4}\phi(\bar{N}[\gamma^\mu,\gamma^\nu]\gamma^5 N)F^{\mu\nu},\end{aligned} \tag{2.41}$$

where N is a nucleon. A dimensional analysis shows that all four coupling constants in the above interaction Lagrangian are either dimensionful or dimensionless but related to dimensionful constants, so that in the end all of them are suppressed by f_a^{-1}[1], what explains the weak coupling of axions to standard model particles, since we assume f_a to be large.

2.2 Axions in inflation theory

As we saw previously, for the existence of the axion, the breaking of the PQ-symmetry is crucial. Since, we assume that ALPs are generated in the early Universe in order to properly account for structure formation, what we discuss in more detail in the next chapter, we are forced to consider the axion in the context of inflation, which is theorized to take part in the radiation-dominated epoch, and thus, in the early Universe. One distinguishes between the axion field that is just being influenced by inflation and the axion field that is actually driving inflation. We now want to briefly discuss both cases. For a brief review of basic inflation theory, I refer to appendix E and to the references therein for way more details.

First of all, we start with the Universe's temperature during inflation, which is set by the *Gibbson-Hawking temperature*

$$T_I = \frac{H_I}{2\pi}, \tag{2.42}$$

where H_I is the inflationary Hubble scale[1]. The basic idea of Gibbson and Hawking goes as follows[21]. One can fully describe a black hole (short: BH) by its mass, M, angular momentum, J, and charge, Q. For given parameter set (M, J, Q) a certain BH can have an infinite amount of internal configurations that represent the different possible initial conditions the body that collapsed to the BH fulfilled because from a classical point of view, the collapsed body could have been made out of an infinite amount of particles with arbitrary small mass. Seen from quantum mechanics, in order to have a gravitational collapse, the energies of the particles would have been restricted by the requirement that their wavelengths have to be smaller than the size of the black hole, so it seems reasonable to say that the number of configurations have to be finite. This in turn means that one can define an entropy $S \sim \log(\text{number of configurations})$ for the BH and this then leads naturally to an associated temperature $T = G^2 \left[\left(\frac{\partial S}{\partial M} \right)_{J,Q} \right]^{-1}$ to the BH, which gives rise to thermal radiation[21]. Gibbson and Hawking discuss this effect then for various cases, eventually leading to (2.42) as the radiation emitted from the de-Sitter horizon[1]. See [21] for details. However, one can map the inflationary Hubble scale, H_I, to measurable parameters[1]

$$\frac{H_I}{2\pi} = M_{\text{pl}} \sqrt{\frac{A_s r_T}{8}}, \tag{2.43}$$

where A_s is the scalar amplitude and r_T is the tensor-to-scalar ratio[9]. Since $r_T < 0.032$, see [22], we know that cosmological fluctuations are dominantly scalar and since $\sqrt{A_s} \sim 10^{-5}$ we can assume that they are adiabatic. We will dig deeper in the evolution of cosmological perturbations in the next chapter. However, this little groundwork was necessary, so that we can distinguish between $f_a < T_I$, for which the PQ-symmetry is unbroken during inflation, and $f_a > T_I$, for which the PQ-symmetry is broken during inflation[1]. Let us discuss both cases briefly in the following.

Let us start with the unbroken, $f_a < T_I$, case. This basically means that during inflation, when the PQ-symmetry is not broken, the vacuum ex-

[9] Both, the scalar amplitude, A_s, and the scalar spectral index, n_s, are parameters resulting from power spectra measurements. Their corresponding tensor pendants, A_t and n_t are set by inflation in order to fit the observations. The tensor-to-scalar ratio is then, as the name suggests, defined as $r_T := A_t / A_s$ and strongly constrained by CMB measurements as described in [22] for instance. They show that the current, very small, upper limit $r_T < 0.032$ can be achieved at a 95% confidence level. See [22] for details.

pectation value of ϕ is zero and it is only after inflation, when T_I drops below f_a that the symmetry breaks. Consider the Universe to be made out of several patches, then patches that are not in causal contact with each other, will pick a random $\theta = \phi/f_a$ value. There is no preferred θ at this point because f_a is larger than the scale of non-perturbative effects and it requires f_a to drop below that scale so that these effects can switch on in order to employ a potential on the axion field, so that it can acquire mass. Since this is not the case, there is no potential and thus, $[-\pi, +\pi]$ is a uniform distribution from which one draws θ randomly. Thus, the average

$$\langle \theta_i^2 \rangle = \pi^2/3 \tag{2.44}$$

is fixed[1]. The subscript i just denotes the different patches we mentioned above. We will see later in subsection 2.3.2 that one can compute a relic axion density out of this phenomenon for QCD-axions[10]. Additionally, the global symmetry breaking we just mentioned gives rise to topological defects, that are the initial condition for the production mechanism briefly discussed in subsection 2.3.3.

Now, let us turn to the broken, $f_a > T_I$, case[11]. Basically, we can start with the same processes, namely that the PQ symmetry breaks and gives several patches with different θ-values. The crucial difference to the previous case is now, that since this symmetry breaking happens during the era of inflation, each patch is expanded rapidly, so that in the end all patches are in causal contact an thus, we have one single value of $\theta \in [-\pi, +\pi]$. Since θ can be any value in this interval, we can regard θ as a free parameter unlike in the unbroken case. Even though we will discuss cosmological perturbations later, let me add, that this case also sources inflationary isocurvature fluctuations, $\delta\phi = T_I$, which react with the initial field displacement, ϕ_i, what sets a minimum value to the otherwise free θ-parameter[1] due to

$$\langle \phi_i^2 \rangle = \phi_i^2 + \langle \delta\phi^2 \rangle = f_a^2 \theta_i^2 + T_I^2 \overset{(2.42)}{=} f_a^2 \theta_i^2 + \left(\frac{H_I}{2\pi} \right)^2 , \tag{2.45}$$

[10] We explicitly refer to QCD-axions since they typically have f_a values of order 10^{12} GeV, which fit to the bound set by T_I[1]. The so called ADMX-experiment[23] was designed specifically for QCD-axions in this mass range. The recent results[24] show that at 90% confidence level, QCD-axions in the mass range of $(65.5 - 69.3)$ μeV can be excluded.

[11] Note, that this, unlike the unbroken case, now allows for very high values of $f_a \sim 10^{16}$ GeV in order to describe general ALPs. This will play an important role later in subsection 2.3.2.

what will become clearer later.

Let us now turn to the other case of the axion field driving inflation. In appendix E the slow-roll parameters of slow-roll inflation were defined that basically say that the inflation potential must be very close to being flat and very shallow. Recall that the axions have a shift-symmetry that protects their potential from all quantum corrections, so the properties of the potential should be preserved. Hence, one could consider the axion field as the ϕ field from appendix E that drives inflation, but since this field is supposed to decay at a certain point, we can exclude the QCD-axion since we have theorized it to be stable, or at least long-lived compared to the age of the Universe. In fact, if axions are supposed to compose DM, they cannot drive inflation simultaneously[1]. There are several models that try to explain inflation using the axion field. Let us very quickly go over them in the way Marsh[1] did as we will focus on ALPs as DM candidates. Please refer to the references therein for detailed discussions of the specific models mentioned.

The simplest model of axion-driven inflation is the so-called *natural inflation*, where the usual axion potential with $N_{\mathrm{DW}} = 1$,

$$V(\phi) = \Lambda_a^4 \left[1 \pm \cos\left(\frac{\phi}{f_a}\right)\right], \tag{2.46}$$

is used. Originally, $\Lambda_a \sim m_{\mathrm{GUT}}$ and $f_a \sim M_{\mathrm{pl}}$ are used, what is exactly what Lyth[25] required for inflation to produce observable values of r_T, namely $r_T \gtrsim 10^{-2}$. In the context of quantum gravity, one can argue that $f_a \gtrsim M_{\mathrm{pl}}$ is technically forbidden, so that one has to enter the string-theory sector with $f_a < M_{\mathrm{pl}}$. Typically, to this bound corresponding models, like *axion monodromy*, do not produce power-law initial power spectra, so one considers *assisted inflation* that indeed do so. In this model one considers a collection of coupled fields. The friction between them due to Hubble expansion adds an additional damping term in the equations that effectively slows the collective motion of the fields down in order to satisfy the slow-roll conditions. In the context of axion fields one speaks of *N-flation* with N identical potentials of the form (2.46). In addition to such a collective behavior the *Kim-Nilles-Peloso* model allows for rotation between the fields. Even more models were reviewed in [26].

2.3 Production mechanisms

Now that we argued for the existence of the axion, let us turn to the question how cosmic populations of axions are produced. One can think of initially producing these axion populations in different ways, from which we like to discuss the thermal production (see subsection 2.3.1) and the production via misalignment (see subsection 2.3.2) in detail in the following and just briefly think of the axion as the decay product of a heavier parent particle or as the decay product of a toplogical defect (for both see subsection 2.3.3).

2.3.1 Thermal production

Let us assume that in the early Universe, axions are in thermal contact with the standard model radiation, so by the same mechanisms to produce relic neutrino or WIMP[12] abundances. A generic ALP[13] couples, in general, more weakly to the standard model particles than the QCD axion, which is the only class of axion models in which the coupling to standard model particles is really specified[1]. However, let us consider different production channels. Baiscally, by considering the effective interaction Lagrangian presented in [28]

$$\mathcal{L}_{\text{int}} = i\frac{g_{\text{aNN}}}{m_N}\partial_\mu a(\bar{N}\gamma^\mu\gamma^5 N) + i\frac{g_{\text{aee}}}{2m_e}\partial_\mu a(\bar{e}\gamma^\mu\gamma^5 e) + g_{\text{a}\gamma\gamma} a\vec{E}\cdot\vec{B} \tag{2.47}$$

we can construct different processes. First of all, there are *Bremsstrahlung-processes* like

$$e^- + e^- \leftrightarrow a + e^- + e^- \text{ or } N + \pi \leftrightarrow N + a, \tag{2.48}$$

where N is a nucleon and in general $N = \pi$ is allowed, so processes of the form $a + 1 + 2 \leftrightarrow 3 + 4$ and $a + 1 \leftrightarrow 2 + 3$. Second, we can consider *Primarkoff-production*

$$\gamma + Q \leftrightarrow Q + a, \tag{2.49}$$

where Q is a quark and third, there is *Compton-scattering*

$$e^- + \gamma \leftrightarrow e^- + a \text{ or } e^- \leftrightarrow e^- + a, \tag{2.50}$$

[12]WIMP is short for *weaky-interacting massive particle*, which is like the axion just another dark matter candidate. So far the only detected WIMP, the massive neutrino, is not sufficient to dominate the dark matter energy density we observe in today's Universe, but there is an ongoing effort to search (in-)directly for WIMPs, what is presented for example in [27].

[13]In this subsection we denote an axion by a instead of ϕ to make its appearance in reactions clear.

where the latter two processes are of the form $a + 1 \leftrightarrow 2 + 3$ again. Processes of the form $a + 1 + 2 \leftrightarrow 3 + 4$ and $a + 1 \leftrightarrow 2$ are ruled out by experiment[28]. So we are left with processes of the form $a + 1 \leftrightarrow 2 + 3$, what we can use to compute the thermal axion abundance, n_a. The resulting abundance relative to the thermal photon abundance is given by

$$\frac{n_a}{n_\gamma} = \frac{1}{2}\frac{g_{*,s}(T_{\mathrm{CMB},0})}{g_{*,s}(T^a_{\mathrm{dec}})}, \tag{2.51}$$

what is derived in full detail in appendix D. Using the same notation, one can consider for example axion-pion conversion with nucleons that exist after QCD phase transition at roughly 200 MeV and are non-relativistic with

$$n_N \approx (mT)^{\frac{3}{2}} e^{-x}. \tag{2.52}$$

Further the absorption cross-section is given by

$$\langle\sigma|v|\rangle_{\mathrm{abs}} \sim \frac{m_N^2}{\left(\frac{f_{\mathrm{PQ}}}{N}\right)^2}\left(\frac{T}{m_N}\right)^2 m_\pi^{-2}, \tag{2.53}$$

where we used $g_{\mathrm{aNN}} \approx m_N/(f_{\mathrm{PQ}}/N)$. The axion absorption rate is

$$\Gamma_{\mathrm{abs}} = m_N\langle\sigma|v|\rangle_{\mathrm{abs}} \sim \frac{T^3}{m_\pi^2}\left[\frac{m_N}{\left(\frac{f_{\mathrm{PQ}}}{N}\right)}\right]^2 x^{-\frac{1}{2}}\exp(-x), \tag{2.54}$$

so that the usual ansatz $\Gamma \sim H$ yields

$$\frac{\Gamma_{\mathrm{abs}}}{H} \sim \left(\frac{m_a}{10^{-4}\,\mathrm{eV}}\right)^2 x^{-\frac{3}{2}}\exp(-x). \tag{2.55}$$

For high temperatures, there are no nucleons available and for low temperatures Γ_{abs}/H cuts off exponentially. However, for high temperatures we can consider photo-/gluon-production in the quark-gluon plasma if the quark species Q is relativistic. Then,

$$\langle\sigma|v|\rangle_{\mathrm{abs}} \sim \frac{\alpha}{T^2}\left[\frac{m_Q}{\left(\frac{f_{\mathrm{PQ}}}{N}\right)}\right]^2 \Rightarrow \Gamma_{\mathrm{abs}} \sim \alpha T\left[\frac{m_Q}{\left(\frac{f_{\mathrm{PQ}}}{N}\right)}\right]^2 \tag{2.56}$$

$$\Rightarrow \frac{\Gamma_{\mathrm{abs}}}{H} \sim x\left(\frac{m_Q}{1\,\mathrm{GeV}}\right)^2\left(\frac{m_a}{0.1\,\mathrm{eV}}\right)^2, \tag{2.57}$$

which is valid for $T \gtrsim m_Q$. The maximum value is obtained at quark-hadron phase transition, so to estimate the relic abundance of thermal axions we can simply integrate from the quark-hadron phase transition onwards to get

$$Y_\infty = \frac{0.278}{g_{*,s}}\left[1 - \exp\left\{-\left(\frac{m_a}{10^{-4}\,\text{eV}}\right)^2 \frac{\exp\{-x_{\text{qh}}\}}{x_{\text{qh}}^{\frac{5}{2}}}\right\}\right], \tag{2.58}$$

where we approximated

$$\int_{x_{\text{qh}}}^{\infty} \exp(-x) x^{-\frac{5}{2}} \approx \frac{\exp\{-x_{\text{qh}}\}}{x_{\text{qh}}^{\frac{5}{2}}} \tag{2.59}$$

and $x_{\text{qh}} = m_N/T_{\text{qh}}$, where the subscript stands for quark-hadron. With these equations at hand we get

$$m_a \overset{!}{\gtrsim} 10^{-3} - 10^{-2}\,\text{eV}. \tag{2.60}$$

For $m_a \lesssim 10^{-3}$ eV they interact too weakly to ever get produced. From this we can get

$$n_a = s_0 Y_\infty \approx 83\text{cm}^{-3}\frac{10}{g_{*,s}} \text{ and } \Omega_{th}h^2 = \frac{m_a}{130\,\text{eV}}\frac{10}{g_{*,s}}. \tag{2.61}$$

Using $\Omega_{th}h^2 \overset{!}{\lesssim} 0.12$ from observations, $h = 0.7$ and $g_{*,s} = 17.25$ after QCD phase transition, we get

$$0.12 \gtrsim \Omega_{\text{th}}h^2 = \frac{m_a}{130\,\text{eV}}\frac{10}{g_{*,s}} \Rightarrow m_a \lesssim 26.91\,\text{eV}, \tag{2.62}$$

so that our considered light axion can only give a fraction of todays observed DM density unless they are in the range of the upper bound[28]. Additionally, we should note, that in order to derive (2.51) we assumed that the axions decoupled while relativistic. The immediate consequence is that they will remain relativistic and therefore can only be considered as HDM, whereas we argued in appendix A that the standard cosmological model strictly demands CDM. We therefore conclude that axions cannot that are produced thermally in the early Universe are negligible for our considerations and hence, we need to search for another production mechanism. Further, the axion couplings in (2.47) scale with f_a^{-1}, so that only low-f_a or, equivalently, high-mass axions are produced[1], what interferes with the computed mass above and again with our desire to seek for light particles.

2.3.2 Non-thermal production via misalignment

In order to be able to treat the production of axions via misalignment properly, we need to make a small lookahead on the next chapter. Here, if not stated otherwise, we follow [1]. Let us start by stating that in order to describe the axion, we consider, as before, a scalar field, ϕ. The action for a minimally coupled real scalar field in the theory of General Relativity is given by[29]

$$S_\phi = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right], \tag{2.63}$$

which is only valid after symmetry breaking since this is necessary to initialize the axion as a pseudo-Goldstone boson and it is only valid after non-perturbative effects switch on since this is necessary for the axion to acquire mass. That the latter does not happen instantaneously will introduce a time-dependence, which can be converted to a temperature-dependence as usual, on the equations. In fact, the axion mass will reach an asymptotic value for $T \ll T_{\text{non-perturbative}}$. If this is the case before the axion contributes a significant amount of the energy density and the Universe is still young enough, so that the axion field is not oscillating, then we can simply take the asymptotic value to be constant over all times. However, we can vary the action as we are used to with respect to ϕ in order to get the equation of motion (short: EOM)

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu)\phi - \frac{\partial V}{\partial\phi} =: \Box\phi - \frac{\partial V}{\partial\phi} = 0 \tag{2.64}$$

and we can also vary the action with respect to the metric in order to get the energy-momentum tensor

$$T^\mu_\nu = g^{\mu\alpha}\partial_\alpha\phi\partial_\nu\phi - \frac{\delta^\mu_\nu}{2}\left[g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + 2V(\phi)\right]. \tag{2.65}$$

In case of a perfect fluid, one can interpret the components of the energy-momentum tensor as follows.

$$T^0_0 = -\rho, \quad T^0_i = (\rho + P)v_i, \quad \text{and} \quad T^i_j = P\delta^i_j + \Sigma^i_j, \tag{2.66}$$

where ρ is the energy-density, P is the pressure, v_i is the velocity and Σ^i_j is the anisotropic stress. This interpretation is possible for axions since we assume that they have very small masses and thus must have very high occupation numbers, which validates the fluid interpretation here. However, recalling the flat RW-metric (D.3), one can express the D'Alembertian,

$\Box$, we found in (2.64), as

$$\ddot{\phi} + 3H\dot{\phi} + m_a^2\phi = 0. \tag{2.67}$$

Note, that this is basically the equation of a harmonic oscillator with an additional friction term counting for the Universe's expansion and note additionally, that the RW-metric implies that the velocity and anisotropic stress in (2.65) vanish, so that we get

$$\bar{\rho}_a = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}m_a^2\phi^2 \quad \text{and} \quad \bar{P}_a = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}m_a^2\phi^2 \tag{2.68}$$

as the background energy density and background pressure, respectively. As discussed in appendix A, the scale factor sclaes as a power-law with time, $a \sim t^p$. Then, the EOM can be solved exactly

$$\phi = a^{-\frac{3}{2}} \left(\frac{t}{t_i}\right)^{\frac{1}{2}} [C_1 J_n(m_a t) + C_2 Y_n(m_a t)], \tag{2.69}$$

where $n = (3p-1)/2$, $J_n(x)$ and $Y_n(x)$ are Bessel functions of the first and second kind, respectively. With the proper initial conditions one can solve now for the coefficients $C_{1,2}$. When the PQ-symmetry breaks, $H \gg m_a$, so that ϕ is overdamped as can be seen directly in the EOM (2.64), so that $\dot{\phi}_i = 0$ is a reasonable initial condition. Corresponding to (2.15) we can additionally state $\phi(t_i) = f_a\theta_i$ as the second initial condition. Note, that in both conditions we used the subscript i for an initial value. The second condition now justifies the name for this production mechanism since the initial axion field is initially displaced from its potential minimum and thus needs realignment of the vacuum in order to get back to its minimum. As already said, for $H > m_a$ the motion of ϕ is overdamped and governed by the friction term due to expansion, so that it has a constant scaling with the scale factor and thus, the axion's equation of state parameter is $w_a = -1$ and the axion basically just contributes to DE. As time goes by, eventually $H < m_a$ occurs, meaning the motion is then underdamped and oscillation of ϕ begins. Since $\rho_a \sim a^{-3}$ now, the axions behave as ordinary matter, so that the oscillation should occur around $w_a = 0$. We will prove this later. Let a_{eq} be the scale factor at matter-radiation equality, then $H(a_{\text{eq}}) \sim 10^{-28}$ eV. This means, that if $m_a > H(a_{\text{eq}})$, the axions start oscillating in the radiation-dominated epoch.

Let a_{osc} be the scale factor when axion field oscillation begins, then since $\rho_a(a) \sim a^{-3}$ we can approximate

$$\rho_a(a)a^3 \approx \rho_a(a_{\text{osc}})a_{\text{osc}}^3 \tag{2.70}$$

for $a > a_{\text{osc}}$. Since we said that the energy density is almost constant for $a < a_{\text{osc}}$, we can also approximate

$$\rho_a(a_{\text{osc}}) \approx \frac{1}{2} m_a^2 \phi_i^2, \tag{2.71}$$

so that we only need the axion mass, m_a, and the initial field displacement, ϕ_i, to get the energy density of the axion population generated via misalignment. By considering the EOM (2.64) one immediately would suggest $3H(a_{\text{osc}}) = m_a$ as a good ansatz for a_{osc} and in fact it turns out that for real-Universe models, this is a well-fitting approximation[30]. One can now use the well-known $H(t)$-solutions during the radiation- and matter-dominated epoch, respectively, one can compute approximately[14] the energy density parameter of ULAs as a function of ϕ_i

$$\Omega_a \approx \left\langle \left(\frac{\phi_i}{M_{\text{pl}}} \right)^2 \right\rangle \cdot \begin{cases} \frac{1}{6}(9\Omega_\gamma)^{\frac{3}{4}} \left(\frac{m_a}{H_0}\right)^{\frac{1}{2}} & a_{\text{osc}} < a_{\text{eq}} \\ \frac{9}{6}\Omega_m & a_{\text{eq}} < a_{\text{osc}} \lesssim 1\,, \\ \frac{1}{6}\left(\frac{m_a}{H_0}\right)^{\frac{1}{2}} & a_{\text{osc}} \gtrsim 1 \end{cases} \tag{2.72}$$

where the angular brackets are used to either average over the θ-interval or simply use the given ϕ_i-value corresponding to the drawn θ in the unbroken or broken PQ-symmetry case, respectively, as we have discussed in section 2.2. Now, for $H \ll m_a$ we can make a WKB-approximation with the ansatz

$$\phi(t) = A(t)\cos(m_a t + \vartheta), \tag{2.73}$$

where ϑ is an arbitrary phase and the amplitude $A(t)$ is, in consistence with the slow-roll inflation model, slowly varying, i.e.

$$\frac{\dot{A}(t)}{m_a} \sim \frac{H(t)}{m_a} \ll 1. \tag{2.74}$$

Pluggin this ansatz in the EOM (2.64), to leading order we get $A(t) \sim a^{-\frac{3}{2}}$. Recalling (2.71) this leads to

$$\rho_a \sim \phi_i^2 \sim A^3 \sim a^{-3}, \tag{2.75}$$

giving the expected scaling behavior for ordinary matter, so the produced axions actually contribute to the matter content of the Universe. The axion field oscillations rapidly with frequency $2m_a$, so $\langle w_a \rangle_t = 0$ is the average

[14] Please refer to [31] for a short, but detailed derivation.

equation-of-state parameter for axions for $t \gg 1/m_a$, which proves that w_a oscillates around $w_a = 0$. Note, that this result is independent of the background evolution. In order to contribute significantly to the DM energy density, ULAs should have $f_a \gtrsim \phi_i > 10^{14}\,\mathrm{GeV}$. Now recall from section 2.2 that this corresponds to $f_a > H_I/2\pi = T_I$ and thus to the broken PQ-symmetry scenario. Thus, for general ALPs, or ULAs in particular, we will always consider them in the broken case.

Let us now consider the QCD-axion, but note, that we will go over the arguments rather quick. The temperature-dependent axion mass is given by

$$m_a^2(T) = \alpha_a \frac{\Lambda_{\mathrm{QCD}}^3 m_u}{f_a^2} \left(\frac{T}{\Lambda_{\mathrm{QCD}}} \right)^{-n}, \tag{2.76}$$

where $n \approx 8$. By using the Friedmann equation in the radiation-dominated epoch, one can additionally show

$$3H^2 M_{\mathrm{pl}}^2 = \frac{\pi^2}{30} g_* T^4. \tag{2.77}$$

When taking $g* = 61.75$ for $T \gtrsim T_{\mathrm{QCD}}$ and recalling that one can approximately use $3H(T_{\mathrm{osc}}) = m_a$, then the QCD-axion with $f_a < 10^{15}\,\mathrm{GeV}$ has $T_{\mathrm{osc}} = 1\,\mathrm{GeV}$. Again, the energy density scales as that of ordinary matter thereafter independent of the axion mass and one can thus compute m_a from (2.76) for large temperatures. However, one can show that

$$\Omega_a h^2 \sim 2 \cdot 10^4 \left(\frac{f_a}{10^{16}\,\mathrm{GeV}} \right)^{\frac{7}{6}} \langle \theta_i^2 \rangle \quad \text{if } f_a < 2 \cdot 10^{15}\,\mathrm{GeV} \text{ and} \tag{2.78}$$

$$\Omega_a h^2 \approx 5 \cdot 10^3 \left(\frac{f_a}{10^{16}\,\mathrm{GeV}} \right)^{\frac{3}{2}} \langle \theta_i^2 \rangle \quad \text{if } f_a > 2 \cdot 10^{17}\,\mathrm{GeV}. \tag{2.79}$$

It was shown that we can safely use (2.78) for all $f_a < 6 \cdot 10^{17}\,\mathrm{GeV}$. Till now, we have used the harmonic potential

$$V(\phi) = \frac{1}{2} m_a^2 \phi^2, \tag{2.80}$$

but for $\theta \gtrsim 1$ we must take anharmonic corrections into account, which are caused by axion self-interactions. We are thus led to replace

$$\langle \theta_i^2 \rangle \to \langle \theta_i^2 F_{\mathrm{anh}}(\theta_i) \rangle, \tag{2.81}$$

where $F(x) \to 1$ for x small and monotonically increasing for $x \to \pi$. For our typical cosine potential (2.26) one can find

$$F_{\text{anh}}(x) = \left[\ln \left(\frac{e}{1 - \frac{x^2}{\pi^2}} \right) \right]^{\frac{7}{6}}. \tag{2.82}$$

With this replacement at hand, recall (2.78) for $\Omega_a h^2$ and further recall that in the broken PQ-symmetry case (2.45) and for the unbroken case (2.44) holds, respectively, then we obtain

$$\Omega_a h^2 \approx 2 \cdot 10^4 \left(\frac{f_a}{10^{16}\,\text{GeV}} \right)^{\frac{7}{6}} \cdot \begin{cases} \frac{\pi^2}{3} F_{\text{anh}} \left(\frac{\pi}{\sqrt{3}} \right) (1 + \alpha_{\text{dec}}) \\ \left(\theta_i^2 + \frac{H_I^2}{(2\pi f_a)^2} \right) F_{\text{anh}} \left(\sqrt{\theta_i^2 + \frac{H_I^2}{(2\pi f_a)^2}} \right) \end{cases}, \tag{2.83}$$

where the upper (lower) equation obviously corresponds to the unbroken (broken) case. It is interesting to note, that $\alpha_{\text{dec}} \in [0.16, 186]$ is a rather large interval, which opens the possibility that all $f_a \lesssim 9 \cdot 10^{10}\,\text{GeV}$ axions could give the correct $\Omega_a h^2$ if α_{dec} is chosen accordingly. Now recall from appendix A that the DM energy density parameter is

$$\Omega_{\text{DM}} h^2 \approx 0.12, \tag{2.84}$$

so that

$$\Omega_a h^2 \leq \Omega_{\text{DM}} h^2 \approx 0.12 \tag{2.85}$$

is an upper bound the QCD-axions should suffice.Let the PQ-symmetry be unbroken during inflation, then $\langle \theta_i \rangle = \pi^2/3$ is fixed and we immediately get an upper bound for f_a. In order to satisfy other known bounds from observation we get

$$1 \cdot 10^9\,\text{GeV} \lesssim f_a \lesssim 8.5 \cdot 10^{10}\,\text{GeV}. \tag{2.86}$$

Let the PQ-symmetry now be broken during inflatin, then θ_i is a free parameter and we can chose it according to f_a to satisfy the dark matter density parameter bound given above. Additionally we can choose $\theta \to \pi$, so that F_{anh} diverges in order to compensate very low f_a values. If one sets the boundary for this tuning process at the order 10^{-2} for θ_i, we get

$$8 \cdot 10^9\,\text{GeV} \lesssim f_a \lesssim 1 \cdot 10^{15}\,\text{GeV}. \tag{2.87}$$

Let us summarize this subsection by appreciating that the axion production via misalignment is model-independent, that fully builds up on the axion-defining properties. Furthermore, it only depends on gravitational interaction as the EOM were derived for an action (2.63) in the setting of pure General Relativity. It thus appears as a very natural and elegant mechanism to provide an explanation for a relevant initial axion population.

2.3.3 Decay product of heavier parent particle X or of toplogical defect/string

Let us assume that there is a heavier parent particle, X, with $m_X > m_a$, which is coupled to the axion field and decays producing a population of relativistic axions. If the decay happens after the axions have decoupled from the standard model, they remain relativistic throughout the history of the Universe and can be used to describe dark radiation, which is parametrized via the effective number of relativistic neutrinos. One assumes the decay to be instantaneous when the parent particle dominates the energy density of the Universe[1]. If parent particle does not dominate the energy density of the Universe when it decays, they may act as so-called *curvatons*, which are theoretical particles that would be responsible for the initial curvature fluctuations. For details on that, please refer to [32]. Dark radiation would affect the CMB, see [33] for details, where the main effect is an additional damping in the high-multiple acoustic peaks. This results if one fixes the angular size of the sound horizon, then one has to compensate the change in matter-radiation equality by either a different Hubble constant or by a different DE density[1]. However, the damping tail of the CMB is well-measured and so it puts strong constraints on the effective number of neutrinos. Typically, the whole idea appears in models with SUSY and extra dimensions, where the main outcome is that the heavy parent particle in fact is a heavy modulus with masses larger than 10 TeV in order to produce a suitable axion population[1]. Note again, that these are way too heavy for our considerations.

As we are already talking about theories with higher dimensions, we can consider the following ideas. One knows that the breaking of a global symmetry results in a so-called *toplogical defect* and in the case of the PQ-symmetry, global axionic strings are produced and the toplogical defects correspond to domain walls if $N_{\mathrm{DW}} > 1$. We can distinguish now wheter the PQ symmetry is broken during inflation or not. If it is broken, the toplogical effects are vanished by inflation, so let us assume it is unbroken during inflation. The decay of axionic strings would produce a population of cold axions with energy densities of the domain walls $\rho_{\mathrm{DW}} \sim a^{-2}$, which can rather quickly can dominate the energy density of the Universe if we keep $\rho_m \sim a^{-3}$ in mind. This would lead to a crucial change of the picture of the Universe's evolution. This is why these models typically need another mechanism to get rid off the domain wall problem[1]. The underlying math is far beyond the scope of this work and thus, we do not consider these models any further, even though in theory, they are capable of producing cold axion and thus CDM populations to account for.

Chapter 3

Dynamics of cosmological perturbations

In this chapter, we are going to first build up the linear cosmological perturbation theory in section 3.1 and then going to apply it on axion fields in section 3.2. In the first section we set up the basic equations, see subsection 3.1.1, holding always based on grounds of GR and then move on to describe conventional hydrodynamical matter, see subsection 3.1.2. Then we are ready to tackle the, for us most interesting, scalar matter field perturbations, see subsection 3.1.3, and their quantization, see subsection 3.1.4. In the second section, we again, first set up the basic equations, see subsection 3.2.1, and then move on to the mentioned application of linear cosmological perturbation theory in subsection 3.2.2. In the end, we consider the non-linear regime in subsection 3.2.3 and with it the Schrödinger-Poisson equation.

3.1 Gauge-invariant cosmological perturbation theory

In this section we will closely follow [2] if not stated otherwise in order to work out a gauge-invariant description of cosmological perturbations. I strongly recommend [2] for a very detailed, but comprehensible, derivation and presentation of the following ideas. However, the following subsections will equip us with a strong and easily applicable set of equations and interpretations in order to put the implications of the axion field we desire to describe in a proper context. First, we are going to consider hydrodynamical and scalar matter field perturbations, respectively, in a classical setup. The description of hydrodynamical matter in subsection 3.1.2 connects well to our intuition about the radiation- and matter-dominated epochs and also sensitizes us to the necessary steps towards a proper description of cosmological perturbations. Then we can tackle the classical description of scalar matter field perturbations in subsection 3.1.3, what will lay down the crucial foundation we need in order to move on to the quantum description of this type of perturbations. Hence, in subsection 3.1.4 we are dealing with the spectrum of density perturbations in inflationary Universe models with scalar-field matter, what will connect well to our axion field description. However, before we can even start thinking about any of these perturbations, we need to setup the basic equations and discuss the basic vocabulary in the first subsection.

3.1.1 Aim and Setup

As stated previously, we seek for a gauge-invariant description of cosmological perturbations. This approach is easier than working in a specific gauge and then transforming to other gauges in order to tackle different problems and to straightforwardly interpret the resulting equations. Initial conditions are treated separately since the quantum theory will give the initial conditions needed for the classical theory. They are necessary since we assume that there were small initial perturbations in the very early Universe that grew over time constituting the large-scale structure we observe today.

Consider a homogeneous and isotropic spacetime, $g_{\mu\nu}^{(0)}$, what we will call the *background*, denoted by a superscript (0), and small deviations, $\delta g_{\mu\nu}$, from the background, so that in total, the spacetime is parametrized by the total metric

$$g_{\mu\nu} = g_{\mu\nu}^{(0)} + \delta g_{\mu\nu}. \tag{3.1}$$

As always, we use the RW-metric,

$$ds^2 = dt^2 - a^2(t)\gamma_{ij}dx^i dx^j = a^2(\eta)(d\eta^2 - \gamma_{ij}dx^i dx^j) \tag{3.2}$$

with

$$\gamma_{ij} := \delta_{ij}[1 + \frac{\kappa}{4}(x^2 + y^2 + z^2)]^{-2}, \tag{3.3}$$

either in coordinate-time, t, or in conformal time, η, where $dt = ad\eta$ connects them. Usually we will denote a derivative with respect to coordinate (conformal) time by a dot (prime). One distinguishes three types of perturbations. *Scalar perturbations* that give growing inhomogeneities, *vector perturbations* that decay if the Universe expands and *tensor perturbations* that give gravitational waves. The names result from the way the corresponding fields transform. The most general form of scalar perturbations

$$\delta^{(s)}_{\mu\nu} = a^2(\eta) \begin{pmatrix} 2\phi & -B_{|i} \\ -B_{|i} & 2(\psi\gamma_{ij} - E_{|ij}) \end{pmatrix}, \tag{3.4}$$

where a vertical bar (comma) in front of an index denotes the covariant (partial) derivative with respect to the three-dimensional background coordinates. Note, that the covariant derivatives become partial derivatives in a flat, $\kappa = 0$, Universe. Since (3.1) holds, the general form of the metric perturbations (3.4) gives

$$ds^2 = a^2(\eta)[(1+2\phi)d\eta^2 - 2B_{|i}dx^i d\eta - ((1-2\psi)\gamma_{ij} + 2E_{|ij}dx^i dx^j)]. \tag{3.5}$$

The four scalar functions, ϕ, ψ, B and E, are functions of the spacetime-coordinates x^μ and are fixed by four equations that are collected in $\delta g_{\mu\nu}$. Since vector and tensor perturbations do not lead to growing inhomogeneities, we are not interested in them in this work, so we do not go deeper in the description of them. In general, one should consider all three types altogether, but in the linear theory, they can be considered separately, what simplifies the following discussions.

In order to be able to derive gauge-invariant equations, we need to express them in gauge-invariant quantities, which we will derive now. Let $\xi^\mu = (\xi^0, \xi^i)$ be an infinitesimal perturbation to the spacetime-coordinates, $x^\mu = (x^0 = \eta, x^i)$, and ξ a from ξ^0 independent scalar function, that solves

$$\xi^{|i}_{|i} = \xi^i_{|i}, \tag{3.6}$$

then the most general coordinate transformation that preserves the scalar nature of the metric perturbations is

$$\eta \rightarrow \tilde{\eta} = \eta + \xi^0 \quad \text{and} \quad x^i \rightarrow \tilde{x}^i = x^i + \gamma^{ij}\xi_{|j}. \tag{3.7}$$

By applying this transformation to the scalar perturbations, one gets additional terms that can be collected in the four scalar functions constituting $\delta g^{(s)}_{\mu\nu}$ by defining

$$\tilde{\phi} = \phi - \frac{a'}{a}\xi^0 - \xi'^0,\ \tilde{\psi} = \psi + \frac{a'}{a}\xi^0,\ \tilde{B} = B + \xi^0 - \xi',\ \tilde{E} = E - \xi. \tag{3.8}$$

However, one can construct two[1] gauge-invariant variables

$$\Phi := \phi + \frac{1}{a}[(B - E')a]' \quad \text{and} \quad \Psi := \psi - \frac{a'}{a}(B - E'), \tag{3.9}$$

which form the basic variables in whose terms we would like to write our equations often during the subsequent subsections and are indeed invariant under the infinitesimal coordinate transformation (3.7).

Since the transformation of the four fundamental scalar variables were dependent solely on the two independent scalar fields ξ^0 and ξ (3.8), one has two degrees of freedom on which one can impose constraints by choosing a certain gauge[2]. Although there is a vast amount of different gauges, we are only interested in the two most famous. The *synchronous gauge* sets $\phi = 0 = B$ and likewise $\tilde{\phi} = 0 = \tilde{B}$ in (3.7), so that one can solve for ξ^0 and ξ to obtain

$$\eta \to \eta_s = \eta + \frac{1}{a}\int a\phi d\eta \text{ and } x^i \to x^i_s = x^i + \gamma^{ij}\left(\int Bd\eta + \int \frac{d\eta}{a}\int a\phi d\eta\right)_{|j}. \tag{3.10}$$

This gauge does not fix the synchronous coordinates totally as one can perform a residual transformation that leaves the synchronous gauge conditions invariant, but changes the transformation above. This can result in unphysical gauge modes that cause trouble when interpreting the results. The *longitudinal gauge* sets $B = 0 = E$, so that the coordinates are totally fixed since these conditions fix ξ^0 and ξ via (3.8). By plugging the conditions in (3.7) we get

$$\eta \to \eta_l = \eta - (B - E') \quad \text{and} \quad x^i \to x^i_l = x^i + \gamma^{ij}E_{|j}. \tag{3.11}$$

Note, that by setting $B = 0 = E$ (3.8) becomes[3]

$$\phi_l = \phi + \frac{a'}{a}(B - E')' \overset{(3.9)}{=} \Phi,\ \psi_l = \phi - \frac{a'}{a}(B - E') \overset{(3.9)}{=} \Psi \tag{3.12}$$

[1]In fact, one can construct an infinite amount of gauge-invariant variables by simply taking linear combinations of them since they should be gauge-invariant again. We choose a particular simple form of them.

[2]It seems to be counterintuitive to choose a specific gauge if our initial goal is to find gauge-invariant equations. However, choosing a gauge can help to get straightforward interpretations, what we will encounter later.

[3]This gauge is also called the *conformal-Newtonian gauge*, what refers to $\phi_l = \Phi$ as one can interpret Φ as the generalization of the Newtonian gravitational potential, what will become clearer in subsection 3.1.2.

and apparently $B_l = 0 = E_l$. Of course one could find transformation laws to switch between the gauges, but let us avoid to use them by carefully choosing the proper gauge for our purposes.

Let us now close our setup by considering the Einstein equations

$$G^\mu_\nu = 8\pi G T^\mu_\nu, \tag{3.13}$$

where

$$G^\mu_\nu := R^\mu_\nu - \frac{1}{2}\delta^\mu_\nu R \tag{3.14}$$

is the Einstein tensor, R^μ_ν is the Ricci tensor and $R = R^\mu_\mu$ is the Ricci curvature scalar. We explicitly use the Einstein tensor instead of writing the Einstein equations in one line because it will come in handy in the following discussions. In a homogeneous and isotropic Universe, the background Einstein equations are

$${}^{(0)}G^0_0 = 3\frac{1}{a^2}(H^2+\kappa), \quad {}^{(0)}G^0_i = 0 \quad \text{and} \quad {}^{(0)}G^i_j = \frac{1}{a^2}(2H' + H^2 + \kappa)\delta^i_j, \tag{3.15}$$

where the Hubble parameter, H, uses conformal time now[4]. Likewise we get

$${}^{(0)}G^\mu_\nu = 8\pi G\, {}^{(0)}T^\mu_\nu, \tag{3.16}$$

where the background energy-momentum tensor satisfies

$${}^{(0)}T^i_0 = 0 = {}^{(0)}T^0_i \quad \text{and} \quad {}^{(0)}T^i_j \sim \delta^i_j. \tag{3.17}$$

For small perturbations we can also split the Einstein tensor in a background and a perturbed part

$$G^\mu_\nu = {}^{(0)}G^\mu_\nu + \delta G^\mu_\nu, \tag{3.18}$$

where the perturbation should also satisfy the EOM for small perturbations linearized about the background metric

$$\delta G^\mu_\nu = 8\pi G \delta T^\mu_\nu. \tag{3.19}$$

[4]It should always be clear out of the context if we use H in conformal or coordinate time.

With this groundwork we can obtain the Einstein tensor for scalar perturbations

$$\delta G^0_0 = 2a^{-2}\{-3H(H\phi+\psi')+\nabla^2[\psi - H(B-E')]+3\kappa\psi\} \tag{3.20}$$

$$\delta G^0_i = 2a^{-2}[H\phi+\psi'-\kappa(B-E')]_{|i} \tag{3.21}$$

$$\delta G^i_j = -2^{-2}\{[(2H'+H^2)\phi+H\phi'+\psi''+2H\psi'-\kappa\psi+\frac{1}{2}\nabla^2 D]\delta^i_j - \frac{1}{2}D_{|ij}\}, \tag{3.22}$$

where

$$D := (\phi-\psi)+2H(B-E')+(B-E')'. \tag{3.23}$$

Since the equations are not gauge-invariant yet we rewrite them in terms of Φ, Ψ and $(B-E')$ to obtain

$$\delta G^0_0 = 2a^{-2}[-3H(H\Phi+\Psi')+\nabla^2\Psi+3\kappa\Psi+3H(-H'+H^2+\kappa)(B-E')] \tag{3.24}$$

$$\delta G^0_i = 2a^{-2}[H\Phi+\Psi'+(H'-H^2-\kappa)(B-E')]_{|i} \tag{3.25}$$

$$\begin{aligned}\delta G^i_j = -2a^{-2}\{&[(2H'+H^2)\Phi+H\Phi'+\Psi''+2H\Psi'-\kappa\Psi+\frac{1}{2}\nabla^2 D]\delta^i_j \\ &+(H''-HH'-H^3-\kappa H)(B-E')\delta^i_j - \frac{1}{2}\gamma^{ik}D_{|kj}\},\end{aligned} \tag{3.26}$$

where

$$D = \Phi - \Psi. \tag{3.27}$$

By applying (3.7) on the Einstein tensor, one finds the corresponding transformation law

$$\delta G^0_0 \to \delta G^0_0 - ({}^{(0)}G^0_0)'\xi^0, \tag{3.28}$$

$$\delta G^0_i \to \delta G^0_i - \left({}^{(0)}G^0_0 - \frac{1}{3}{}^{(0)}G^j_j\right)\xi^0_{|i} \quad \text{and} \tag{3.29}$$

$$\delta G^i_j \to \delta G^i_j - ({}^{(0)}G^i_j)'\xi^0, \tag{3.30}$$

so that with $\xi^0 = -(B-E')$ one immediately can construct the gauge-invariant, denoted by a superscript (gi), variable $\delta G^{(\mathrm{gi})\mu}_\nu$ by

$$\delta G^{(\mathrm{gi})0}_0 = \delta G^0_0 + ({}^{(0)}G^0_0)'(B-E'), \tag{3.31}$$

$$\delta G^{(\mathrm{gi})0}_i = \delta G^0_i + \left({}^{(0)}G^0_0 - \frac{1}{3}{}^{(0)}G^j_j\right)(B-E')_{|i} \quad \text{and} \tag{3.32}$$

$$\delta G^{(\mathrm{gi})i}_j = \delta G^i_j + ({}^{(0)}G^i_j)'(B-E') \tag{3.33}$$

and likewise $\delta T_\nu^{(\mathrm{gi})\mu}$ by simply replacing G by T in the above equations. The EOM (3.19) can now be written in the gauge-invariant form

$$\delta G_\nu^{(\mathrm{gi})\mu} = 8\pi G \delta T_\nu^{(\mathrm{gi})\mu}. \tag{3.34}$$

Finally, by plugging (3.24) in (3.31) all terms on the left hand side proportional to $(B - E')$ and its derivatives cancel, so that by plugging the result then in (3.34) one gets the general form of the gauge-invariant equations of cosmological perturbations

$$4\pi G a^2 \delta T_0^{(\mathrm{gi})0} = -3H(H\Phi + \Psi') + \nabla^2\Psi + 3\kappa\Psi, \tag{3.35}$$

$$4\pi G a^2 \delta T_i^{(\mathrm{gi})0} = (H\Phi + \Psi')_{|i} \quad \text{and} \tag{3.36}$$

$$\begin{aligned} -4\pi G a^2 \delta T_0^{(\mathrm{gi})0} = {} & [(2H' + H^2)\Phi + H\Phi' + \Psi'' + 2H\Psi' \\ & - \kappa\Psi + \frac{1}{2}\nabla^2 D]\delta_j^i - \frac{1}{2}\gamma^{ik} D_{|kj}\,, \end{aligned} \tag{3.37}$$

where $D = \Phi - \Psi$. We are now prepared to discuss hydrodynamical and scalar matter field perturbations in the subsequent subsections.

3.1.2 Classical description of hydrodynamical perturbations

We begin by considering hydrodynamical matter as first step to gain insights in how well-suited our approach is in terms of how well it matches the expected results. The evolution of initial *adiabatic fluctuations*[5] or entropy perturbations[6] depends on whether we consider the radiation- or matter-dominated epoch. By careful analysis one can consider the perturbations in the radiation- and matter-dominated epoch separately and then make an effort in smoothly connecting the results. However, we just want to very quickly give a brief overview over the basic equations and the final results to give a short impression of the application of the formalism before going deeper into for us relevant applications in the subsequent subsections.

The energy-momentum tensor for hydrodynamical matter fluctuations is given by

$$\delta T_0^{(\mathrm{gi})0} = \delta\varepsilon^{(\mathrm{gi})}, \quad \delta T_i^{(\mathrm{gi})0} = (\varepsilon_0 + p_0)a^{-1}\delta u_i^{(\mathrm{gi})}, \quad \delta_j^{(\mathrm{gi})i} = -\delta p^{(\mathrm{gi})}\delta_j^i, \tag{3.38}$$

[5]These are the typical density fluctuations as opposed to the isocurvature fluctuations we are going to investigate later.

[6]We will discuss them later in more detail.

where $\delta\varepsilon^{(\mathrm{gi})}$ are the energy density perturbations, $\delta p^{(\mathrm{gi})}$ the pressure perturbations and $\delta u_i^{(\mathrm{gi})}$ the velocity perturbations in gauge-invariant form. Likewise, the EOM reads

$$\Phi'' + 3H(1+c_s^2)\Phi' - c_s^2\nabla^2\Phi + [2H' + (1+3c_s^2)(H^2-\kappa)]\Phi = 4\pi G a^2 \tau\delta S, \tag{3.39}$$

where Φ is the generalization of the Newtonian gravitational potential and c_s is the sound speed. Alternatively, one can write the EOM in terms of the velocity u as

$$u'' - c_s^2\nabla^2 u - \frac{\theta''}{\theta}u = \mathcal{N} \tag{3.40}$$

with

$$\theta = \frac{H}{a}\left[\frac{2}{3}(H^2 - H' + \kappa)\right]^{-\frac{1}{2}} = \frac{1}{a}\left(\frac{\varepsilon_0}{\varepsilon_0 + p_0}\right)^{\frac{1}{2}}\left(1 - \frac{3\kappa}{8\pi G a^2 \varepsilon_0}\right)^{\frac{1}{2}}, \tag{3.41}$$

$$\mathcal{N} = (4\pi G)^{\frac{1}{2}} a^3 (H^2 - H' + \kappa)^{-\frac{1}{2}}\tau\delta S = a^2(\varepsilon_0 + p_0)^{-\frac{1}{2}}\tau\delta S. \tag{3.42}$$

Additionally, one finds

$$\frac{\delta\varepsilon^{(\mathrm{gi})}}{\varepsilon_0} = 2[3(H^2+\kappa)]^{-1}[\nabla^2\Phi - 3H\Phi' - 3(H^2-\kappa)\Phi] \tag{3.43}$$

and

$$\delta u_i^{(\mathrm{gi})} = -a^{-2}(H^2 - H' + \kappa)^{-1}(a\Phi)'_{,i}. \tag{3.44}$$

Let us investigate adiabatic perturbations. The right-hand side of the EOM vanishes, so that the EOM becomes homogeneous and solvable, for example in the radiation-dominated, $p = \frac{1}{3}\varepsilon$ or the matter-dominated, $p = 0$, epoch[7]. Just for an impression, let me directly give the results for non-decaying long-wavelength, $\lambda > r_{\mathrm{Hubble}}$, mode inhomogeneities by first noting $|\Phi| = $ const. in both epochs, $|\delta\varepsilon^{(\mathrm{gi})}/\varepsilon_0| \approx -2\Phi$ in the radiation-dominated and $|\delta\varepsilon^{(\mathrm{gi})}/\varepsilon_0| = $ const. in the matter-dominated epoch. Then consider the transition between the epochs, where the equation of state parameter changes from $w_\gamma = \frac{1}{3}$ to $w_m = 0$, where both functions drop by a factor 9/10. Lastly, one has to distinguish if the perturbations enter the horizon or if they remain outside. In the former case $|\Phi| = $ const. and $|\delta\varepsilon^{(\mathrm{gi})}/\varepsilon_0| \sim \eta^2$ holds, whereas in the latter case $|\Phi| = $ const. and $|\delta\varepsilon^{(\mathrm{gi})}/\varepsilon_0| = $ const. holds. In this example, both results were obtained in

[7]To stick to the author's notation, I will adapt to ε as the energy density rather than ρ. Since both are commonly used, one should not be confused.

the matter-dominated epoch.

Now, let us turn to the following fact. Additionally to adiabatic perturbations, entropy perturbations generically arise in all multi-component systems and generate scalar-type perturbations. The right hand side[8] does not vanish. Let us assume that at a certain initial time, there are no adiabatic perturbations and we only consider entropy perturbations. Their initial condition[9] is $\Phi \to 0$ as $t \to 0$. They can be produced in axion models and in non-simple inflationary Universes and in phase transitions producing topological defects. Causality forbids creation of adiabatic perturbations on scales larger than r_{Hubble}, so only entropy perturbations are possible. Let me, again to just give you an impression, directly state the results for long-wavelength perturbations, for which we start in the radiation-dominated epoch, in which $|\Phi|$ and $|\delta\varepsilon^{(\mathrm{gi})}/\varepsilon_0|$ increase linearly, where the latter one increases slightly faster than the former one, and both are $\sim \frac{\delta S}{S}$ from 0 till η_{eq}. Likewise in the matter-dominated epoch one obtains $|\delta\varepsilon_m/\varepsilon_{m,0}| = \mathrm{const.}$. Now, cold-matter perturbations drop to $2/5$ of initial value at η_{eq} and coincide with radiation perturbations/$\delta\varepsilon$ in general. Lastly, for $\eta > \eta_{\mathrm{eq}}$ we get the same evolution for entropy perturbations as for adiabatic perturbations. Note, that there is a difference of a factor 2.

3.1.3 Classical description of scalar matter field perturbations

At very large energies, the hydrodynamical matter description fails and we need a description in terms of fields, so that we now turn to describe scalar-matter fields, which are far more relevant for our ALP considerations as we already made the mental step to consider the axion field, ϕ, rather than the ALP itself. However, let us set up the basic equations and then treat their implications.

[8]In general, all source terms $\sim \delta S$ do not vanish.

[9]Let me note here the following. The initial conditions for isocurvature perturbations are gauge-invariant curvature perturbations $\zeta \to 0$ at t_i. Entropy and isocurvature perturbations coincide for $t_i = 0$, which is why we will usually speak of isocurvature rather than entropy perturbations in the rest of this work.

The energy-momentum tensor in this case is

$$\delta T_0^{(\text{gi})0} = a^{-2}[-(\varphi_0')^2\Phi + \varphi_0'\delta\varphi^{(\text{gi})\prime} + V_{,\varphi}a^2\delta\varphi^{(\text{gi})}] \tag{3.45}$$

$$\delta T_i^{(\text{gi})0} = a^{-2}\varphi_0'\delta\varphi_{,i}^{(\text{gi})}, \tag{3.46}$$

$$\delta T_j^{(\text{gi})i} = a^{-2}[(\varphi_0')^2\Phi - \varphi_0'\delta\varphi^{(\text{gi})\prime} + V_{,\varphi}a^2\delta\varphi^{(\text{gi})}]\delta_j^i. \tag{3.47}$$

The EOM thus reads

$$\ddot{\varphi} + 3H\dot{\varphi} + V_{,\varphi} = 0 \tag{3.48}$$

First, let us investigate the backgroune, in which an important difference to hydrodynamical matter appears since the background part of $\varphi_0(t)$ is time-dependent and has non-trivial dynamics, so that the equation of state is time-dependent as well. This in turn results in a complicated dynamics of the time-dependent scale factor. For example, consider the $0-0$-Einstein equation

$$H^2 = l^2\left[\frac{1}{2}\dot{\varphi}_0^2 + V(\varphi)\right] \tag{3.49}$$

and assume that φ is roughly static, so that $\dot{\varphi}_0^2 \approx 0$, but $V(\varphi)$ is very large, so that $l^2V(\varphi) \approx$const., what gives $a(t) \sim \exp\{Ht\}$ since $H := \dot{a}/a$. This exponential inflation corresponds to $p \approx -\varepsilon$, i.e. $w = -1$, and $l = \frac{8\pi G}{3}$. With the transformation

$$x := ln(a) \Rightarrow \dot{x} = \frac{1}{a}\dot{a} \equiv H \Leftrightarrow \frac{d}{dt} = H\frac{d}{dx} \tag{3.50}$$

the $0-0$-Einstein equation above transforms to

$$H = lV^{\frac{1}{2}}\left[1 - \frac{1}{2}l^2\left(\frac{d\varphi}{dx}\right)^2\right]^{-\frac{1}{2}}. \tag{3.51}$$

and further by using the EOM (3.48) we get

$$\left\{\frac{d^2\varphi}{dx^2} + 3\left[1 - \frac{1}{2}l^2\left(\frac{d\varphi}{dx}\right)^2\right]\frac{d\varphi}{dx}\right\}\frac{l^2V}{1 - \frac{1}{2}l^2\left(\frac{d\varphi}{dx}\right)^2} + V_{,\varphi} = 0. \tag{3.52}$$

Now, by defining the variable

$$y := \frac{d\varphi}{dx} \tag{3.53}$$

we obtain

$$\frac{dy}{d\varphi} = -\left(1 - \frac{1}{2}l^2y^2\right)\left[3 + \frac{\frac{V_{,\varphi}}{V}}{l^2 y}\right] \tag{3.54}$$

and since $H \in \mathbb{R} : |y| \leq \frac{\sqrt{2}}{l}$ follows immediately by demanding that the term in square brackets is positive semidefinite. When $|y| > \frac{\sqrt{2}}{l}$ one describes Euclidean solutions since $t \in \mathbb{C}$. The resulting $\varphi - y$ diagram is $V(\varphi)$-independent if $V(\varphi)$ is symmetric in φ and $|\varphi| \to \infty$ and $V(\varphi)$ increases slower than an exponential, so that $\frac{V_{,\varphi}}{V} \to 0$ as $|\varphi| \to \infty$. For the solution close to $y = \pm\frac{\sqrt{2}}{l}$ holds $V \ll \dot{\varphi}^2$ and $a(t)$ behaves as in the case of hydrodynamical matter, i.e. $p = +\varepsilon$, $a(t) \sim t^{1/3}$, but then diverges with growing t/φ and converges for φ being close to zero back towards the $|\varphi| = \frac{\sqrt{2}}{l}$ line. Between $y = \pm\frac{\sqrt{2}}{l}$ exists a *separatrix*. For $|\varphi| \gg \frac{1}{l}$ holds that the separatrix is close to $\frac{dy}{d\varphi} = 0$, so that

$$y = -\frac{1}{3l^2}\frac{V_{,\varphi}}{V} + \mathcal{O}\left(\left(\frac{V_{,\phi}}{V}\right)^2, \frac{V_{,\varphi\varphi}}{V}, \dots\right). \tag{3.55}$$

In the vicinity of this separatrix one gets $\frac{|\dot{H}|}{H^2} = \frac{3}{2}l^2y^2 \ll 1$ for $|\varphi| \gg \frac{1}{l}$, so that these solutions describe an inflationary period[10] with effective equation of state $p \approx -\varepsilon$. Note, that for exponential background expansion one gets a power-law inflation, which will come in handy in the application of the formalism on axions in the next section. If trajectories get close to the stationary line at $|\varphi| > \frac{1}{l}$ when $\Delta y \sim \mathcal{O}(1)$, so that $\dot{\varphi}_i^2(\varphi_i)\exp\{3\sqrt{2}|l\varphi_i|\}$, what are constraints on the initial conditions of inflation, a quasi de Sitter period establishes. The end of inflation is given when φ drops below the Planck scale at $1/l$. Then, φ begins to oscillate. In general, you get the expected $a(t) \sim t^{\frac{2}{3}}$ and $R \sim -\frac{4}{3}t$ behavior with additional correction terms due to the oscillation[11].

Let us now recast the EOM in a gauge-invariant form. For cosmological perturbations this can be achieved by inverting the corresponding $\delta_\nu^{(\mathrm{gi})\mu}$ in (3.35) by plugging in the energy-momentum tensor (3.45) and setting

[10] This is also called *quasi de Sitter solution*.

[11] One gets $a(t) \sim (t-t_0)^{\frac{2}{3}}\left(1 + \frac{\cos(2m(t-t_0))}{6m^2(t-t_0)^2} - \frac{1}{24m^2(t-t_0)^2} + \mathcal{O}((t-t_0)^{-3})\right)$ and $R = -\left[\frac{4}{3}(t-t_0)\right]\{1 - 3\cos(2m(t-t_0)) + \mathcal{O}((t-t_0)^{-1})$ for the correction terms.

$\Phi = \Psi$, i.e.

$$\nabla^2\Phi - 3H\Phi' - (H' + 2H^2)\Phi = \frac{3}{2}l^2(\varphi_0'\delta^{(\mathrm{gi})\prime} + V_{,\varphi}a^2\delta\varphi^{(\mathrm{gi})}), \tag{3.56}$$

$$\Phi' + H\Phi = \frac{3}{2}l^2\varphi_0'\delta\varphi^{(\mathrm{gi})}, \tag{3.57}$$

$$\Phi'' + 3H\Phi' + (H' + 2H^2)\Phi = \frac{3}{2}l^2(\varphi_0'\delta\varphi^{(\mathrm{gi})\prime} - V_{,\varphi}a^2\delta\varphi^{(\mathrm{gi})}). \tag{3.58}$$

By combining these, the EOM linearized over the background solutions are

$$\delta\varphi^{(\mathrm{gi})\prime\prime} + 2H\delta\varphi^{(\mathrm{gi})\prime} - \nabla^2\delta\varphi^{(\mathrm{gi})} + V_{,\varphi\varphi}a^2\delta\varphi^{(\mathrm{gi})} - 4\varphi_0'\Phi' + 2V_{,\varphi}a^2\Phi = 0, \tag{3.59}$$

where the latter two terms describe the gravitational fluctuations. Now, use $\delta\varphi^{(\mathrm{gi})}$ and obtain

$$\Phi'' + 2\left(H - \frac{\varphi_0''}{\varphi_0'}\right)\Phi' - \nabla^2\Phi + 2\left(H' - H\frac{\varphi_0''}{\varphi_0'}\right)\Phi = 0 \tag{3.60}$$

$$\Leftrightarrow \Phi'' + 2\left(\frac{a}{\varphi_0'}\right)'\left(\frac{\varphi_0'}{a}\right)\Phi' - \nabla^2\Phi + 2\varphi_0'\left(\frac{H}{\varphi_0'}\right)'\Phi = 0. \tag{3.61}$$

Further, use

$$u := \frac{a}{\varphi_0'}\Phi \tag{3.62}$$

to get

$$u'' - \nabla^2 u - \frac{\theta''}{\theta}u = 0 \tag{3.63}$$

with

$$\theta := \frac{H}{a\varphi_0'}. \tag{3.64}$$

In the asymptotic limit for short-wavelengths we get

$$u \sim e^{\pm ik\eta},\ k^2 \gg \frac{\theta''}{\theta}, \tag{3.65}$$

so that

$$\Phi \approx \left[C_1\sin\left(k\int a^{-1}dt\right) + C_2\cos\left(k\int a^{-1}dt\right)\right]e^{ikx}, \tag{3.66}$$

$$\delta\varphi^{(\mathrm{gi})} \approx \frac{2}{3l^2}\frac{k}{a}\left[C_1\cos\left(k\int a^{-1}dt\right) - C_2\sin\left(k\int a^{-1}dt\right)\right]e^{ikx}. \tag{3.67}$$

Likewise in the asymptotic limit for long-wavelengths we get

$$u \approx C_1\theta + C_2\theta \int \frac{d\eta}{\theta^2} = \frac{A}{\varphi_0'} \left(\frac{1}{a} \int^{\eta} a^2(\eta')d\eta' \right), k \ll \frac{\theta''}{\theta}, \tag{3.68}$$

so that

$$\Phi = \approx A \left(\frac{1}{a} \int adt \right)^{\cdot} = A \left(1 - \frac{H}{a} \int adt \right), \tag{3.69}$$

$$\delta\varphi^{(\text{gi})} \approx A\dot{\varphi}_0 \left(a^{-1} \int adt \right). \tag{3.70}$$

With these results we can define the following constant of motion

$$\zeta := \frac{2}{3} \frac{H^{-1}\dot{\Phi} + \Phi}{1+w} + \Phi, \tag{3.71}$$

where $w = p/\varepsilon$ is the known equation of state parameter. That this indeed is conserved can be checked straightforwardly since for λ outside r_H one can neglect $\nabla^2\Phi$.

As we have seen above, a period of inflation arises rather naturally, so let us now turn to inflationary Universe models in general. Say, there is a period of exponential expansion before reheating at η_{re} for which $p = -\varepsilon$ and $H^{-1} = r_H$ =const. holds. Then at η_{re} the vacuum energy density is converted into usual matter, i.e. massive particles and radiation, in $\Delta t < H^{-1}$, what we call one Hubble expansion. Further, at $\eta > \eta_{\text{re}}$ the Universe evolves as if it is in the radiation-dominated epoch till η_{eq} as usual. In the radiation- and matter-dominated epochs, r_H increases faster than a fixed cosmological scale, k. Thus, as we have discussed in appendix E, initial perturbations inside the horizon are allowed to be produced, which are then inflated away to scales of galaxies and clusters today if inflation lasted at least 62 Hubble expansions. Consider short wavelength perturbations, for which we get $\Phi \sim \dot{\varphi}_0$. Now, during inflation, when $\dot{\varphi}_0 \approx -\frac{V_{,\varphi}}{3H}$ holds, $\dot{\Phi}$ is negligible, so that for a quadratic potential one gets $\dot{\varphi}_0 \sim m$. Further, $|\delta\varphi_0| \sim a^{-1}$, so for perturbations with fixed comoving wavenumber, k, $|\delta\varphi_0|$ is large and eventually the linear theory breaks down as $\frac{\delta\varphi_0}{\varphi_0} < 1$ does not hold any longer. Say, enough e-foldings took place, then the mentioned perturbations with fixed k should now be considered as long-wavelength modes, described by (3.69), so that one obtains the asymptotic series

$$\Phi \approx A\left(a^{-1}\int a dt\right)^{\cdot} = A\left(H^{-1} - \int a(H^{-1})^{\cdot} dt\right)^{\cdot}$$
$$= A([H^{-1}]^{\cdot} - [H^{-1}[H^{-1}]^{\cdot}]^{\cdot} + H^{-1}[H^{-1}[H^{-1}]^{\cdot}]^{\cdot} - \dots), \qquad (3.72)$$
$$\delta\varphi^{(\mathrm{gi})} \approx A\dot{\varphi}_0(H^{-1} - H^{-1}[H^{-1}]^{\cdot} + H^{-1}[H^{-1}[H^{-1}]\cdot]^{\cdot} - \dots). \qquad (3.73)$$

During inflation $|\dot{H}| \ll H^2$ holds, so that only the first terms in the serieses are relevant. Further, $\dot{\varphi}_0 \ll V(\varphi)$ and $|\ddot{\varphi}| \ll |V_{,\varphi}(\varphi)|$ during inflation, what gives

$$\Phi \approx -A\frac{\dot{H}}{H^2} \quad \text{and} \quad \delta\varphi^{(\mathrm{gi})} \approx A\frac{\dot{\varphi}_0}{H}. \qquad (3.74)$$

After inflation the scalar field oscillates. Since $a(t) \sim t^m$ one gets according to (3.69) that $\Phi \approx \frac{A}{m+1}$ is time-independent. With the approximated solutions (3.74) we get

$$\Phi \approx \frac{1}{m+1} H \frac{\delta\varphi^{(\mathrm{gi})}}{\dot{\varphi}_0} \qquad (3.75)$$

evaluated at $k^2 = \theta''/\theta$, what in many cases occurs when $\lambda \gtrsim r_H$. Note,

$$\Phi(t_f) = \frac{1 + \frac{2}{3}[1 + w(t_f)]^{-1}}{1 + \frac{2}{3}[1 + w(t_i)]^{-1}}\Phi(t_i), \qquad (3.76)$$

i.e. the spectrum is nearly scale-invariant and solely depends on Δw.

3.1.4 Quantum mechanical description of scalar matter field matter perturbations

Before being able to lift the previously done description of scalar field matter perturbations to a quantum mechanical one, we first need to quantize the metric and matter fluctuations simultaneously, what requires a non-vanishing matter component, what is only possible in an expanding Universe. Luckily, this condition is naturally satisfied. As we already know, quantizing matter fields in a non-trivial background leads to particle production, which is the basis for structure formation in an inflationary Universe. We need canonical commutation relations and cannot just quantize the classical EOM since we only need to quantize physical degrees of freedom. Fluctuations are small, typically of Gaussian type, so the computation of metric and density perturbations reduces to the determination of the two-point-correlation functions and power spectra. The main physical observable that is connected with density perturbations is the root-mean-square relative mass function because the two-point correlation function can be converted to the relative density perturbations and they can in turn be converted to the root-mean-square relative mass functions. We will denote the power spectrum by $|\delta_\varepsilon(k)|^2$ and the two-point correlation function is given by

$$\xi_\varepsilon(r) = \frac{\delta\varepsilon}{\varepsilon}(x) \cdot \frac{\delta\varepsilon}{\varepsilon}(x+r) \tag{3.77}$$

where

$$\xi_\varepsilon(r) = 4\pi \int_0^\infty \frac{\sin(kr)}{kr} |\delta_\varepsilon(k)|^2 \frac{dk}{k}. \tag{3.78}$$

Our approach will be the following. We express the action in terms of a single gauge-invariant variable and quantize only this rather than quantizing multiple variables. The initial action is given by

$$S = -\frac{1}{16\pi G} \int R\sqrt{-g}d^4x + \int \mathcal{L}_m(g)\sqrt{-g}d^4x, \tag{3.79}$$

where the first term corresponds to curvature and gravity and the second term to matter. Now, one needs to expand this action to second order in perturbation variables since this will lead to the first order perturbation equations. For the purely gravitational part, $\delta_2 S_{\rm gr}$, of the action, $\delta_2 S$, one can do the calculation in the so-called ADM-formalism. We do not want to go deeper into it, but please refer to [2] and references therein for details. In the end we obtain

$$\begin{aligned}\delta_2 S_{\rm gr} = \frac{1}{16\pi G} \int \{ & a^2[-6(\psi')^2 - 12H(\phi+\psi)\psi' - 9H^2(\phi+\psi)^2 \\ & - 2\psi_{,i}(2\phi_{,i} - \psi_{,i}) - 4H(\phi+\psi)(B-E')_{,ii} + 4H\psi' E_{,ii} \\ & - 4\psi'(B-E')_{,ii} - 4H\psi_{,i}B_{,i} + 6H^2(\phi+\psi)E_{,ii} - 4HE_{,ii}(B-E')_{,jj} \\ & + 4HE_{,ii}B_{,jj} + 3H^2E_{,ii}^2 + 3H^2B_{,i}B_{,i}] + \mathcal{D}_1^{\rm gr} + \mathcal{D}_2^{\rm gr}\}d^4x, \end{aligned} \tag{3.80}$$

where $\mathcal{D}_{1,2}^{\rm gr}$ are total derivative terms not affecting the EOM as usual. Now, let us turn to the matter part of the action for a flat Universe. We use the background equations

$$H^2 = l^2\left[\frac{1}{2}(\varphi_0')^2 + V(\varphi_0)a^2\right], \quad 2H' + H^2 = 3l^2\left[-\frac{1}{2}(\varphi_0')^2 + V(\varphi_0)a^2\right], \tag{3.81}$$

which can be combined to

$$H^2 - H' = \frac{3}{2}l^2(\varphi_0')^2. \tag{3.82}$$

With these information at hand we can determine the matter part of the action (3.79), i.e.

$$\delta_2 S_m = \int d^4x \sqrt{-g_0} \left(\frac{\delta_2\sqrt{-g}}{\sqrt{-g_0}} \mathcal{L}_0 + \frac{2\delta_1\sqrt{-g}\delta_1\mathcal{L}}{\sqrt{-g_0}} + \delta_2\mathcal{L} \right), \tag{3.83}$$

where, as always, a subscript zero denotes the homogeneous background values and

$$\mathcal{L}(\varphi) = \frac{1}{2}\varphi_{,\mu}\varphi^{,\mu} - V(\varphi), \tag{3.84}$$

what one can Taylor expand to read off $\delta_{1,2}\mathcal{L}$. Combining the matter part (3.83) with the previously obtained gravitational part (3.80) one constructs the total action $\delta_2 S$ as we wanted to do. Now, one can vary this action with respect to $(B - E')$ to get the constraint

$$\psi' + H\phi = \frac{3}{2} l^2 \varphi_0' \delta\varphi. \tag{3.85}$$

Further, we can construct the gauge-invariant variable

$$v := a \left[\delta\varphi^{(\mathrm{gi})} + \frac{\varphi_0'}{H}\psi \right], \tag{3.86}$$

where $\delta\varphi^{(\mathrm{gi})} = \delta\varphi + \varphi_0'(B - E')$. With this gauge-invariant variable at hand we can rewrite

$$\delta_2 S = \frac{1}{2}\int \left((v')^2 - v_{,i}v_{,i} + \frac{z''}{z}v^2 + \frac{1}{3l^2}\sum_{i=1}^{4}\mathcal{D}_i \right) d^4x, \tag{3.87}$$

as the scalar field action in flat spacetime with time-dependent mass, $m^2 = -z''/z$, where $z := \frac{a}{H}\varphi_0'$ and $\mathcal{D}_i$ are total derivatives. We are set to perform the quantization. The canonically conjugate momentum to v is

$$\pi(\eta, x) = \frac{\partial\mathcal{L}}{\partial v'} = v'(\eta, x), \tag{3.88}$$

with one gets the Hamiltonian

$$\mathcal{H} = \int (v'\pi - \mathcal{L}) d^4x = \frac{1}{2}\int \left(\pi^2 + v_{,i}v_{,i} - \frac{z''}{z}v^2 \right) d^3x. \tag{3.89}$$

Now, we lift the variables v and π to operators $\hat{v}$ and $\hat{\pi}$ that satisfy the standard commutation relations

$$[\hat{v}(\eta, x), \hat{v}(\eta, x')] = 0 = [\hat{\pi}(\eta, x), \hat{\pi}(\eta, x')] \tag{3.90}$$

$$\text{and} \quad [\hat{v}(\eta, x), \hat{\pi}(\eta, x')] = i\delta(x - x'). \tag{3.91}$$

By varying the action (3.87) with respect to $\hat{v}$, we get the EOM for $\hat{v}$, namely

$$\hat{v}'' - \Delta\hat{v} - \frac{z''}{z}\hat{v} = 0, \tag{3.92}$$

which is equivalent to the Heisenberg equations

$$i\hat{v}' = [\hat{v}, \hat{\mathcal{H}}],\ i\hat{\pi} = [\hat{\pi}, \hat{\mathcal{H}}]. \tag{3.93}$$

We would like to work in the Heisenberg picture with time-dependent state vectors instead of time-dependent operators. Note, that $\hat{v}$ can be expressed in terms of annihilation and creation operators that satisfy the commutation relations

$$[\hat{a}_k, \hat{a}_k] = 0 = [\hat{a}_k^\dagger, \hat{a}_k^\dagger] \quad \text{and} \quad [\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'}, \tag{3.94}$$

so that we are now able to construct the Fock-representation of the Hilbert space, i.e. the space in which the for us interesting operators $\hat{v}$ and $\hat{\pi}$ act. For a free scalar field in flat spacetime with constant mass, m, exists a unique vacuum state, $|0\rangle$, defined by

$$\hat{a}_k|0\rangle = 0 \quad \forall k, \tag{3.95}$$

where in the mode expansion, the annihilation operators are the operator coefficients of the positive-frequency modes,

$$v_k(\eta) \sim e^{i\omega_k \eta}, \quad \text{with } \omega_k^2 = k^2 + m^2. \tag{3.96}$$

All other states can be obtained by acting on $|0\rangle$ with the proper combination of creation and annihilation operators. Note, that a unique vacuum with distinguished time direction gives a time-invariant notion of positive and negative frequency modes in contrast to what happens when quantizing the homogeneous component of a scalar field in an expanding Universe, where there is no definite notion of time, so there cannot be a unique vacuum state. Hence, start by picking a time, η_0. Now, find a linear combination of the two fundamental solutions

$$v_l(\eta_0) = \left(k_l^2 - \frac{z''}{z}\right)^{-\frac{1}{2}}_{\eta=\eta_0} \quad \text{and} \quad v_l'(\eta_0) = i\left(k_l^2 - \frac{z''}{z}\right)^{\frac{1}{2}}_{\eta=\eta_0} \tag{3.97}$$

of the time-dependent mode equation

$$v_l''(\eta) + \left(k_l^2 - \frac{z''}{z}\right) v_l(\eta) = 0. \tag{3.98}$$

if

$$\left(k_l^2 - \frac{z''}{z}\right) > 0 \quad \forall \text{modes } l. \tag{3.99}$$

A vacuum state can, again, be constructed by

$$a_l|0_{\eta_0}\rangle = 0 \quad \forall l, \text{ where } |0_{\eta_0}\rangle \equiv |\psi_0\rangle. \tag{3.100}$$

After some calculation one can obtain for the number operator of the l-th mode, $\hat{N}_l$

$$\langle\psi_0|\hat{N}_l^1|\psi_0\rangle = |\beta_l|^2, \tag{3.101}$$

meaning that the initial vacuum state gives a non-vanishing expectation value for the number operator at η_1, denoted by the superscript 1 at the number operator. This is only true if the positive (+) and negative (-) frequency modes are related by a *Bogoliubov-transformation*

$$v_l^{(1)+} = \alpha_l v_l^{(0)+} + \beta_l v_l^{(0)-}, \quad v_l^{(1)-} = \beta_l^* v_l^{(0)-}, \quad |\alpha|^2 - |\beta|^2 = 1, \tag{3.102}$$

from which one can derive the corresponding transformation of the creation and annihilation operators

$$\hat{a}_l = \alpha_l \hat{b}_l + \beta_l^* \hat{b}_l^\dagger \quad \hat{a}_l^\dagger = \beta_l \hat{b}_l + \alpha_l^* \hat{b}_l^\dagger, \tag{3.103}$$

where the creation and annihilation operators denoted by a $\hat{b}$ correspond to the mode expansion of $\hat{v}$. However, the non-vanishing expectation value of the number operator can be interpreted as follows. The observer at η_1 sees a non-vanishing number of particles in the vacuum state, $|\psi_0\rangle$. This is the process responsible for the production of initial perturbations in the early Universe in inflationary Universe models[12].

In scalar field Universe the two fundamental solutions (3.97) of the time-dependent mode equation are no longer applicable since during inflationary period

$$\frac{z''}{z} \approx \frac{a''}{a} > 0 \Rightarrow \left(k_l^2 - \frac{z''}{z}\right) \not> 0. \tag{3.104}$$

If this happens, one defines the so-called de-Sitter invariant vacuum given by the conditions

$$v_k(\eta_0) = \frac{1}{k^{3/2}}(H_0 + ik)e^{ik\eta_0}, \; v_k'(\eta_0) = \frac{i}{k^{1/2}}\left(H_0 + ik - i\frac{H_0'}{k}\right)e^{ik\eta_0}. \tag{3.105}$$

[12]Note, that the quantum theory sets the initial conditions for the classical considerations.

Note, that these conditions converge with the old solutions for $k \gg H_0$. The basic problem is now, that in general, for large k the leading terms in the expansions of the solutions agree[13], but not for small k, in which case the terms may depend sensitively on the definition of the vacuum. Luckily, we only need the short-wavelength, large k, part of the initial vacuum spectrum. One can use the following initial conditions.

$$v_k(\eta_0) = k^{-\frac{1}{2}} M(k\eta_0) \quad \text{and} \quad v_k'(\eta_0) i k^{\frac{1}{2}} N(k\eta_0), \tag{3.106}$$

where the normalization condition

$$NM^* + N^* M = 2, \quad |M(k\eta_0)| \to 1, \, |N(k\eta_0)| \to 1 \tag{3.107}$$

for $k\eta_0 \gg 1$ holds.

With this prework we are now ready to tackle the spectrum of density perturbations in inflationary Universe models with scalar field matter. First of all, start with the EOM (3.61)

$$\Phi'' + 2\left(\frac{a}{\varphi_0'}\right)'\left(\frac{\varphi_0'}{a}\right)\Phi' - \Delta\Phi + 2\varphi_0'\left(\frac{H}{\varphi_0'}\right)'\Phi = 0. \tag{3.108}$$

Note, that the EOM could have been derived directly after we have constructed the proper action, but the same EOM come out, so we spare us this alternative derivation. However, the significant difference is, that $\hat{\Phi}$ is an operator now, whose mode expansion reads

$$\hat{\Phi}(\eta, x) = \frac{1}{\sqrt{2}} \frac{\varphi_0'}{a} \int \frac{d^3k}{(2\pi)^{\frac{3}{2}}} \left[u_k^*(\eta) e^{ikx} a_k + u_k(\eta) e^{-ikx} a_k^\dagger\right]. \tag{3.109}$$

The EOM puts the constraint

$$u_k''(\eta) + \left[k^2 - \left(\frac{1}{z}\right)'' z\right] u_k(\eta) = 0 \tag{3.110}$$

on the mode functions $u_k(\eta)$. The Fourier mode coefficients functions, $u_k(\eta)$ and $v_k(\eta)$, are related via

$$u_k(\eta) = -\frac{3}{2} l^2 \frac{z}{k^2} \left(\frac{v_k}{z}\right)' \tag{3.111}$$

[13]Typically, one gets $v_k(\eta_0) \sim k^{-1/2}$ and $v_k'(\eta_0) \sim k^{1/2}$ for $k \to \infty$.

and one can use the initial conditions

$$u_k(\eta_i) = -\frac{3}{2}l^2\left(\frac{i}{k^{3/2}}N(k\eta_i) - \frac{z'(\eta_i)}{z(\eta_i)}\frac{1}{k^{5/2}}M(k\eta_i)\right) \tag{3.112}$$

$$u_k'(\eta_i) = -\frac{3}{2}l^2\left[\frac{1}{k^{3/2}}M(k\eta_i) + 3\frac{z'(\eta_i)}{z(\eta_i)}\left(\frac{i}{k^{3/2}}N(k\eta_i) - \frac{z'(\eta_i)}{z(\eta_i)}\frac{1}{k^{5/2}}M(k\eta_i)\right)\right], \tag{3.113}$$

which are extremely important asymptotic conditions because they tend to one for $k\eta_i \gg 1$ and ensure independence of the vacuum definition, what will become clear later. Now, consider the definition of the power spectrum of metric perturbations, $|\delta_k|^2$ as a measure of the two-point correlation function of $\hat{\Phi}$, i.e.

$$\langle 0|\hat{\Phi}(\eta, x)\hat{\Phi}(\eta, x+r)|0\rangle = \int_0^x \frac{dk}{k}\frac{\sin(kr)}{kr}|\delta_k|^2 \tag{3.114}$$

$$\Rightarrow |\delta_k(\eta)|^2 = \frac{1}{4\pi^2}\frac{(\varphi_0')^2}{a^2}|u_k(\eta)|^2k^3, \tag{3.115}$$

where the upper equation was plugged into the mode expansion (3.109). The solution of the constraint on the mode functions (3.110) in inflationary Universe models is

$$u_k(\eta) = \begin{cases} u_k(\eta_i)\cos[k(\eta-\eta_i)] + \frac{u_k'(\eta_i)}{k}\sin[k(\eta-\eta_i)] \\ \frac{u_k(\eta_i)\cos(k\eta_i) - u_k'(\eta_i)k^{-1}\sin(k\eta_i)}{\{(\varphi_0')^{-1}[a^{-1}\int a^2 d\eta]'\}_{\eta_H(\eta)}}(\varphi_0')^{-1}\left(\frac{1}{a}\int a^2 d\eta\right)' \end{cases}, \tag{3.116}$$

where the upper equation corresponds to the short-wavelength and the lower equation to the long-wavelength case, respectively, so that

$$|\delta_k| \approx \begin{cases} \frac{l}{4\pi}V_{,\varphi}/V^{\frac{1}{2}} & k_{\text{ph}} > H(t) \\ \frac{l}{4\pi}(V^{\frac{3}{2}}/V_{,\varphi})_{\eta_H(k)}\, V_{,\varphi}^2/V^2 & H(t) > k_{\text{ph}} > H_i\frac{a(t_i)}{a(t)} \end{cases}. \tag{3.117}$$

Typical examples to investigate are $V(\varphi) = \frac{\lambda}{n}\varphi^n$ and $V(\varphi) = \frac{1}{2}m^2\varphi^2$, for which I like to refer to [2]. Note, that initial fluctuations inside the horizon are well-described by the short-wavelength case in (3.116) and then grow out of the horizon, so that they are then well-described by the long-wavelength case in (3.116). However, after inflation, one approximately gets

$$|\delta_k| \approx \frac{3l^2}{4\pi(p+1)}\left(\frac{\dot{\varphi}_0 H^2}{\dot{H}}\right)_{t_H(k)} \approx \frac{3l^2}{2\pi(p+1)}\left(\frac{V^{\frac{3}{2}}}{V_{,\varphi}}\right)_{t_H(k)}, \tag{3.118}$$

which is directly connected to the CMB spectrum. A key result is that the spectrum of adiabatic perturbations is close to being scale-invariant with additional model-dependent logarithmic correction factors. However, From CMB observations, one knows $|\delta_k| \lesssim 10^{-5}$ on large scales, k. For the quadratic model, this gives $\frac{m}{m_{\mathrm{pl}}} < 10^{-6}$, what in turn gives $\lambda < 10^{-14}$ as a coupling constant. In order to achieve that quantum fluctuations from inflation do not produce too large density perturbations, there are either extremely small values of the coupling constant, λ, or a certain mass hierarchy exists.

Let us sum up the main results of our above considerations. We have a gauge-invariant description of metric perturbations power spectra caused by fluctuations of scalar field matter, on which we can apply a gauge that suits our problem under consideration the best. We see that metric perturbations always exist as we have seen that the non-trivial vacuum-structure gives naturally rise to particle production. Note, that the results do not depend on the choice of the vacuum state as long as basic initial asymptotic limit conditions apply. Further, inflation, i.e. its exponential expansion, increases the amplitude of the metric perturbations and the final power spectrum appears to be nearly scale-invariant.

3.2 The axion field

If not stated otherwise, we follow [1] closely throughout all the subsections in this section.

3.2.1 Aim and Setup

We now, finally, want to investigate the dynamics of initial density perturbations due to the for us most interesting scalar field matter, namely the axion field, ϕ. For this purpose, we would like to, first, set up the system of basic equations in this subsection and then go on to apply linear cosmological perturbation theory in subsection 3.2.2, what will be a more intense, but rather quick discussion, in which we would like to cover the application of different gauges for their corresponding purposes, discuss isocurvature perturbations, the sound speed in different cases, discuss transfer functions and the Halo-Mass functions. This will set a proper groundwork for the very brief discussion of the non-linear theory in subsection 3.2.3, in which we leave the extensively described linear cosmological perturbation theory, since we will realize that it fails for typical fluctuation sizes, so that we are naturally forced to deal with the non-linear effects in form of the Schrödinger-Poisson equation. Its implications are briefly discussed, mostly in the context of halo density profiles. Let us get started.

As was already mentioned in subsection 2.3.2, the action for a minimally coupled scalar field, like the axion field ϕ, is given by[29]

$$S_\phi = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right], \tag{3.119}$$

but note, that apparently, this holds only after the PQ-symmetry breaking, which makes sense since this is the necessary step to establish the axion as a pseudo-Goldstone boson and it is further only valid after non-perturbative effects apply since this is necessary to give mass to the axions[1]. However, By varying the action with respect to ϕ we get the EOM

$$\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial\nu) =: \Box\phi - \frac{\partial V}{\partial\phi} = 0 \tag{3.120}$$

and moreover, by varying the action with respect to $g^{\mu\nu}$ one gets the energy momentum tensor

$$T^\mu_\nu = g^{\mu\alpha}\partial_\alpha\phi\partial_\nu\phi - \frac{1}{2}\delta^\mu_\nu[g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + 2V(\phi)]. \tag{3.121}$$

Recall, that we use the RW-metric (D.4) for a flat Universe, $\kappa = 0 \Rightarrow S_0(r) = r$, along with the Hubble parameter, $H := \dot{a}/a$, so that the EOM reads

$$\ddot{\phi} + 3H\dot{\phi} + m_a^2\phi = 0, \tag{3.122}$$

where we assumed the potential to be of the form $V(\phi) = \frac{1}{2}m_a^2\phi^2$. Note, that this is precisely the same form as the EOM for a general scalar matter field as (3.48), so that we can apply the treatment of scalar matter field perturbations to the axion field, ϕ, from section 3.1. The RW geometry further dictates the background evolution of the density, $\rho_{a,0}$, and the pressure, $P_{a,0}$, i.e.

$$\rho_{a,0} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}m_a^2\phi^2 \quad \text{and} \quad P_{a,0} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}m_a^2\phi^2. \tag{3.123}$$

3.2.2 Application of cosmological perturbation theory

In section 3.1 we focused on keeping the discussion and the main equations in a gauge-invariant form, so that we can apply the proper gauges whenever they are needed for a smooth discussion and interpretation of the results. Applied to axions, the two gauges we discussed[14], namely the synchronous and longitudinal (Newtonian) gauge, respectively, where the former one comes in handy because $\theta = 0$ and the latter one is useful in the Newtonian limit. Note, that θ corresponds to the fluid description as follows

$$\theta := ik^i v_i. \tag{3.124}$$

Additionally, we already define σ in the same fashion to be

$$(\rho_0 + P_0)\sigma := -\left(\hat{k}^j \hat{k}_i - \frac{1}{3}\delta^j_i\right)\Sigma^i_j. \tag{3.125}$$

However, consider a perturbed axion field, $\delta\phi$, and denote, for the sake of simplicity, the background axion field simply by ϕ, then the EOM (3.34) become

$$\delta\phi'' + 2H\delta\phi' + (k^2 + m_a^2 a^2)\delta\phi = -\frac{1}{2}\phi' h' \tag{3.126}$$

in the synchronous gauge, where h is the corresponding gauge-potential discussed in subsection 3.1.3, and

$$\delta\phi'' + 2H\delta\phi' + (k^2 + m_a^2 a^2)\delta\phi = (\Psi' + 3\Phi')\phi' - 2m_a^2 a^2 \phi\Psi \tag{3.127}$$

in the Newtonian gauge, where Φ and Ψ are the corresponding gauge-potentials also discussed in subsection 3.1.3. Note, that a prime denotes the derivative with respect to conformal time, $d\tau = dt/a$, as usual. "If all cosmological perturbations are seeded by single field inflation, the initial conditions are *adiabatic*"[1], which is a useful information because this allows us to relate the overdensity in photons with the overdensity of any fluid component[15], i, via

$$\delta_i = \frac{3}{4}(1 + w_i)\delta_\gamma, \tag{3.128}$$

where $w_a \approx -1$ and $\delta\phi = \delta_a \approx 0$ in the early Universe. Recall, that for DM we expect $w_m = 0$ as was discussed in appendix A. This is not in conflict with $w_a \approx -1$ at early times since as we have discussed in subsection

[14]There are several so-called *Boltzmann-solvers* to numerically solve the relevant equations for us. One of them is called *CAMB*, which uses the synchronous gauge and another one is *AXIONCAMB*, which additionally has a full ALP-treatment implemented[1]. This will be of greater relevance in the follow-up project, so that we do not go deeper into them in this work.

[15]Recall, that the fluid description is valid, which was discussed in subsection 2.3.2 around equation (2.66).

3.1.3 we obvserve an oscillatory behavior for the axion field, ϕ, as one can expect from the EOM (3.120). Further, we found out that the field oscillates around $w_a = 0$, fitting to the expected DM behaviour, which is fixed at later times, when the oscillations decay since the axions start to cluster in the potential wells set by the photons in the radiation-dominated epoch. Hogan and Rees[34] discuss the formation of axion miniclusters from initial isocurvature perturbations[16], which is a typical problem one has to deal with in the unbroken PQ-symmetry case from section 2.2. Such an isocurvature perturbation between two species i and j can be written as

$$S_{ij} = 3(\zeta_i - \zeta_j), \tag{3.129}$$

where ζ_i is the curvature perturbation corresponding to species i,

$$\zeta_i := -\Psi - H\frac{\delta\rho_i}{\dot{\rho}_i}, \tag{3.130}$$

so that the total curvature perturbation can be described as

$$\zeta = \frac{\sum_i(\rho_i + P_i)\zeta_i}{\sum_i(\rho_i + P_i)}. \tag{3.131}$$

Note, that these equations are gauge-invariant. We can set initial conditions for $k \ll aH$ for all modes k and for $\tau \ll 1$. In the broken PQ-symmetry case, the only perturbed species are the axions with $\delta_a = 1$. Since the equations are linear, one can first solve the equations and the multiply the corresponding spectrum and normalization, where the spectrum is a typical power law with spectral index $(1 - n_I) = 2\varepsilon_{\text{inf}}$, with ε_{inf} defined in (E.27). Nevertheless, as we have already seen in subsection 3.1.3, the equation of state parameter, w_a, is time-dependent and satisfies the Friedmann equation (A.3)

$$\rho_a' = -3H\rho_a(1 + w_a). \tag{3.132}$$

With the equation of state parameter at hand, one gets the adiabatic background sound speed

$$c_{\text{ad}}^2 = w_a - \frac{w_a'}{3H(1 + w_a)}. \tag{3.133}$$

Note, that equation of state parameter and sound speed are the relevant initial conditions for the background equations. Note further, that one

[16] Think of isocurvature perturbations as perturbations in the relative number density of a particle species that leaves the total curvature unperturbed[1].

is now able to rewrite the EOM as two first order equations for δ_a and $u_a = (1+w_a)v_a$, namely the heat flux,

$$\delta_a' = -ku_a - \frac{(1+w_a)h'}{2} - 3H(1-w_a)\delta_a - 9H^2\frac{(1-c_{\rm ad}^2)u_a}{k} \tag{3.134}$$

$$\text{and} \quad u_a' = 2Hu_a + k\delta_a + 3H(w_a - c_{\rm ad}^2)u_a, \tag{3.135}$$

where we used $c_s^2 = 1$, what will become clear in a bit and $\delta\phi = 0$ as explained above. Later, when $a > a_{\rm osc}$, we already have observed the rapid oscillation of w_a. The same is true for $c_{\rm ad}^2$. Due to these rapid oscillations, one needs to make an approximation similar to $w_a = 0$ for a proper description and to spare numerical calculation costs. In the synchronous gauge, the EOM for the fluid description read

$$\delta' = -(1+w)(\theta + \frac{h'}{2}) - 3H(c_s^2 - w)\delta \tag{3.136}$$

$$\text{and} \quad \theta' = -H(1-3w)\theta - \frac{w'}{1+w}\theta + \frac{c_s^2}{1+w}k^2\delta, \tag{3.137}$$

since the anisotropic stress vanishes in the RW-geometry we assume due to the isotropy of space. The same equations in Newtonian gauge[35] read

$$\delta' = -(1+w)(\theta - 3\dot{\phi}) - 3H(c_s^2 - w)\delta \tag{3.138}$$

$$\theta' = -H(1-3w)\theta - \frac{w'}{1+w}\theta + \frac{c_s^2}{1+w}k^2\delta - k^2\sigma + k^2\Psi. \tag{3.139}$$

One can define the sound speed in perturbations by

$$c_s^2 := \frac{\delta P}{\delta \rho}. \tag{3.140}$$

From the Einstein equations

$$\delta\rho_a = \rho_a\delta_a, \tag{3.141}$$

$$\delta P_a = \rho_a\left[\delta_a + 3H(1-c_{\rm ad}^2)(1+w_a)\frac{u_a}{k}\right], \tag{3.142}$$

$$\text{and} \quad \rho_a(1+w_a)v_a = \rho_a u_a \tag{3.143}$$

we get

$$\delta P_a = \langle c_s^2\rangle_t \rho_a\delta_a \tag{3.144}$$

and after using a WKB-approximation as before, we can write

$$\phi = a^{-\frac{3}{2}}[\phi_+\cos(m_a t) + \phi_-\sin(m_a t)] \tag{3.145}$$

$$\text{and} \quad \delta\phi = \delta\phi_+(t,k)\cos(m_a t) + \delta\phi_-(t,k) - \sin(m_a t), \tag{3.146}$$

so that we can find for $\delta\phi = 0$ the effective sound speed

$$c^2_{\text{s,eff}} := \langle c_s^2 \rangle_t = \frac{\frac{k^2}{4m_a^2 a^2}}{1 + \frac{k^2}{4m_a^2 a^2}}, \tag{3.147}$$

which is crucial to describe the differences of structure formation in the ULA- and CDM-picture, respectively. However, with $c^2_{\text{s,eff}}$ at hand, we can rewrite the EOM (3.136) as

$$\delta_a' = -ku_a - \frac{h'}{2} - 3Hc^2_{\text{s,eff}}\delta_a - 9H^2 c^2_{\text{s,eff}}\frac{u_a}{k} \tag{3.148}$$

$$\text{and} \quad u_a' = -Hu_a + c^2_{\text{s,eff}}k\delta_a + 3c^2_{\text{s,eff}}H^2 u_a. \tag{3.149}$$

Let us now assume that axions make a significant amount of the observed DM content of the Universe. The Poisson equation in Newtonian gauge reads

$$k^2\Psi^2 = -4\pi G a^2 \rho\delta. \tag{3.150}$$

Using the Poisson equation and the EOM (3.138) in Newtonian gauge, one gets a single equation for δ, i.e.

$$\ddot{\delta}_a + 2H\dot{\delta}_a + \left(\frac{k^2 c^2_{\text{s,eff}}}{a^2} - 4\pi G\rho_a\right)\delta_a = 0. \tag{3.151}$$

Note, that we switched from conformal time to coordinate time[17]. Note additionally, that the mass term consists of a contribution due to pressure, $\frac{k^2 c_s^2}{a^2}$, and a counteracting contribution due to the density, $-4\pi G\rho_a$. Both contributions are to be understood during gravitational collapse, what becomes clear for $k^2 c^2_{\text{s,eff}} \to 0$ because obviously we then observe the domination of the density over the pressure term, i.e. a so-called *Jeans instability* causing the collapse. In the opposite case, the pressure term dominates over the density term and the fluctuations can oscillate without further growth. Note, that for $k^2 c_s^2 = 4\pi G\rho_a$, the pressure and density contributions are in equilibrium, which is why we denote this situation as the *axion Jeans scale*, what we can express as

$$k_J^2 = (16\pi G a \rho_{a,0})^{\frac{1}{2}} m_a = 66.5^2 a^{\frac{1}{2}} \left(\frac{\Omega_a h^2}{0.12}\right)^{\frac{1}{2}} \frac{m_a}{10^{-22}\,\text{eV}}\,\text{Mpc}^{-2}, \tag{3.152}$$

[17]Once again, I would like to highlight, that this is precisely the same EOM as for CDM with exactly the same growing, $\delta_a \sim a$, and decaying, $\delta_a \sim a^{-3/2}$, of perturbations.

where we evaluated $c^2_{\text{s,eff}}$ in the limit $k/(m_a a) < 1$ to obtain

$$c^2_{\text{s,eff}} \approx \frac{k^2}{4m_a^2 a^2}. \tag{3.153}$$

In the same limit and by considering the matter-dominated epoch with the known $\dot{\rho}_a = \rho_{\text{crit}} a^{-3}$ scaling, one can solve the EOM (3.151) for δ by

$$\delta_a = C_1 D_+(k,a) + C_2 D_-(k,a) \tag{3.154}$$

with the linear growth functions

$$D_+(k,a) = \frac{3a^{1/2}}{\tilde{k}^2} \sin\left(\frac{\tilde{k}^2}{a^{1/2}}\right) + \left(\frac{3a}{\tilde{k}^4} - 1\right) \cos\left(\frac{\tilde{k}^2}{a^{1/2}}\right) \tag{3.155}$$

$$D_-(k,a) = \left(\frac{3a}{\tilde{k}^4} - 1\right) \sin\left(\frac{\tilde{k}^2}{a^{1/2}}\right) - \frac{3a^{1/2}}{\tilde{k}^2} \cos\left(\frac{\tilde{k}^2}{a^{1/2}}\right), \tag{3.156}$$

where

$$\tilde{k} = \frac{k}{(m_a H_0)^{1/2}} \sim \frac{k}{k_J}. \tag{3.157}$$

Note, that for low $\tilde{k}$ one gets

$$D_+(k,a) \sim a \quad \text{and} \quad D_-(k,a) \sim a^{-3/2} \tag{3.158}$$

as we expected. This behavior changes with growing scales, $\tilde{k}$, so that at some intermediate scale, the actual behavior depends on time. At early times, there is a mixture of the typical power laws with oscillations whereas at late times, the oscillations decay and we observe, again, the power law behaviors alone. However, at large scales, i.e. above the Jeans scale, k_J, we obtain the expected oscillatory behavior with constant amplitude, i.e. no perturbation growth. Linear growth functions are especially useful because of the following phenomenology. In the standard cosmological ΛCDM model, growth is scale-independent, i.e. independent of k, for redshifts $z \lesssim \mathcal{O}(10^2)$ because the baryon acoustic oscillations have frozen in and the density parameter of radiation has become sufficiently small to be neglected. This gives us the opportunity to measure the current power spectrum at $z = 0$ and then modify it with a so-called *transfer-function*, $T_X(k,z)$, to a $z > 0$. In general, one can use transfer functions, which are also scale-dependent, so that

$$P_X(k,z) = T_X^2(k,z) P_{\Lambda\text{CDM}}(k,z) \tag{3.159}$$

holds for a model X, but as already mentioned, we can simply state $T^2_{\Lambda\mathrm{CDM}}(k,z) = T^2_{\Lambda\mathrm{CDM}}(z)$. Then, for the linear growth from $z = 0$ to a $z > 0$, we obtain

$$P_{\Lambda\mathrm{CDM}}(k,z) = \left(\frac{D_+(z)}{D_+(0)}\right)^2 P_{\Lambda\mathrm{CDM}}(k) \tag{3.160}$$

where

$$D_+(z) = \frac{5\Omega_m}{2H(z)} \int_0^{a(z)} \frac{H_0^3 da'}{a' H(a')}. \tag{3.161}$$

Further, we are now able to make a connection to WDM and DM composed by ULAs. Recall, that in subsection 2.3.1 we found that thermal axions decouple whilst being relativistic, so that they accouont for HDM causing structure formation suppression, which is not what we observe. However, besides axions, there are other theoretical particles that could have been decoupled at an intermediate temperature forming WDM, e.g. sterile neutrinos or gravitinos with $m_X \sim 1\,\mathrm{keV}$. The redshift-independent transfer functions

$$T_{\mathrm{WDM}}(k) = (1 + (ak)^{2\mu})^{-\frac{5}{\mu}} \quad \text{and} \quad T_{\mathrm{ULA}}(k) = \frac{\cos(x_J^3(k))}{1 + x_J^8(k)} \tag{3.162}$$

give good results[18] for WDM with $m_X \gtrsim 0.1\,\mathrm{keV}$ and for ULAs with $m_a \gtrsim 10^{-24}\,\mathrm{eV}$, respectively, where $\mu = 1.12$,

$$\alpha = 0.074 \left(\frac{m_X}{\mathrm{keV}}\right)^{-1.15} \left(\frac{0.7}{h}\right) \mathrm{Mpc}, \tag{3.163}$$

$$x_J(k) = 1.61 \left(\frac{m_a}{10^{-22}\,\mathrm{eV}}\right)^{\frac{1}{18}} \frac{k}{k_{J,\mathrm{eq}}} \tag{3.164}$$

$$\text{with } k_{J,\mathrm{eq}} = 9 \left(\frac{m_a}{10^{-22}\,\mathrm{eV}}\right)^{\frac{1}{2}} \mathrm{Mpc}^{-1}. \tag{3.165}$$

Note, that the four parameters above are fitting parameters for the numerical simulations done by [1]. By considering the transfer functions in more detail, one observes that T_{WDM} is a power-law, what makes sense since WDM is composed of thermally produced particles with $T \sim a^{-1}$ in general and comoving scales of order of the horizon when the temperature is roughly of the same order as m_X. Further, one sees that the transfer function, T_{ULA}, for ULAs describes the same behavior as we discussed

[18]The calculations were made under the basic assumption that all the DM content is composed by a single DM species.

above, namely that the Jeans scale, $k_{J,\text{eq}}$, sets the scale when the exponential growth transforms to an oscillation.

The power spectrum itself is interesting to know for the description of galaxy formation. The Halo-Mass Function (short: HMF) is given by

$$\frac{dn}{d\ln(M)} = -\frac{1}{2}\frac{\rho_m}{M}f(\nu)\frac{d\ln(\sigma^2)}{d\ln(MI'}, \tag{3.166}$$

with $\nu := \frac{\delta_{\text{crit}}}{\sigma}$, where $\sigma^2(M,z)$ is the *variance of fluctuations*, $\delta_{\text{crit}}(M,z)$ is the linearly extrapolated critical density and M is the halo mass. The HMF "gives the expected number of halos per logarithmic mass bin, per unit volume, for a given cosmology"[1]. For $f(\nu)$ one can use the so-called *Sheth-Tormen* function

$$f(\nu) = A\sqrt{\frac{2q}{\pi}}\nu(1+(\sqrt{q}\nu)^{-2p})\exp\left\{-\frac{q\nu^2}{2}\right\} \tag{3.167}$$

with $A = 0.3222$, $p = 0.3$ and $q = 0.707$. Further, one defines

$$\sigma^2(M,z) := \frac{1}{2\pi^2}\int_0^\infty \frac{dk}{k}(k^3P(k,z))^2W^2(k|R(M)), \tag{3.168}$$

where W is a window-function that is used to smooth the power spectrum, i.e.

$$W(k|R) = \frac{3}{(kR)^3}(\sin(kR) - kR\cos(kR)), \quad \text{with } M = \frac{4}{3}\pi\rho_m R^3. \tag{3.169}$$

We see, that the power spectrum gives the variance and so we already have one of the two HMF parameters. The second one is obtained as follows. Consider a $\Lambda = 0$, i.e. an Einstein-de Sitter Universe. Then, one can solve spherical collapse analytically obtaining a constant, mass-independent critical density for scale-independent growth as we have discussed above, that can be scaled to any redshift by

$$\delta_{\text{crit}}(z) = \frac{1.686\,D_+(0)}{D_+(z)}. \tag{3.170}$$

One can replace the scale factor fraction with a scale-dependent growth factor, $G(M,z)$, i.e.

$$\delta_{\text{crit}}(M,z) = 1.686\,G(M,z), \tag{3.171}$$

to allow for $\Lambda \neq 0$, where G imposes a direct connection between the typical CDM-model we want to obey and the axion model by

$$G(k,z) = \frac{\delta_a(k_0,z)\delta_a(k,z_{\text{early}})}{\delta_a(k,z)\delta_a(k_0,z_{\text{early}})} \cdot \frac{\delta_{\text{CDM}}(k_0,z)\delta_{\text{CDM}}(k,z_{\text{early}})}{\delta_{\text{CDM}}(k,z)\delta_{\text{CDM}}(k_0,z_{\text{early}})}. \tag{3.172}$$

Note, that $k_0 < k_J(z_{\text{early}})$ and $\Omega_a = \Omega_{\text{CDM}}$ should be satisfied, where in the former condition we should note additionally, that if k_0 is too small, then Λ becomes the dominant density amount and in the latter condition, z_{early} should be chosen, so that BAO have been frozen in already. Marsh states, that for "DM axions in a close-to-ΛCDM cosmology, reasonable choices are $k_0 = 0.002h\,\text{Mpc}^{-1}$ and $z_{\text{early}} \approx 300$[1]. With the second parameter of the HMF at hand, one can now investigate the HMF in detail. The main result is, that at low M and at high z, ULAs suppress halo formation compared to CDM. Recall from section 2.2 that in the unbroken PQ-symmetry case, there is the possibility that axion miniclusters form. For QCD-axions, these miniclusters satisfy $M \approx 10^{-9} M_\odot$. These miniclusters could be denser than halos if they are more massive than corresponding halos and thus, would be of interest to observe today because after fulfilling these conditions, they should have been able to exist till today. For QCD-axions we also observe a cut-off in the HMF at $M < 10^{-9} M_\odot$ as a result of the Jeans scale.

3.2.3 Brief treatment of the non-linear theory

Up till now we were working in linear theory and now want to make a first step in the non-linear theory[19], that will be a crucial foundation for the follow-up project. However, let us consider non-relativistic velocities, which is reasonable since virial velocities in galaxies are typically, $v_{\text{vir}} \ll c$ and additonally, it is reasonable to work in the Newtonian limit since $\Psi \ll 1$ everywhere except when one is close to a BH. Note, that Ψ must satisfy the Poisson equation (3.150). Consider only wavelengths above the axion Compton wavelength, then

$$\Box = -(1-2\Psi)(\partial_t^2 + 3H\partial_t) + \frac{1}{a^2}(1+2\Psi)\nabla^2 - 4\dot{\Psi}\partial_t \tag{3.173}$$

holds to leading order in Ψ. Further,

$$\rho_a = \frac{1}{2}\left[(1-2\Psi)\dot{\Phi}^2 + m_a^2\phi^2 + \frac{1}{a^2}(1+2\Psi)\partial^i\phi\partial_i\phi\right] \tag{3.174}$$

[19]Note, that $\delta_{\text{galaxies}} \gtrsim \mathcal{O}(10^5)$, so that we cannot perform perturbation theory, obviously. We are simply forced to take an alternative approach in this case.

and since ϕ oscillates in the early Universe, we make a WKB-approximation

$$\phi = \frac{1}{m_a\sqrt{2}} \left(\psi e^{-im_a t} + \psi^* e^{im_a t} \right), \tag{3.175}$$

where ψ is a complex scalar field. In the limits mentioned above, we can write

$$\Psi \sim \varepsilon^2_{\text{non-rel.}}, \quad \frac{k}{m_a} \sim \varepsilon_{\text{non-rel.}} \quad \text{and} \quad \frac{H}{m_a} \sim \varepsilon_{\text{WKB}} \tag{3.176}$$

and work to $\mathcal{O}(\varepsilon^2_{\text{non-rel,WKB}})$, so that $\rho_a = |\psi|^2$, what in turn recasts the Poisson equation (3.150) to the form[20]

$$\nabla^2\Psi^2 = 4\pi G a^2 |\psi|^2 \delta, \tag{3.177}$$

so that one clearly sees how Ψ is generated out of $|\psi|^2$. All in all, by using the d'Alembertian (3.173) instead of the familiar one in the Schrödinger equation, we arrive, after applying the Poisson equation above at the *non-linear Schrödinger-Poisson equation*

$$i\dot{\psi} - \frac{3}{2} iH\psi + \frac{1}{2m_a a^2}\nabla^2\psi - m_a\Psi\psi = 0, \tag{3.178}$$

namely the EOM for the complex scalar field ψ. Let us write $\psi = R\exp\{iS\}$ and plug it in, then we find the velocity

$$\vec{v}_a := \frac{1}{m_a a}\nabla S \tag{3.179}$$

of an effective fluid. By splitting the actual fluctuations from the background, one finds the EOM in terms of δ_a

$$\dot{\delta}_a + \frac{\vec{v}}{a}\nabla\delta_a = -\frac{1+\delta_a}{a}\nabla\cdot\vec{v}_a, \tag{3.180}$$

$$\text{and} \quad \dot{\vec{v}}_a + \frac{\vec{v}_a\cdot\nabla}{a}\vec{v}_a = -\frac{\nabla(\Psi+Q)}{a} - H\vec{v}, \tag{3.181}$$

where we defined the quantum potential

$$Q := -\frac{1}{2m_a^2 a^2}\frac{\nabla^2\sqrt{1+\delta_a}}{\sqrt{1+\delta_a}}, \tag{3.182}$$

[20] We write the Poisson equation in coordinate-space here. To go to Fourier space, we simply replace $\nabla \to -ik$ as usual.

which is in fact the only model parameter necessary for the axion gradient energy and the Jeans scale because it introduces an additional term in the force equation on a fluid element, i.e.

$$F = -\frac{\nabla(\Psi + Q)}{a} \tag{3.183}$$

as was seen in the EOM above. Via this effective fluid description, one carefully can connect the linear perturbation theory with non-linear simulation tools, which typically rely on fluid descriptions. Like in the previous subsection, where we dealt with the linear theory, let us take a quick look on halos. From pure CDM N-body simulations one knows that the *Navarro-Frenk-White* (short: NFW) density profile seems to be universal and has the shape

$$\frac{\rho_{\mathrm{NFW}}(r)}{\rho_{\mathrm{crit}}} = \frac{\delta_{\mathrm{NFW}}(c,r)}{\frac{r}{r_s}\left(1+\frac{r}{r_s}\right)^2}, \tag{3.184}$$

where c is the *halo concentration*, defined so that for the virial radius $r_{\mathrm{vir}} = cr_s$ holds and r_s is the scale radius. ULAs have a direct impact on the density profile due to their wave-like effects we discussed above, what implies that we cannot just simply consider CDM, but have to take an ULA modification into account. However, we expect that the resulting density profile gives the NFW profile again if we smooth it over sufficiently many Jeans scales. One typcially considers so-called *soliton*-solutions of the EOM (3.120) with stationary wave and constant energy, E, solutions

$$\psi = \chi(r)e^{-iEt} \tag{3.185}$$

since this system possesses the scaling symmetry

$$(r, \chi, \Psi, E, M(<r), \rho) \rightarrow \left(\frac{r}{\lambda}, \lambda^2\chi, \lambda^2\Psi, \lambda^2 E, \lambda M(<R), \lambda^4\rho\right), \tag{3.186}$$

where λ is the scale factor, $\rho = \chi^2$ is the soliton density and $M(<r)$ is the soliton mass inside r. Due to the scaling symmetry, one has to solve the resulting system of differential equations after plugging in the stationary wave ansatz in just once. Marsh finds that the initial condition $\chi(0) = 1$ gives $E_0 = -0.692$ for the zeroth energy-eigenvalue[1]. By considering the scaling symmetry, one can find the characteristic radius, r_{sol}, for the soliton solution to be

$$r_{\mathrm{sol}} \sim m_a^{-\frac{1}{2}}\rho_{\mathrm{sol}}^{-\frac{1}{4}}, \tag{3.187}$$

which is the same scaling as the jeans scale, what is expected because they are derived in the non-relativistic limit after dimensional analysis. With r_{sol} one can write down the soliton density profile

$$\rho_{\mathrm{sol}}(r) = \frac{\rho_{\mathrm{sol}}(0)}{\left(1+\left(\frac{r}{r_{\mathrm{sol}}}\right)^2\right)^8}, \quad \text{with } r_{\mathrm{sol}} = 22\left(\frac{\rho_{\mathrm{sol}}(0)}{\rho_{\mathrm{crit}}}\right)^{-\frac{1}{4}}\left(\frac{m_a}{10^{-22}\,\mathrm{eV}}\right)^{-\frac{1}{2}}\ \mathrm{kpc}. \tag{3.188}$$

Apparently, the soliton density must match the NFW profile continuously after smoothing as explained above, what is a current problem to solve. Nevertheless, rather generically, one can simply write

$$\rho(r) = \theta(r_\varepsilon - r)\rho_{\mathrm{sol}}(r) + \theta(r - r_\varepsilon)\rho_{\mathrm{NFW}}(r) \tag{3.189}$$

with the familiar Heavyside-function, $\theta(r)$. With the density profile at hand, one can now match observational data to fix the axion mass for instance.

Chapter 4

Summary

In the following I would like to sum up what we have done throughout this work. I will give two equivalent summaries in English and German.

4.1 In English

We basically told two stories in this work. First, we had two different problems, namely the dark matter question in the standard cosmological ΛCDM model, described in appendix A, and the strong-CP problem in QCD, described in subsection 2.1.1. For a brief summary of the for us interesting results of QCD, see appendix B. In the ΛCDM model we seek for a proper DM candidate, whose properties are constrained by a variety of cosmological observations and numerical simulations. Although, we usually try to fit one of the known particles of the standard model of particle physics to such problems, there is apparently no appropriate candidate to solve the CDM problem. However, by solving the strong-CP problem of QCD, one of the key components of the standard model of particle physics, we find a naturally occuring particle in subsection 2.1.2, the axion, whose properties fit beautific to the properties of the desired CDM. The basic idea of the solution is that the vacuum structure is non-trivial and that different vacua exist, between which tunneling processes should be possible in some way, which then leads to instantons describing these effects, which are manifestly non-perturbative. We discussed the properties of instantons at length in appendix C. This is crucial since the global $U(1)_{PQ}$-symmetry is not exact, so that the spontaneous symmetry breaking establishes the axions as pseudo-Goldstone bosons, whereas the non-perturbative effects give rise to a potential that leads to a non-vanishing axion mass, m_a. We then distinguished the class of QCD-axions coming, obviously, from the QCD story we told so far and the ALPs in general, that are the generalization of axions in the context of GUTs, like string theory. Nevertheless, our analysis of the QCD-axions was sufficient to impose general conditions on the properties of ALPs. We discussed this topic along with three different QCD-axion models in subsection 2.1.3. After we have established the idea of ALPs, we discussed the axion field, ϕ, as an observer field during inflation in section 2.2, where we had to distinguish between the breaking of the PQ-symmetry during or after inflation, what leads to different phenomenology. We briefly touched upon inflation driven by axions, but did not go deeper into that, since this leads to the loss of axions as DM candidates, which is not what we want. A brief review of the idea of inflation was done in appendix E. To close the setup of axions in chapter 2, we considered four different production mechanisms. We started with the thermal production in subsection 2.3.1, that led to the production of axions as HDM or dark radiation, so that we rule this production channel out for our work. A very detailed derivation of the thermally produced axion abundance was done in appendix D. Then we went on to the production via misalignment in subsection 2.3.2, that perfectly fits our initial ideas of different vacuum states, that have to be realigned. This production channel is, in opposite to the thermal production, strictly non-thermal, what underlines, how well suited it appears to be for a reasonable initial axion DM population on cosmological scales. Lastly, we just briefly discussed the production of ALPs as a deacay product of a heavier parent particle and of topological defects (strings) in subsection 2.3.3, but since this production mechanism is far beyond the scope of this work and well-suited in a string theory setup, we kept the discussion at a minimum.

In the second part of the work, we focused mainly on two great review articles to present the fundamental dynamics of the initial density perturbations. Therefore chapter 3 starts with the extensive discussion of linear cosmological perturbation theory in section 3.1 with the presentation of the main aspects of [2], who have derived all relevant equations in a manifestly gauge-invariant form, giving us a maximum amount of flexibility to continue our work. In subsection 3.1.2, we have seen how conventional hydrodynamical perturbations behave in different cases, mostly to give us an impression of the conventional theory and to get used to the machinery of the formalism. We then quickly went over to scalar matter field perturbations in subsection 3.1.3, first, in the context of a classical theory, that already gave insights in the natural appearance of a period of inflation and the fundamental dynamics of the scalar field, what after inflation begins to oscillate. Although, we carefully have to distinguish between short- and long-wavelength perturbations, respectively, we found out, that the spectrum of density perturbations approximately does not depend on the exact dynamics during inflation rather than solely on the difference in the equation of state parameters between the end and start of inflation, giving rise to a nearly-scale invariant spectrum. Afterwards, in subsection 3.1.4, we turned to the task of lifting the classical theory to a quantum mechanical level. Whilst doing so, we saw that again, it cannot hold anymore, that the vacuum has a trivial structure, what immediately leads to the existence of particles in an initially empty vacuum state, when observed at some later time. We thus, independently of the strong-CP problem solution come back to the crucial idea of non-trivial vacuum states. However, we then went on to describe density perturbations in inflationary Universe models. With this groundwork, we are set to discuss the axion field as the most interesting example of a scalar matter field, following the example of [1] in section 3.2. Again, after setting up the basic framework in subsection 3.2.1, we turned to small initial fluctuations in subsection 3.2.2, that are well-described by linear cosmological perturbation theory. We investigated the dynamics of isocurvature perturbations and discussed the sound speed, what us then again led to the idea of an oscillation axion field after inflation. Note, that the sound speed is used to describe the differences between a ULA-DM and a CDM cosmology. The results match exactly with the known picture of appendix A. We then considered transfer function, which base on linear growth functions, that also give the same scaling laws as in the ΛCDM picture. However, this discussion can be used to describe the differences between the power spectra of different cosmologies, like CDM, ULA-DM and as an interesting opponent, even WDM. This discussion was closed by an investigation of the HMF. However, nowadays we typically observe large fluctuations, so that the linear perturbation theory is no longer applicable. Hence, we are forced to consider the non-linear theory, which is luckily for us, a quick outcome of a very few basic assumptions that are justified well by observation. They immediately lead to the Schrödinger-Poisson equation that underlines the wavelike character of the axions in a non-relativistic setup and its impact on gravity and thus, structure formation. We then went on to use the wavelike behavior to discuss the shape of halo density profiles, that are closely related to N-body simulations, what leaves us with an ideal starting point to pick up work in the next project.

4.2 In German

Grundlegend haben wir in dieser Arbeit zwei Geschichten erzählt. Zuerst haben wir zwei Probleme vorliegen, nämlich die Frage, was Dunkle Materie (kurz: DM) im kosmologischen ΛCDM Standardmodell ist, was wir in Anhang A beschrieben haben. Das zweite Problem ist das starke CP-Problem der Quantenchromodynamik (kurz: QCD), welches wir in Unterabschnitt 2.1.1 dargestellt haben. Für eine kurze Zusammenfassung der wichtigsten Ergebnisse der QCD siehe Anhang B. Im ΛCDM Standardmodell suchen wir nach einem geeigneten DM Kandidaten, wessen Eigenschaften durch eine Vielzahl kosmologischer Beobachtungen und numerischer Simulationen eingeschränkt sind. Obwohl wir normalerweise im Standardmodell der Teilchenphysik nach einem geeigneten Kandidaten suchen, scheint es so, als sei keiner der bekannten Standardmodell-Teilchen ein geeigneter Kandidat. Daher widmen wir uns zunächst dem Lösen des starken CP-Problems der QCD. Auf ganz natürliche weise, kommt das Axion in Unterabschnitt 2.1.2 als Teilchen bei der Lösung des starken CP-Problems heraus, dessen Eigenschaften erschreckend gut auf die der verlangten kalten DM passen. Die grundlegende Idee ist, dass das Vakuum eine nicht-triviale Struktur aufweist und dass es im Grund verschiedene Vakuumszustände gibt, zwischen denen auf bestimmte Art und Weise getunnelt werden kann, die wiederum durch Instantonen beschrieben werden, die manifest nicht-perturbativ[1] sind. Wir haben die Eigenschaften der Instantonen in Länge in Anhang C diskutiert. Das ist sehr wichtig, denn die globale $U(1)_{PQ}$-Symmetrie is nicht exak, so dass die spontane Brechung das Axion als Pseudo-Golstone Boson etabliert, wohingegen die nicht-perturbativen Effekte erst das Potential aufwerfen, aus welchem die Axionen-Masse abgeleitet werden kann. Nachfolgend haben wir die Klasse der QCD-Axionen von der der axionartigen Teilchen (kurz: ALP für "axion like particles") unterschieden. Erstere umfasst logischerweise Axionen, die aus der QCD resultieren, während letztere die Verallgemeinerung aus den großen vereinheitlichen Theorien, wie zum Beispiel der String-Theorie, ist. Nichtsdestotrotz war unsere Analyse der QCD-Axionen ausreichend, um allgemeine Eigenschaften von ALPs zu formulieren. Das gesamte Thema haben wir in Unterabschnitt 2.1.3 diskutiert. Nachdem

[1] Die direkte Übersetzung von perturbation aus dem Englischen ist *Störung*, was allerdings zu sehr abstrusen Formulierungen wie "nicht-störerisch" führt, die ich hier durch den generischen Gebrauch des eingedeutschten perturbativ vermeiden werde. Eine Ausnahme hierbei stellt die Störungsrechnung/-theorie im Allgemeinen dar, da dies der geläufige Ausdruck in der deutschsprachigen Literatur ist und wir von diesem nicht abweichen wollen.

wir die Idee der ALP etabliert hatten, haben das Axionenfeld, ϕ, als Beobachterfeld der Inflation in Unterabschnitt 2.2 betrachtet, wo wir zwischen der Brechung der PQ-Symmetry während und nach der Inflation unterscheiden mußten, was zu unterschiedlichen Phänomenologien führt. Wir haben nur kurz das Axionenfeld als den eigentlichen Treiber der Inflation betrachtet, wofür wir allerdings das Axion als DM Kandidaten eintauschen müßten, was uns widerstrebt und wir diesen Pfad daher nicht weiter beschritten. Ein kurzer Überblick über Inflationstheorie wurde in Anhang E gegeben. Um Kapitel 2 über die grundlegende Axionen-Theorie zu schließen, haben wir vier verschiedene Produktionsmechanismen betrachtet. Wir starteten in Unterabschnitt 2.3.1 mit der thermalen Produktion, die allerdings zu heißer DM und dunkler Strahlung führte, weshalb wir diesen Produktionskanal ausschlossen. Eine sehr detaillierte Herleitung der thermalen Axionenpopulation ist in Anhang D gegeben. Dann haben wir die Produktion durch Falschausrichtung des Vakuums in Unterabschnitt 2.3.2 betrachtet, was perfekt auf unsere ursprünglichen Ideen einer nicht-trivialen Vakuumsstruktur passt. Dieser Produktionskanal ist im Gegensatz zu der thermalen Produktion strikt nicht-thermal, was unterstreicht, wie gut dieser für eine nennenswerte Axionen-DM Population im frühen Universum geeignet ist. Abschließend haben wir noch das Axion als Zerfallsprodukt eines schwereren Elternteilchens bzw. eines toplogischen Defekts (Strings) in Unterabschnitt 2.3.3 betrachtet, was allerdings weit ausserhalb der Reichweite dieser Arbeit liegt und typischer Bestandteil in String-Theorien ist, weshalb wir die Diskussionen auf einem Minimum gehalten haben.

Im zweiten Teil der Arbeit fokussierten wir uns auf zwei sehr empfehlenswerte Artikel, die einen Überblick über die Dynamiken der jeweiligen ursprünglichen Dicht-Perturbationen geben. Daher startet Kapitel 3 mit einer ausgedehnten Diksussion der linearen kosmologischen Störungstheorie in Abschnitt 3.1 mit einer Prs̈entation der wichtigsten Ergebnisse von [2], die alle Gleichungen in eichinvarianter Form hergeleitet haben, was uns ein maximales Maßan Flexibilität bietet, um unsere Arbeit fortzusetzen. In Unterabschnitt 3.1.2 betrachten wir erst einmal konventionelly hydrodynamische Perturbationen in verschiedenen Fällen, hauptsächlich, um uns einen ersten Eindruck der Maschinerie des Formalismus zu geben. Danach gehen wir in Unterabschnitt 3.1.3 zu massiven Skalarfeld Perturbationen über, erst einmal in einem klassichen Kontext, was uns bereits das natürliche Auftreten von Inflation und die fundamentale Dynamik aufzeigt, nämlich, dass das Axionenfeld nach dem Ende der Inflation zu oszillieren beginnt. Obwohl wir vorsichtig zwsichen kurz- und langwelligen Perturbationen unterscheiden müssen, finden wir heraus, dass das Spektrum der Dichte-Perturbationen approximativ nicht von der genauen Dynamik während der Inflation, sondern viel mehr von der Differenz der Zustandsgleichungsparameter von Ende zu Beginn der Inflation abhängt. Danach haben wir uns in Unterabschnitt 3.1.4 der Aufgabe gewidmet, die klassische Theorie auf eine quantenmechanische Theorie zu erheben. Im Zuge dessen sahen wir erneut, dass die Idee von der trivialen Vakuumsstruktur nicht länger aufrecht erhalten werden kann. So kommen wir unabhängig der starken CP-Problem Lösung wieder bei der wichtigen Idee der nicht-trivialen Vakuumsstruktur an. Wir konnten dann Dichte-Perturbationen in inflationären Universumsmodellen beschreiben. Mit diesem Grundstein, waren wir bereit das Axionenfeld als das interessanteste massive Skalarfeld zu betrachten, wobei wir dem Beispiel von [1] in Abschnitt 3.2 gefolgt sind. Wieder haben wir zuerst die Rahmenbedingungen für die weiteren Analysen in Unterabschnitt 3.2.1 geschaffen. Danach haben wir uns zuerst kleinen Perturbationen in Unterabschnitt 3.2.2 gewidmet, die sehr gut durch lineare kosmologische Störungstheorie beschrieben werden. Wir haben die Dynmik von Iso-Krümmungs Perturbationen untersucht und die Schallgeschwindigkeit diskutiert, die uns wieder zum Oszillationsverhalten des Axionenfelds nach Ende der Inflation führte. Man beachte, dass die Schallgeschwindigkeit benutzt wird, um Unterschiede zwischen ULA-DM und CDM Kosmologien aufzuzeigen Die Ergebnisse stimmen mit dem bekannten Bild aus Anhang A überein. Dann haben wir die Transferfunktionen betrachtet, die auf linearen Wachstumsfunktionen beruhen, die ebenfalls das gleiche Skalierungsverhalten wir im ΛCDM Modell widergeben. Diese Diskussion kann auch benutzt werden, um die Unterschiede zwischen den Leistungsspektren verschiedener Kosmologien aufzuzeigen. Dieser Abschnitt wurde durch die Untersuchung von Halo-Massen-Funktionen abgeschlossen. Heutzutage beobachten wir allerdings typischerweise große Perturbationen, so dass die lineare Störungstheorie nicht mehr anwendbar ist und wir gezwungen werden, uns mit der nicht-linearen Theorie zu beschäftigen, welche recht schnell aus nur wenigen grundlegenden Annahmen folgt, die gut durch Beobachtungen gerechtfertig sind, und direkt die Schrödinger-Poisson Gleichung ergeben, die den Wellencharakter der Axionen in einem nicht-relativistischen Rahmen und ihren Einfluss auf Gravitation und die daraus resultierende Strukturformation widerspiegelt. Darauf aufbauend haben wir die Form von Halo-Dichteprofilen untersucht, die eng mit N-Körper Simulationen verknüpft sind, was uns einen idealen Startunkt für das nachfolgende Projekt bietet.

Chapter 5

Conclusion and Outlook

One can conclude that the axion appears to be a naturally arising and very suitable candidate to solve the DM question in ΛCDM cosmology. Its fundamental shift-symmetry protects its mass to all orders in perturbation theory. Its coupling to standard model particles shows that it couples very weakly to them since all coupling constants scale with powers of the axion mass, which thought to be extremely small for ULAs. The axion can be produced on cosmologically relevant scales in the early Universe via the misalignment of the non-trivial vacuum states, what naturally explains the existence of initial density perturbations. The evolution of these perturbations is well-described by linear cosmological perturbation theory at first, but later on must be traded for the non-linear treatment via the Schrödinger-Poisson equation, when the perturbations become to large. Nevertheless, the results match pretty well the observed results and standard cosmological theory, so that in fact, their existence should be observable.

Marsh[1] states, that one should pay particular attention to large scale structure measurements since "ULAs suppress structure formation on cluster scales"[1] and as we have briefly treated in the context of thermal production, thermally produced axions could contribute to HDM and dark radiation. Additionally, the direct detection experiments of axions should be rethought to allow for new approaches since the conventional methods make rather slow progress in scanning the huge possible mass range of m_a. Finally, there should also be a huge effort in numerical simulations. A vast investigation of CDM and WDM models exists already in great detail, but it lacks of a proper treatment of low-mass cases, like the ULA mass range. In order to simulate CDM, one just needs the initial conditions and the dynamics of the initial density perturbations. We are basically set to redo the simulations with modified initial conditions, e.g. the suppressed power spectrum, and modified dynamics, e.g. the impact of wavelike behavior. This is going to be our task in the upcoming work.

Chapter 6

Acknowledgments

I would like to thank my supervisor, Dr. Subodh Patil, very much for the guidance through the project and the many short discussion sessions in your office. Additionally, I would like to thank my second supervisor, Dr. Matthieu Schaller, for accepting the proposal to make two joint projects to allow for the opportunity to go even deeper into the material than just scratching the surface on two distinct topics.

Me gustaría agradecer a mi novia Ellie, qui eres el principal apoyo para mí en tiempos difíciles.

Hiermit gilt mein Dank auch meiner Familie, auf die stets bei Engpässen Verlass ist und die auch nicht vor den hohen Portokosten zurückscheut, um Lebensmittelvorräte an mich zu senden, damit ich auch ja keinen Hunger zu fürchten brauche.

Auch gilt mein Dank der Altherrenrunde, die auf ganz unterschiedliche Arten und Weisen einen Lichtblick an das Ende eines langen Tages bringen kann. Sei es in Form von sehr wichtigen Pfadi-Videkonferenzen, StaFü-Besprechungen, kleinen Atmen-Spaziergängen am Abend, Krankenhausfahrten, Doppelkopfabenden oder vielen mehr, ihr seid nicht in wenige Worte zu fassen.

Duizend dank wil ik ook graag aan twee van mijn Nederlandse vrienden zeggen. Twee van de eerste Nederlanders die ik ontmoette waren Hugo and Vince. Het is interessant om te weten dat we toevallig allemaal dezelfde verjaardag hebben. Ik ben heel blij dat ik met Hugo bijna elke week een lekker koppje koffie mag drinken en altijd iemand heb om te praten. Hetzelfde klopt voor Falco met wie ik samen studeer en ook het bureau deel. Dankzij hem had ik vaak minder problemen mijn studiemateriaal te vergaren en had ook vaak de mogelijkheid lange interessante conversaties te hebben.

I would like to thank the University Leiden for this awesome opportunity to pursue my studies here. It is an ongoing pleasure to gain knowledge and learn from the very best in their fields.

Zum Schluss möchte ich dem Ev. Studienwerk Villigst meinen tiefsten Dank aussprechen, ohne dessen Unterstützung dieses Vorhaben unter keinen erdenklichen Umständen elegant umsetzbar gewesen wäre. Vielen lieben Dank für das, was ihr mir ermöglicht.

Appendix A

Brief review of the ΛCDM model

This appendix is dedicated to the standard model of cosmology, namely the ΛCDM model, whose properties we would like to present in a condensed way. For a more detailed story on the historical development I recommend chapters 1.4.2. and 1.4.4. in [36] and the references therein for the scientific details.

However, one could start the story with Albert Einstein who published his theory of general relativity in 1916 and studied solutions of the corresponding field equations[36].

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{A.1}$$

Interestingly, he immediately found that all the solutions require an expanding or contracting Universe, but since he was convinced of a static Universe, he introduced a cosmological constant, Λ, to the field equations

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{A.2}$$

which solely purpose is to compromise the gravitational attraction, so that the Universe he studied can be static[36]. Note, that this is allowed since $\nabla^{\mu} T_{\mu\nu}$ does not change since $\Lambda = \text{const.}$ and metric compatibility, $\nabla^{\mu} g_{\mu\nu} = 0$, is a fundamental metric property in general relativity. Note, that these static solutions are unstable, what we would like to demonstrate quickly. First, from Einsteins Field equations (A.1) and the fact that the cosmological principle dictates that $T_{\mu\nu}$ can only be time-dependent, since the Universe is isotropic in space, so that we can write it as $T_{00} = \rho(t), T_{i0} = 0, T_{ij} = -P(t)g_{ij}$, where $\rho(t)$ is the density and $P(t)$ is the pressure as usual, we can derive

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{\kappa}{R_0^2}\frac{1}{(a/a_0)^2} \quad \textit{Friedmann eq.} \text{ and} \tag{A.3}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \quad \textit{Acceleration eq.}, \tag{A.4}$$

where R_0 is the present-day curvature of the Universe and $\kappa \in \{-1, 0, +1\}$ depending on if we describe a negatively curved, flat or positively curved Universe, respectively, a is the scale factor, where $a_0 = 1$ is its present-day value and a dot denotes the derivative with respect to cosmological time t. Now, the first law of thermodynamics is

$$dQ = dE + PdV. \tag{A.5}$$

The cosmological principle implies that the net energy transfer between neighboring infinitesimal elements of the Universe is zero, $dQ = 0$. Note that $\rho c^2 \overset{c=1}{=} \rho$ gives an energy density, where $\rho \sim a^{-3}$, so that ρV should give an energy since $V \sim a^3$ is a volume. Collecting these information the first law of thermodynamics becomes

$$0 = dE + PdV = d(\rho V) + PdV = Vd\rho + \rho dV + PdV.$$
$$\overset{:\rho V}{\Rightarrow} 0 = \frac{d\rho}{\rho} + \frac{dV}{V} + \frac{P}{\rho}\frac{dV}{V} = \frac{d\rho}{\rho} + \left(1 + \frac{P}{\rho}\right)\frac{dV}{V}. \tag{A.6}$$

Further,

$$V \sim a^3 \Rightarrow \frac{dV}{V} = 3\frac{da}{a} \Rightarrow 0 = \frac{d\rho}{\rho} + 3\left(1 + \frac{P}{\rho}\right)\frac{da}{a}$$
$$\overset{\cdot\rho/dt}{\Rightarrow} \dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0 \quad \textit{Fluid eq.}, \tag{A.7}$$

which we will use now. Recall that the equation of state is $P = w\rho$, where w is the equation-of-state parameter. We have $w_m = 0$ for non-relativistic matter, $w_\gamma = \frac{1}{3}$ for relativistic matter, respectively. From (A.7) we see that the second term vanishes if $P = -\rho$ or, equivalently, if $w = -1$. Then we get $\dot{\rho} = 0$ and since the energy density is constant we associate this case with he cosmological constant, i.e. $w_\Lambda = -1$. Note, that from (A.1) we can read off $\rho_\Lambda = \frac{\Lambda}{8\pi G}$, which is indeed constant. We are thus led to say that the cosmological constant is related to *dark energy*. Now, we rewrite the Friedmann equation (A.3) to include the energy density contribution of the cosmological constant explicitly, leading to

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{\kappa}{R_0^2}\frac{1}{a^2} + \frac{\Lambda}{3}, \tag{A.8}$$

which is possible since ρ is just the sum of all energy density contributions and now we exclude the dark energy contribution from the sum and make its constant contribution explicit. The same holds for pressure, which is why we do the same for the acceleration equation (A.4)

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3} \tag{A.9}$$

For a static Universe, we demand $\dot{a} = 0 = \ddot{a}$ and from observations we know that today non-relativistic matter is the dominant energy density contribution, so that we use $w_m = 0$ or likewise $P = w\rho = 0$. Plugging everything into the expanded acceleration equation we get

$$0 \stackrel{!}{=} \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho_m + \frac{\Lambda}{3}, \tag{A.10}$$

where we safely neglected relativistic matter as explained. From this we see the instability of a static Universe solution immediately because if the Universe is expanded (compressed) just infinitesimally, ρ_m decreases (increases) slightly, since $\rho_m \sim V^{-1}$, leading to $\ddot{a}$ becoming positive (negative) leading to a non-vanishing $\dot{a}$ or, equivalently, to a Universe expanding (compressing). The expanding behaviour of the Universe was later observed by Hubble so that a static Universe is ruled out[36]. Note, that Einstein included Λ in (A.1) to be able to describe a static Universe, but as we discussed above, even though the Universe is expanding, Λ itself is related to a constant energy density ρ_Λ, namely the dark energy density. In fact, for all $w < -\frac{1}{3}$ the acceleration equation (A.4) yields an accelerating expansion for which we call the corresponding energy density collectively as dark energy, but the $w_\Lambda = -1$ case is in that sense special that it corresponds to a constant energy density. However, *dark* implies that right now we do not understand the physics behind dark energy even though there are several experiments, described for instance in [37], that put observational constraints on the dark energy density parameter, $\Omega_{\Lambda,0} = \frac{\rho_{\Lambda,0}}{\rho_{\text{crit},0}}$, where $\rho_{\text{crit}} = \frac{3H^2}{8\pi G}$ is the critical density and H is the Hubble parameter. The subscript 0 denotes, as usual, the present-day value. Today we approximately observe $\Omega_{\Lambda,0} \approx 0.7$.

Let us turn to the next building block of the ΛCDM model, namely the *Cold Dark Matter*, CDM, part. There are several experiments, described in [37], that lead to the existence of dark matter. Historically relevant is the first evidence studied by Zwicky in 1933, where he studied velocities of galaxies in the Coma Cluster with the outcome that they cannot be explained solely by the observed matter content of the galaxies[36]. A second evidence, which gives a nice visualization, comes from observing X-rays from merging galaxy clusters and simultaneously measuring the weak gravitational lensing effect. The X-rays emitted by the colliding hot gas shows the expected shock formed by the two colliding clusters. Additionally, the weak gravitational lensing shows two massive clusters passing through each other, forming a huge offset related to the colliding hot gas. This offset can be explained by the existence of non-baryonic dark matter haloes in the galaxy clusters, where the dark matter is observed to be collisionless in order for the haloes to pass through each other as was observed. The natural question arising from this observations is what the dark matter actually is. I do not want to go deeper into this discussion, but the main results are that no known particle in the standard model of particle physics is an appropriate candidate and so the field is open to all kinds of speculations[38], like the axions we discuss in this project. However, the observed dark matter has properties we should consider now, so that we can align our theoretical considerations with these observational constraints to be consistent with experiment. First of all, thinking of dark matter as particles, it is relevant to know when the dark matter particles decoupled from the thermal bath. If they decoupled whilst being relativistic, like the first propositions of dark matter particles being massive neutrinos suggested[36], they would have high kinetic energies allowing them to wipe out small-scale structure in the early Universe, which leads to top-down structure formation which is inconsistent with large-scale structure observations[37]. That leaves us with dark matter being cold or warm, both able to give the observed structure formation from small to large scales. Numerical simulations and observations, e.g. cosmic shear[37] based on gravitational lensing, tend to agree perfectly with the predictions of cold dark matter[36], so we go along with them and adapt to a cold dark matter (CDM) model. From observations we know that $\Omega_{m,0} \approx 0.3$, but the baryonic matter contributes only $\Omega_{b,0} \approx 0.05$. This gives a baryon-matter-ratio $\Omega_{b,0}/\Omega_{m,0} \approx 0.16$, so that one sees immediately that most of the Universes matter is present in the form of non-baryonic dark matter. Recall that the baryonic matter consists of stars, interstellar matter and mostly the diffuse intergalactic medium[39].

Appendix B

Brief review of QCD

In this appendix we want to quickly go over the for us most important results of *Quantum Chromodynamics*, QCD, the quantum field theoretical description of the strong interaction. We will construct the QCD-Lagrangian $\mathcal{L}_{\text{QCD}}$ and discuss the flavour- and chiral-symmetry, respectively. For a way more detailed derivation of these results and far more information, I strongly recommend the reader to the standard textbook of Quantum Field Theory by Peskin and Schröder[12]. Here, I will follow the example of Münster [3] for a brief review of QCD.

The building blocks of hadrons, e.g. protons and neutrons, are *quarks* that are described by Dirac-fields, q, which is why quarks have a fermionic nature. Note, that one is typically familiar with denoting Dirac-fields by ψ, but for clarity we refer to quark-fields directly as q, which can be thus written as

$$q = \begin{pmatrix} q_1 \\ \vdots \\ q_4 \end{pmatrix} = (q_\alpha), \quad \alpha = 1, \ldots, 4. \tag{B.1}$$

There are six *quark-flavours*, $N_f = 6$, categorized in the three duplets *up* and *down*, *charm* and *strange*, *top* and *bottom*, so that we have to expand our previous quark-notation

$$q_\alpha \to q_{\alpha f}, \tag{B.2}$$

where $f = u, d, c, s, t, b$ denotes the corresponding flavours up, down, charm, strange, top, bottom, respectively. Since quarks are fermions, in order to satisfy Pauli's exclusion principle, we finally have to introduce a third quantum number, the *colour*, denoted by $i = 1, 2, 3$ or $i = r, g, b$ for the basic colours *red*, *green* and *blue*, respectively. These colours, that are described by colour-triplets, are responsible for the characteristic SU(3)-symmetry group of QCD. Most of the times, we do not need to include all three quantum numbers at the same time in our considerations. Now that we have set up the quark fields, let us introduce their masses, m_f, and assume that different quark flavours carry different masses. Let us construct the quark Langrangian, $\mathcal{L}_{\text{quark}}$, that can contain only lorentz-invariant couplings, as usual, and must satisfy the above described SU(3) gauge symmetry. So for free quarks we can immediately write down

$$\mathcal{L}_{\text{quark}} = \bar{q}(i\gamma^\mu \partial_\mu - m_f)q, \tag{B.3}$$

where a bar denotes, as usual, the Dirac adjoint, i.e. the composition of transposition and complex conjugation of a Dirac spinor. We denote anti-quarks with this notation. Let us check if $\mathcal{L}_{\text{quark}}$ is invariant under SU(3). For that, recall that for all $U \in SU(3)$ the identity $U^\dagger U = \mathbb{1} = UU^\dagger$ holds, so that the colour transformation

$$q_{if} \to q'_{if} = U_{ij} q_{if}, \tag{B.4}$$

where i and j are colour-indices, indeed gives

$$\mathcal{L}_{\text{quark}} \to \mathcal{L}'_{\text{quark}} = \bar{q}'_{if}(i\gamma^\mu \partial_\mu - m_f)q'_{if} = U^\dagger_{ij}\bar{q}_{if}(i\gamma^\mu \partial_\mu - m_f)U_{ij}q_{if} \tag{B.5}$$

$$= \bar{q}_{if}(i\gamma^\mu \partial_\mu - m_f)q_{if} = \mathcal{L}_{\text{quark}}, \tag{B.6}$$

where we used that U commutes with all operators, acting on flavour- and spinor-indices. Having the gauge group SU(3) at hand we can explore the transformation (B.4) a bit more. Note, that we can write all $U \in SU(3)$ as

$$U(x) = \exp\{-iA^a(x)T_a\}, \tag{B.7}$$

where $A^a(x)$ are eight parameters and T^a are the eight generators of the SU(3) symmetry group. Eight is the number of choice as in general SU(N) has $N^2 - 1$ generators and for each generator we get a corresponding parameter. Specifically for QCD we can identify $A^a_\mu(x)$ as the eight *gluon* fields, describing the exchange particles of the strong interaction, and we can further identify $T_a = \lambda_a/2$, where λ_a are the Gell-Mann matrices, a set of eight traceless hermitian matrices corresponding to the Lie-algebra of SU(3), which satisfy the commutation relation $[\lambda_a, \lambda_b] = 2f_{abc}\lambda_c$, where the f's are structure constants of the algebra. Knowing all that we can replace ∂_μ in (B.3) by

$$\partial_\mu \to D_\mu = \partial_\mu - igA^a_\mu(x)T_a, \tag{B.8}$$

where g is a coupling constant. With this replacement we construct the local SU(3) symmetry. Further, we can construct the field-strength tensor, $G^a_{\mu\nu}$, of gluons. For this, we omit the explicit spacetime-dependence of the gluon-fields and compute

$$\begin{aligned}
[D_\mu, D_\nu] &\overset{(B.8)}{=} \left[\partial_\mu - igA^a_\mu \frac{\lambda_a}{2}, \partial_\nu - igA^b_\nu \frac{\lambda_b}{2}\right] \\
&= \left[\partial_\mu, \partial_\nu\right] - ig\partial_\mu A^b_\nu \frac{\lambda_b}{2} + ig\partial_\nu A^a_\mu \frac{\lambda_a}{2} - (ig)^2 A^a_\mu A^b_\nu \left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] \\
&= -ig\partial_\mu A^b_\nu \frac{\lambda_b}{2} + ig\partial_\nu A^a_\mu \frac{\lambda_a}{2} + g^2 A^a_\mu A^b_\nu \frac{1}{4} \cdot 2f_{abc}\lambda_c \\
&=: -igG^a_{\mu\nu}\frac{\lambda_a}{2},
\end{aligned} \tag{B.9}$$

where in the end we just relabeled some indices and factored out some factors to define the gluon-field-strength tensor, i.e.

$$G^a_{\mu\nu}(x) = \partial_\mu A^a_\nu(x) - \partial_\nu A^a_\mu(x) - gf_{abc}A^b_\mu(x)A^c_\nu(x). \tag{B.10}$$

From this form it is immediately clear that it transforms properly under SU(3) flavour-transformations (B.4). With the gluon field-strength tensor we can construct the Yang-Mills Lagrangian

$$\mathcal{L}_{\text{YM}} := \frac{-1}{4}G^a_{\mu\nu}G^{\mu\nu,a} = -\frac{1}{2}Tr\left[G_{\mu\nu}G^{\mu\nu}\right], \tag{B.11}$$

which is basically the kinetic term of the gluons. We have rewritten the Yang-Mills Lagrangian in terms of a trace for usage in the main text of this project. Adding the free quark Lagrangian (B.3) to the Yang-Mills Lagrangian (B.11) then gives us the desired QCD-Lagrangian

$$\mathcal{L}_{\text{QCD}} = \bar{q}(i\gamma^\mu\partial_\mu - m_f)q - \frac{1}{2}Tr\left[G_{\mu\nu}G^{\mu\nu}\right]. \tag{B.12}$$

Let us now discuss some interesting symmetries of this theory. We begin with the already introduced **flavour symmetry**, given by (B.4). From experiment we know that the quark masses differ from each other, so that m_f in (B.12) is in fact a mass matrix

$$m_f \to M := \text{diag}(m_u, m_d, m_c, m_s, m_t, m_b). \tag{B.13}$$

In order to keep up the global flavour symmetry (B.4) with the free quark Langrangian, we demand that the mass matrix (B.13) and the transformation matrix (B.7) commute, which is no longer given, since this requires $m_f = m$ for all flavours f. Depending on what flavours one is considering one can construct an approximate flavour symmetry since the up- and down-quarks have rather similiar masses and are light compared to the other quarks and the strange-quark is heavy compared to the other quarks, but this is of no interest for us. In the case that we would have a global flavour symmetry, we would find the corresponding conserved Noether-currents to be

$$j_a^\mu(x) = \sum_{f,f'} \bar{q}_f(x)\gamma^\mu(T_a)_{ff'}q_{f'}(x), \tag{B.14}$$

which we call *vector-current*, since it properly transforms as a Lorentz-vector.

We continue by considering a **chiral symmetry**, which is given by the axial transformation

$$q_f \to q'_f = [\exp(-i\omega^a T_a\gamma_5)]_{ff'}q_{f'}. \tag{B.15}$$

Infinitesimally we find

$$\delta q = -i\delta\omega^a T_a\gamma^5 q \text{ and } \delta\bar{q} = -i\delta\omega^a T_a\bar{q}\gamma^5, \tag{B.16}$$

what we use along with the gamma matrix identity $\{\gamma^\mu, \gamma^5\} = 0$ to get the infinitesimal forms of the two terms of the free quark Lagrangian (B.3)

$$\delta(\bar{q}i\gamma^\mu\partial_\mu q) = \delta\omega^a\bar{q}(\gamma^\mu + \gamma^5 + \gamma^5\gamma^\mu)T_a\partial_\mu q = 0, \tag{B.17}$$

$$\delta(\bar{q}Mq) = -i\delta\omega^a\bar{q}\gamma^5(MT_a + T_aM)q, \tag{B.18}$$

where we see that once again because of the massive quarks, M is not commuting with the generators T_a and thus the mass term is non-vanishing and so for massive quarks we find no axial symmetry. Let us construct, more generally, the *chiral symmetry group* with the chiral projectors

$$P_L := \frac{1-\gamma^5}{2} \text{ and } P_R := \frac{1+\gamma^5}{2}, \tag{B.19}$$

where the subscripts L and R mean left and right, respectively. They satisfy the projector conditions

$$P_{L,R}^2 = P_{L,R},\, P_LP_R = 0 = P_RP_L \text{ and } P_L + P_R = \mathbb{1}. \tag{B.20}$$

With the projectors, we can seperate the quarks in left- and right-handed quarks, simply by applying the projectors to the general quark fields, i.e.

$$q = q_L + q_R, \text{ where } q_L = P_L q \text{ and } q_R = P_R q. \tag{B.21}$$

Note, that

$$\bar{q}_L = \bar{q} P_R \text{ and } \bar{q}_R = \bar{q} P_L \tag{B.22}$$

due to the fact that $\{\gamma^\mu, \gamma^5\} = 0$ and that transposition is expressed by multiplying with γ^0. With these relations and $(\gamma^5)^2 = \mathbb{1}$ we immediately can write down

$$\bar{q}_L \gamma^\mu q_R = 0 = \bar{q}_R \gamma^\mu q_L, \tag{B.23}$$

so that

$$\bar{q} i \gamma^\mu D_\mu q = \bar{q}_L i \gamma^\mu D_\mu q_L + \bar{q}_R i \gamma^\mu D_\mu q_R, \tag{B.24}$$

$$\bar{q} q = \bar{q}_L q_R + \bar{q}_R q_L. \tag{B.25}$$

We now can rewrite the original axial transformation (B.15) as a chiral transformation simply by replacing ω^a with ω^a_L for left-handed and ω^a_R for right-handed quarks, respectively. The corresponding symmetry groups are just $\mathrm{SU}(N_f)_L$ and $\mathrm{SU}(N_f)_R$, which are both isomorphic to $\mathrm{SU}(N_f)$. For vanishing quark masses we thus obtain a chiral symmetry with $2(N_f^2 - 1)$ parameters under $\mathrm{SU}(N_f)_L \otimes \mathrm{SU}(N_f)_R$. In the special case of $\omega^a_L = \omega^a_R = \omega^a$ we recover the axial transformation and are thus led to say that the flavour symmetry is the diagonal subgroup of the chiral symmetry group. Again, quarks are massive, but one can consider quark subsets and thus an approximate chiral symmetry. This is particularly interesting for the chiral symmetry transformations because by considering the up- and down-quark subset one observes a spontaneous symmetry breaking from the massless symmetry group $\mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R$ to the flavour symmetry subgroup SU(2), so that we expect three massless Goldstone-bosons, which we find to be the three pions. In reality, even the up- and down-quarks carry non-negligible masses and so in fact we have some additional explicit symmetry breaking as described above in case of massive quarks and thus, the pions carry some mass as well and are then called pseudo Goldstone-bosons.

Recall that for $\omega^a_L = \omega^a_R = \Omega^a$ we would recover the flavour transformations (B.4) from the axial transformation (B.15). If we instead took the infinitesimal form of the axial transformation separated in a left- and right-handed part, we can define $\omega^a_L = -\omega^a_R =: \omega^a$. Let us plug this in the infinitesimal axial transformation to get

$$\begin{aligned} q' = (1 - i\omega^a T_a)q &= (1 - i(\omega_L^a P_L + \omega_R^a P_R)T_a)q \\ &= (1 - i(-\Omega^a)(P_L - P_R)T_a)q \\ &= (1 - i\Omega^a \gamma^5 T_a)q, \end{aligned} \tag{B.26}$$

where in the last line we used $P_L - P_R = -\gamma^5$, which can be seen immediately from the definition of the projectors (B.19). Now, recall from (B.14) that we can write the vector-current in the form $j^\mu = \bar{q}\gamma^5 T_a q$. The additional γ^5 we just found in the infinitesimal axial transformation leads us thus to the corresponding conserved Noether-currents

$$j^\mu_{5,a} = \bar{q}\gamma^\mu\gamma^5 T_a q = j^\mu_{R,a} - j^\mu_{L,a}, \tag{B.27}$$

which be call *axial-current* since they properly transform as axial-vectors under Lorentz-transformations.

Appendix C

Selected properties of instantons

The purpose of this chapter is simply to explicitly compute, show or motivate some selected properties of instantons. I.e. we want to compute the winding number

$$\nu = \frac{1}{32\pi^2} \int d^4x (F, \tilde{F}) \,, \tag{C.1}$$

for a given field configuration, expressed by the field-strength tensor F and its dual $\tilde{F}$, the expectation value of e^{-HT} for a Hamiltonian H and total time T in a so-called θ-vacuum $|\theta\rangle$

$$\langle\theta|e^{-HT}|\theta\rangle \sim \int \mathcal{D}A e^{-S} e^{i\nu\theta} \,, \tag{C.2}$$

the θ-vacuum action

$$S_\theta = \frac{\theta}{32\pi^2} \int d^4x (F, \tilde{F}) \equiv \theta\nu \,, \tag{C.3}$$

the one-instanton action

$$S_0 = \frac{8\pi^2}{g_{\mathrm{G}}^2} \,, \tag{C.4}$$

where g_G is the coupling constant of the gauge group G and finally, the energy of a θ-vacuum

$$E(\theta) \sim cos(\theta) e^{-S_0}. \tag{C.5}$$

If not stated otherwise, we closely follow [40] throughout this chapter even though we differ slightly in notation and rearrange the given arguments.

C.1 Winding number ν

Let us consider the gauge group SU(2)={ $M \in \text{Mat}(2,\mathbb{C})|MM^\dagger = \mathbb{1} = M^\dagger M, \det M = 1$ }, which is a non-abelian compact Lie group. Since all $M \in$SU(2) has the form

$$M = \begin{pmatrix} a_0 + ia_3 & a_2 + ia_1 \\ a_2 - ia_1 & a_0 - ia_3 \end{pmatrix}, \tag{C.6}$$

where $a_{0,...,3} \in \mathbb{R}$ and $|a_0|^2 + ... + |a_3|^2 = 1$, we can use the Pauli matrices $\sigma_{1,2,3}$ to rewrite M in the form

$$M = a_0\mathbb{1} + i(a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3) =: a_0\mathbb{1} + i\vec{a}\cdot\vec{\sigma}\,, \tag{C.7}$$

where in the last equation we defined the vectors $\vec{a}$ and $\vec{\sigma}$ as a shorthand notation with $|a_0|^2 + |\vec{a}|^2 = 1$. We see that topologically, SU(2) is S^3={ $(a_0, ..., a_3) \in \mathbb{R}|\, |a_0|^2 + ... + |a_3|^2 = 1$ } and we have to study homotopy classes of mappings S^3 →S^3. For this class of mappings we can immediately write down some standard mappings

$$g^{(0)}(x) = 1 \quad \text{(trivial mapping)}, \tag{C.8}$$

$$g^{(1)}(x) = \frac{a_0 + i\vec{a}\cdot\vec{\sigma}}{r}, r := |\vec{a}|^2 \quad \text{(identity mapping)}, \tag{C.9}$$

$$g^{(\nu)}(x) = \left[g^{(1)}(x)\right], \nu \in \mathbb{Z} \quad \text{(family of mappings)}. \tag{C.10}$$

We call ν the *winding number*, which measures how often the hypersphere is wrapped around G. To get a better intuition for this, let us consider U(1)= $\{z \in \mathbb{C}|\, |z| = 1\}$ for a moment. We are familiar with this gauge group from electromagnetism and U(1) describes the unit circle in the complex plane, which is topologically S^1 and thus, we consider mappings S^1 →S^1. The standard way of parametrizing a circle is in polar coordinates (r, θ), where $r = 1$ is fixed for S^1 and $\theta \in [0, 2\pi)$, where we explicitly exclude 2π to ensure that a function $g : [0, 2\pi) \to S^1$ is single-valued, i.e. is periodic in θ with $g(\theta + n \cdot 2\pi) = g(\theta)$, where $n \in \mathbb{Z}$. Again, we can immediately write down some standard mappings

$$g^{(0)}(x) = 1 \quad \text{(trivial mapping)}, \tag{C.11}$$

$$g^{(1)}(x) = e^{i\theta} \quad \text{(identity mapping)}, \tag{C.12}$$

$$g^{(\nu)}(x) = \left[g^{(1)}(x)\right] = e^{i\nu\theta}, \nu \in \mathbb{Z} \quad \text{(family of mappings)}. \tag{C.13}$$

Again, ν is the winding number. Now imagine a two-dimensional plane with S^1 depicted as a circle and additionally an arbitrarily formed closed curve γ_1 around this circle. We can continuously deform γ_1 so that we are only left with the turns the curve made around S^1. We denote the new curve by γ_1'. Note, that since γ_1 and γ_1' are connected by a continuous transformation, they are homotopically equivalent and hence, the winding number of both curves is the same. By convention, we count the clockwise (counterclockwise) turns around S^1 positively (negatively). The total amount of turns around S^1 is the winding number ν. In the special case that the closed curve does not contain S^1, γ is *contractable*, i.e. continuously deformable to a single point, and $\nu = 0$. An example for $\nu = 1$ is given in fig.C.1 and an example for $\nu = -2$ is given in fig.C.2.

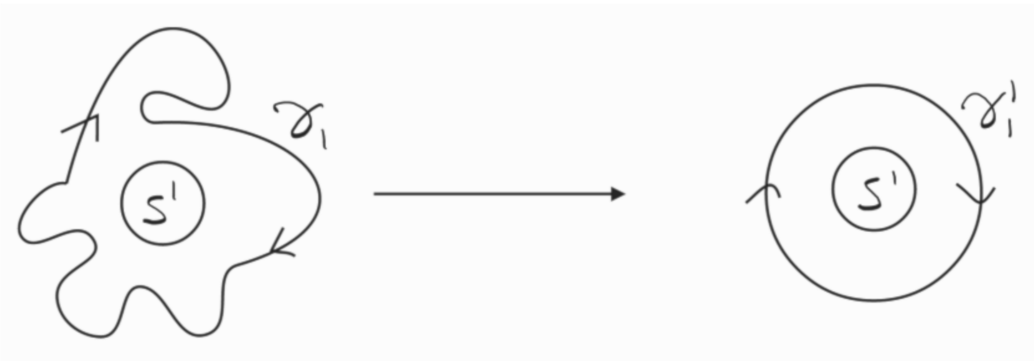


Figure C.1: *On the left hand side you can see one realization γ_1 of an arbitrary closed curve in two dimensions around the center S^1. After a continuous deformation γ_1 is transformed into γ_1', visualized on the right hand side. The new curve contains no deformations, so that you can read off $\nu = 1$ as the winding number. The arrows on the curves indicate the clockwise orientation of the curves around their center.*

Now that we got an intuition for the winding number, we can come back to SU(2). One can show that every mapping $S^3 \rightarrow S^3$ is homotopic to one of our standard mappings, which we do not want to prove here. It is basically the mathematical statement of what we visualized before in U(1), that we can easily read off the winding number after continuous deformation of an arbitrary closed curve γ, which is then immediately associated to an element of the family of mappings $g^{(\nu)}$. Define

$$\nu := \frac{1}{48\pi^2} \int d\theta_1 d\theta_2 d\theta_3 \varepsilon^{ijk} (g\partial_i g^{-1}, g\partial_j g^{-1} g\partial_k g^{-1}) \,, \tag{C.14}$$

where $\theta_{1,2,3}$ are the three angles that parametrize S^3 since the radius $r = 1$ is fixed by definition. The Jacobian of ε cancels the Jacobian of the angles,

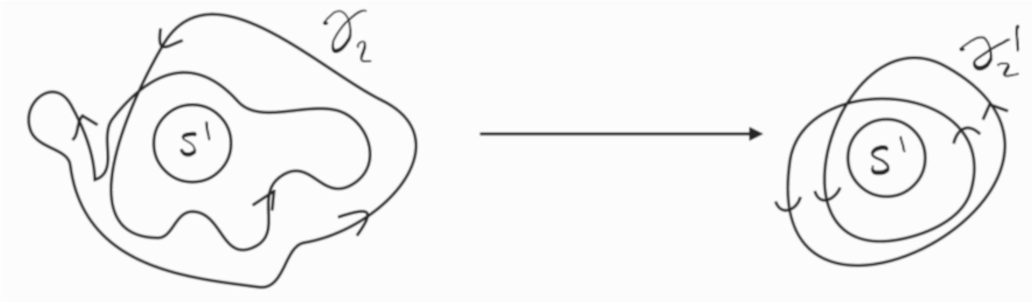


Figure C.2: *On the left hand side you can see one realization γ_2 of an arbitrary closed curve in two dimensions around the center S^1. After a continuous deformation γ_2 is transformed into γ_2', visualized on the right hand side. The new curve contains no deformations, so that you can read off $\nu = -2$ as the winding number. The arrows on the curves indicate the counterclockwise orientation of the curves around their center.*

so the explicit choice of the angles is irrelevant to the definition. We explicitly use the Cartan product to be representation-independent. We define the Cartan product below.

An algebra of a Lie group G is called the *Lie algebra*, which has a commutative connection, denoted by "+" and a multiplication, denoted by the Lie brackets "[.,.]". The latter satisfies the usual axioms for commutators and the Jacobi identity

$$[[x,y],z] + [[y,z],x] + [[z,x],y] = 0\,, \tag{C.15}$$

where x, y, z are elements of the Lie algebra, which are called *generators* if the algebra is irreducible. Generators are traceless. One commonly used *representation* T^a of the algebra elements is the matrix representation we are familiar with, for instance, by describing rotations via a rotation matrix.[41]. For SU(2) the generators are $T^a = -\frac{i}{2}\sigma^a$, which satisfy $[T^a, T^b] = \varepsilon^{abc}T^c$, where ε^{abc} is called the *structure constant*. In fact, all anti-hermitean matrices satisfy this condition and it is always possible to choose the generators such that $Tr[T^aT^b] \sim \delta^{ab}$, where the constant of proportionality is dependent on the chosen representation. Further, we define the *Cartan inner product* to be

$$(T^a, T^b) := \delta^{ab}\,, \tag{C.16}$$

so that $(T^a, T^b) \sim Tr[T^aT^b]$. Recall that for Pauli matrices $Tr[\sigma^a\sigma^b] = 2\delta^{ab}$

holds, so that for the specific example of SU(2) we obtain

$$Tr[T^a T^b] = \frac{i^2}{4} Tr[\sigma^a \sigma^b] = -\frac{1}{2}\delta^{ab} \tag{C.17}$$

$$\Rightarrow (T^a, T^b) = -2Tr[T^a T^b]. \tag{C.18}$$

In case of SU(3) we choose $T^a = -\frac{i}{2}\lambda^a$, where λ^a are the eight Gell-Mann matrices, for which $Tr[\lambda^a \lambda^b] = 2\delta^{ab}$, so that we get the same relation as above for SU(3). Note that we will later see that this is no coincidence.

Getting back to our integral form of the winding number (C.14), we can now with the previously obtained relation between the Cartan inner product and the trace for SU(2), we get

$$\nu := \frac{1}{24\pi^2} \int d\theta_1 d\theta_2 d\theta_3 Tr[\varepsilon^{ijk} g\partial_i g^{-1} g\partial_j g^{-1} g\partial_k g^{-1}]. \tag{C.19}$$

We could now use this property to show that ν is a homotopy invariant and that it is in fact the winding number we introduced before. We recommend the short and elegant proof in [40] and do not give it here. Let us now define

$$A_\mu := g A^a_\mu T_a\,, \tag{C.20}$$

where $g = $ const. is called the *gauge coupling constant* and $A^a_\mu(x)$ are the gauge potentials. We further define the field-strength tensor

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{C.21}$$

Note that $[A_\mu, A_\nu] = 0$ in electromagnetism. In a pure gauge field theory, the Euclidean action is simply given by

$$S = \frac{1}{4g^2} \int d^4x (F_{\mu\nu}, F_{\mu\nu}). \tag{C.22}$$

A *gauge transformation* is a function $g(x)$ from Euclidean space into the gauge group G, i.e. $g(x) = \exp\{\alpha_a(x) T^a\}$, where $\alpha_a(x)$ are arbitrary functions. Under such transformation, the potential A_μ and the field-strength tensor $F_{\mu\nu}$ transform as

$$A_\mu \to g A_\mu g^{-1} + g\partial_\mu g^{-1} \quad \text{and} \quad F_{\mu\nu} \to g F_{\mu\nu} g^{-1}\,, \tag{C.23}$$

respectively, which is why the action S is gauge-invariant. Let us quickly check that S converges properly. We assume that we can expand A_μ in powers of r^{-1}. We must ensure that $F_{\mu\nu}$ vanishes, which implies that A_μ is a gauge transformation of the null potential, so that $A_\mu = g\partial_\mu g^{-1} + \mathcal{O}(r^{-2})$, where g is a function of the four-space to the gauge group G of order one, i.e. a function of angular variables only. Hence, for large r, $F_{\mu\nu} \sim \partial_\mu A_\nu \sim \mathcal{O}(r^{-3})$ and $d^4x \sim \mathcal{O}(r^4)$, S converges properly for large r.

The *dual of an antisymmetric tensor*, denoted by a tilde, is conventionally defined as

$$\tilde{F}_{\mu\nu} := \frac{1}{2}\varepsilon_{\mu\nu}{}^{\lambda\sigma}F_{\mu\nu}\,, \tag{C.24}$$

where the factor $1/2$ is chosen, so that constructing the dual of a dual tensor gives the original tensor back, since the full contraction of two ε gives a factor four. With this definition at hand, we further define

$$G_\mu := 2\varepsilon_\mu{}^{\nu\lambda\sigma}(A_\nu, \partial_\lambda A_\sigma + \frac{2}{3}A_\lambda A_\sigma) = \varepsilon_\mu{}^{\nu\lambda\sigma}(A_\nu, F_{\lambda\sigma} - \frac{2}{3}A_{\lambda\sigma}) \tag{C.25}$$

$$\rightarrow \partial^\mu G_\mu = \frac{1}{2}\varepsilon^{\mu\nu\lambda\sigma}(F_{\mu\nu}, F_{\lambda\sigma}) = (F_{\mu\nu}, \tilde{F}^{\mu\nu}). \tag{C.26}$$

With these definitions we are finally able to compute (C.19). We start with

$$\int d^4x(F_{\mu\nu}, \tilde{F}_{\mu\nu}) = \int d^4x\partial^\mu G_\mu = \int d^3S\,\hat{r}^\mu G_\mu\,, \tag{C.27}$$

where in the last equation we used Gauss' Theorem, d^3S is the surface element of the hypersphere and $\hat{r}$ is the normal vector of the same surface. By considering the explicit expression for G_μ (C.25), we see that the first term is of $\mathcal{O}(r^{-4})$ and thus vanishes for large r, whilst the second term gives (C.14) up to a multiplicative constant, in which we bring the $48\pi^2$ to the other side. Now we replace the integral by $48\pi^2\nu$ and together with the additional factor $2/3$ in G_μ we obtain

$$\nu = \frac{1}{32\pi^2}\int d^4x(F, \tilde{F}), \tag{C.28}$$

which is precisely (C.1) what we wanted to show. We computed ν explicitly for the gauge group $SU(2)$, but due to a theorem of R. Bott[42], everything we computed for SU(2) above holds for an arbitrary simple Lie group, most importantly for us, it holds for SU(N).

C.2 Vacuum expectation value and θ-vacuum action

Consider a finite four-dimensional Euclidean spacetime volume V at a given time T. Like always, we can generalize our results in the end by sending V and T to infinity. We start with a finite volume because it eases our intuitive picture of what we will compute. For example, in a scalar field theory with spontaneous symmetry breaking, the expectation value of the field in the center of the box depends on the applied boundary conditions only, not on the actual boxsize. Fortunately, the only relic of the boundary conditions for a sufficiently large box is the winding number. Before we attempt to prove this assertion, we set our system up.

Say, the considered volume is a box with sides $L_0, ..., L_3$ with one corner placed at the origin of the coordinate system. On the walls of the box, the tangential components of A_μ are given in a way consistent with the convergence condition of the action, i.e. $A_\mu = g\partial_\mu g^{-1}$. For the sake of simplicity we drop terms of higher orders in r^{-1}. Since the tangential components of A_μ are given on the walls, g is given on the walls as well up to a multiplicative constant. The functional integral (C.2) we get in the end is gauge-invariant, thus we can choose a gauge solely for the sake of simplicity. Let us choose the gauge $A_3 = 0$, which still allows for arbitrary x_3-independent gauge transformation. Since $A_3 = 0$ it follows that $\partial_3 g = 0$, what means that g is automatically constant on the walls of the box, except for one of the 3-walls, on which it is given as a function $g_1(x_0, x_1, x_2)$, that is equal to one on the boundary of this wall. Now we imagine a second box that is set up in the same fashion, but with $g_2(x_0, x_1, x_2)$ and sides $L_0, L_1, L_2, L_3 + \Delta$, where $0 < \Delta \in \mathbb{R}^+$. Now we can prove the following theorem. "If g_1 and g_2 are in the same homotopy class, then any field configuration defined in the original box, consistent with the boundary conditions, can be extended to a field configuration in the larger box, consistent with its boundary conditions, and the same gauge at the cost of an increase in action of $\mathcal{O}\left(\frac{1}{\Delta}\right)$"[40].

Proof: By assumption, g_1 and g_2 are in the same homotopy class. Thus, it exists a continuous function of four variables, $g(x_0, x_1, x_2, s), 0 \leq s \leq 1$, so that $g(x_0, x_1, x_2, 0) = g_1$ and $g(x_0, x_1, x_2, 1) = g_2$. Let $g(x)$ be a function defined in the expanded volume by

$$g(x) = g\left(x_0, x_1, x_2, \frac{x_3 - L_3}{\Delta}\right). \tag{C.29}$$

The initially assumed vector potential $A_\mu = g\partial_\mu g^{-1}$ for $\mu \neq 3$ and zero otherwise would give the transition without any addition to the action, but is, unfortunately, inconsistent with the chosen gauge $A_3 = 0$. Hence, we first perform a gauge transformation with which $A_\mu = g\partial_\mu g^{-1}$ for $\mu = 3$ and zero otherwise. Now we see that $A_3 \sim \Delta^{-1}$. Recall the definition of $F_{\mu\nu}$ (C.21), so that the only non-vanishing components of $F_{\mu\nu}$ are $F_{\mu 3}$, i.e.

$$F_{\mu 3} = \partial_\mu A_3 - \partial_3 A_\mu + [A_\mu, A_3], \tag{C.30}$$

where the first term is proportional to Δ^{-1} by definition of $g(x)$, the second term vanishes for $\mu \neq 3$ and is also proportional to Δ^{-1} for $\mu = 3$ and the third term is always zero, since for $\mu = 3$ we get $[A_3, A_3] = 0$ and for $\mu \neq 3$ we get $[0, A_3] = 0$. Taking all this into account, we see that $F \sim \Delta^{-1}$, $F^2 \sim \Delta^{-2}$ and $d^4x = dx_0 dx_1 dx_2 dx_3 \sim \Delta$, so that we get that $S \sim \Delta^{-1}$, which is what we wanted to show.

Note, that if $\Delta \to \infty$, then $S \to 0$, so that indeed the only property of the boundary conditions that matter for large boxes is the winding number that gives the homotopy class of the boundary conditions that must be the same. If this would not be given, we need at least one additional instanton in the volume, which automatically gives an additional increase by at least $8\pi^2/g^2$, independent of Δ. We compute this value in the next section. Now that we realized that we only have to care about the winding number ν we can consider the functional integral and integrate over all field configurations, where $\nu < \infty$. To achieve this, we multiply an additional Kronecker delta (recall that $\nu \in \mathbb{Z}$) $\delta_{\nu n}$ in the integrand. The result of the integral for a fixed volume V, total time T and winding number n is

$$F(V, T, n) := N \int \mathcal{D}A e^{-S} \delta_{\nu n}\,, \tag{C.31}$$

where $\mathcal{D}A = \mathcal{D}A_0 \mathcal{D}A_1 \mathcal{D}A_2$ since A_3 is fixed by the gauge $A_3 = 0$. From (C.1) follows that for large times $T = T_1 + T_2$,

$$F(V, T_1 + T_2, n) = \sum_{n = n_1 + n_2} F(V, T_1, n_1) F(V, T_2, n_2). \tag{C.32}$$

To illustrate this, let us consider U(1) for a moment. The winding number is given analog to the SU(2) case as

$$\nu = \frac{i}{2\pi} \int_0^{2\pi} d\theta \frac{dg^{-1}}{d\theta}, \tag{C.33}$$

so that if $g(\theta) = g_1(\theta) g_2(\theta)$, then $\nu = \nu_1 + \nu_2$. This can be seen easily since ν is unchanged by continuous deformation, so we can deform g so that $g_1 = 1$ for $\theta \in [0, \pi)$ and likewise $g_2 = 1$ for $\theta \in [\pi, 2\pi)$. The integrand then is the sum of a part due to g_1 giving ν_1 and a part due to g_2 giving ν_2. The same argument holds for SU(2) just with semihyperspheres instead of semicircles and finally explains the behavior of F for large T. Since the original winding number n has to be constant for large T, $n = n_1 + n_2$ must hold. Thus, we observe a convolution of the two F which is unexpected. We want to compute the expectation value for $\exp(-HT)$ in one energy eigenstate. From basic quantum mechanics we know that this expectation value behaves like an exponential with the well known multiplication behavior,

$$\exp(a + b) = \exp(a) \cdot \exp(b). \tag{C.34}$$

We now reinstate this behavior by Fourier transformation of F which transforms a convolution in n-space into a multiplication in θ-space, i.e.

$$F(V, T, \theta) = \sum_n e^{in\theta} F(V, T, n) = \sum_n e^{in\theta} N \int \mathcal{D}A e^{-S} \delta_{\nu n} = N \int \mathcal{D}A e^{-S} e^{i\nu\theta}, \tag{C.35}$$

where in the second equality we change the order of integration and summation and then use the Kronecker delta to get rid of the summation over n, so that our convolution (C.32) becomes

$$F(V, T_1 + T_2, \theta) = F(V, T_1, \theta_1) \cdot F(V, T_2, \theta_2). \tag{C.36}$$

As we found the correct composition law now, we can identify F with the expectation value of $\exp(-HT)$ up to a normalization constant N in the energy eigenstate $|\theta\rangle$. We justify this notation later when computing the one-instanton action. With this identification we write

$$F(V, T, n) = N \langle \theta | e^{-HT} | \theta \rangle = N' \int \mathcal{D}A e^{-S} e^{i\nu\theta} \tag{C.37}$$

where N' is another normalization constant. However, by neglecting this constant, we find

$$\langle \theta | e^{-HT} | \theta \rangle \sim \int \mathcal{D}A e^{-S} e^{i\nu\theta}, \tag{C.38}$$

which is the property (C.2) we wanted to show. By carefully observing this property, we can further identify the exponent of the second exponential as an action in Minkowski space, which we call the θ-*vacuum* action due to its explicit θ-dependence, namely

$$S_\theta := \nu\theta = \frac{\theta}{32\pi^2}\int d^4x(F,\tilde{F})\,, \tag{C.39}$$

where we inserted the integral formula (C.1) in the second equality to obtain the third property (C.3).

C.3 One-instanton action S_0

Consider the integral (C.22) with the shorthand notation $(F_{\mu\nu}, F_{\mu\nu}) \equiv (F, F)$,

$$\int (F,F)d^4x = \left[\int (F,F)d^4x \int (F,F)d^4x\right]^{\frac{1}{2}} \tag{C.40}$$

$$= \left[\int (F,F)d^4x \int (\tilde{F},\tilde{F})d^4x\right]^{\frac{1}{2}} \geq \left|\int (F,\tilde{F})d^4x\right|\,, \tag{C.41}$$

where the first equality is trivial, the second equality holds by construction of the dual field strength tensor (C.24) and the inequality is the well-known Schwartz inequality, which is only a true equality if and only if $F = \pm\tilde{F}$. Now we plug in the explicit expressions (C.22) for the (F, F) and (C.1) for the $(F, \tilde{F})$ integral, respectively, to obtain

$$4g^2 S \geq 32\pi^2|\nu| \quad \Rightarrow S \geq \frac{8\pi^2}{g^2}|\nu| \tag{C.42}$$

where the equality is only given in the case $\nu = \pm 1$. For $\nu = 1$ we obtain the one-instanton action

$$S_0 = \frac{8\pi^2}{g^2} \tag{C.43}$$

which is (C.4), what we wanted to show. Note, that we exclude $\nu = 0$ as the trivial mapping and the choice of the positive ν sign is convention. Note additionally that the notation S_0 is by this argumentation slightly misleading, but we stick to it to be consistent with the literature.

C.4 Energy of a θ-vacuum

To compute the energy of a θ-vacuum, we will approach the problem in three steps. First, we consider a single potential well, then a double potential well and finally, a periodic potential. We proceed like this to first check that our technique works as it should, then to introduce the idea of an instanton and then to obtain the general energy formula for a θ-vacuum.

Let us start by considering a spinless particle of unit mass in a one-dimensional potential $V(x)$ with Hamiltonian $H = \frac{p^2}{2} + V(x)$. We tackle this system with Feynman's Path Integral formalism, which is easy generalizable to quantum field theory. Instead of working in Minkowski space, we rather would like to work in Euclidean space, which is simpler to treat. To go from Minkowski coordinates x^μ Euclidean coordinates x_E^μ we simple have to perform a Wick rotation on the time-coordinate x^0, i.e. $x_E^0 = -ix^0$, which is nothing else than an analytic continuation that now allows for complex times. By this transformation, we now work with the Euclidean metric $g_{\mu\nu} \equiv \delta_{\mu\nu}$. This transformation also applies to the action S in Minkowski space that becomes the action $S_E = -iS$ in Euclidean space. We drop the subscript E for convenience. For more details on analytic continuation, refer to [41]. In Euclidean space, we get the following path integral formula.

$$\langle x_f | e^{-HT} | x_i \rangle = N \int \mathcal{D}x e^{-S}, \tag{C.44}$$

where the left hand side is the propagator of the particle to propagate from an initial state $|x_i\rangle$ to a final state $|x_f\rangle$ with Hamiltonian H in time T and the right hand side is the path integral with normalization constant N. On the left hand side, we can insert a complete orthonormal set of energy eigenstates $|n\rangle, H|n\rangle = E_n|n\rangle, \sum_n |n\rangle\langle n| = \mathbb{1}$ to expand the propagator in terms of the systems energies E_n and wave-functions as

$$\langle x_f | e^{-HT} | x_i \rangle = \sum_n \langle x_f | e^{-HT} | n \rangle \langle n | x_i \rangle \tag{C.45}$$

$$= \sum e^{-E_n T} \langle x_f | n \rangle \langle n | x_i \rangle. \tag{C.46}$$

Thus, the leading term in this expansion for large T gives the energy and wave-function of the lowest-lying energy eigenstate. The right hand side can be easily evaluated in the semiclassical limit ($\hbar \to 0$). Note that we have set $\hbar = 1$, but we can keep track of the powers of $\hbar$ in our equations, e.g the exponential with $\hbar \neq 1$ reads as $\exp(-\frac{HT}{\hbar})$. Consider a potential well centered at the origin and choose $x_i = 0 = x_f$, so that the only solution that satisfies this condition is $\bar{x} = 0$, from which $S = 0$ follows immediately. This gives $E_0 = \frac{1}{2}\hbar\omega[1 + \mathcal{O}(\hbar)]$, where we explicitly kept the $\hbar$ to quickly compare the result with the expected result from basic quantum mechanics. Indeed, we find the expected ground state energy up to a correction of $\mathcal{O}(\hbar)$ due to the used path integral formalism that takes into account all paths with their corresponding actions as weighs instead of just the classical path one would choose in basic quantum mechanics. For a detailed calculation, see [40] for example.

Let us now take the next step and consider a double-well potential. Assume it is symmetric around the origin, i.e. $V(-x) = V(x)$, and that the minima are located at $x = \pm a$ with $V(\pm a) = 0$ and $V''(\pm a) = \omega^2$. We attempt to compute propagators from one minimum into itself but also propagators from one minimum to the other. Hence, the first step is to find solutions of the classical equations of motion consistent with the boundary conditions. We already mentioned the two obvious solutions. The particle either remains in the minimum or propagates to the other one. However, we take the big times limit $T \to \infty$. In the case $x(t = -\infty)$ =one minimum and $x(t = +\infty)$ =the other minimum we deal with a solution of the equations of motion

$$0 = E = \frac{1}{2}\left(\frac{dx}{dt}\right)^2 - V(x) \Rightarrow \frac{dx}{dt} = (2V(x))^{\frac{1}{2}} \Rightarrow t = t_1 + \int_0^x dx'(2V(x))^{-\frac{1}{2}}\,, \tag{C.47}$$

where t_1 is an integration constant, i.e. the time at which x vanishes. We call this solution an *instanton* with center at t_1[40]. Let us quickly give this name some content. In fact, all the time, we talk about certain field configurations. The especially interesting potential minima are pure gauge configurations we search for as introduced earlier. The propagation from one such gauge configuration to another is then nothing else than the propagation of the field configurations between different vacua. Since this is a process in time rather than a space translation, we give it the name prefix *instant-* to stress this behavior[13]. By replacing $t \to -t$ we find the solutions with $a \to -a$ that are called *anti-instantons*. First, let us compute the action

$$S_0 = \int_{-\frac{T}{2}}^{+\frac{T}{2}} dt \left[\frac{1}{2}\left(\frac{dx}{dt}\right)^2 + V\right] = \int_{-\frac{T}{2}}^{+\frac{T}{2}} dt \left(\frac{dx}{dt}\right)^2 = \int_{-a}^{+a} dx(2V(x))^{\frac{1}{2}}\,, \tag{C.48}$$

where we first replace $V(x)$ with (C.47) and then use the same equation to replace one $\frac{dx}{dt}$ and the other is used to replace dt by dx in the integral. We now wish to evaluate the functional integral by summing over all configurations with n instantons (or anti-instantons) centered at $t_1, ..., t_n$, where $\frac{T}{2} > t_1 > ... > t_n > -\frac{T}{2}$ with T large compared to the instanton scale.

First, for n widely separated instantons (or anti-instantons), we get $S = n \cdot S_0$. Second, consider the time-evolution operator $\exp(-HT)$ as a product of operators associated with the evolution between the points in time shortly before and after the position of an (anti-)instanton. This breaks the problem in n single-well potential problems. For one of them we obtain $\left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} \exp(-\omega T/2) K^n$, where K is a constant that ensures that we get the correct result for a single (anti-)instanton. Third, we integrate over the locations of the particle centers described above as

$$\int_{-\frac{T}{2}}^{+\frac{T}{2}} dt_1 \int_{-\frac{T}{2}}^{t_1} dt_2 ... \int_{-\frac{T}{2}}^{t_{n-1}} dt_n = \frac{T^n}{n!}. \tag{C.49}$$

Fourth, we are not free to distribute instantons and anti-instantons arbitrarily. E.g. if we start at $-a$, the first object we encounter must be an instanton, the next an anti-instanton and so on and so forth. Furthermore, if we want to end in this case at $-a$, n must be even and likewise if we want to end at $+a$ in this case, n must be odd. Now we take all our findings into account to get

$$\langle -a | e^{-HT} | -a \rangle = \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \sum_{n \text{ even}} \frac{\left(K e^{-S_0} T\right)^n}{n!} (1 + \mathcal{O}(\hbar)). \tag{C.50}$$

Equivalently, we get $\langle +a | \exp(-HT) | +a \rangle$ as an expansion summed over odd n. Whilst the summation over even n gives a cosine term, the summation over odd n gives a sinus term, so that for the mixed propagator we obtain

$$\langle \pm a | e^{-HT} | -a \rangle = \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \frac{1}{2} \left(e^{K e^{-S_0} T} \mp e^{K e^{-S_0} T} \right) (1 + \mathcal{O}(\hbar)). \tag{C.51}$$

Comparison with (C.46) reveals that we have two low-lying energy eigenstates with the energies

$$E_{\pm} = \frac{1}{2}\omega \pm K e^{-S_0}. \tag{C.52}$$

For our purposes, the explicit form of K is irrelevant in this case. Note that we have set $\hbar = 1$ again since we already demonstrated earlier that our formalism gives the correct results.

Let us finally consider a periodic potential with minima $a \in \mathbb{Z}$ for simplicity. The n instantons and $\tilde{n}$ anti-instantons are distributed arbitrarily about the real axis with initial position $x = m$ and instantons (anti-instantons) go to $x = m + 1$ ($x = m - 1$). Each (anti-)instanton must begin where its predecessor ended. Thus, $\Delta n := n - \tilde{n}$ must equal the change in x between the initial and final position eigenstates, $\Delta j := j_+ - j_-$, which is ensured by an additional Kronecker delta (recall that we explicitly constructed the positions as integers) in the double sum in the propagator. By the same considerations as for the double-well potential we find

$$\langle j_+|e^{-HT}|j_-\rangle = \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \sum_n \sum_{\tilde{n}} \frac{\left(Ke^{-S_0}T\right)^{n+\tilde{n}}}{n!\tilde{n}!} \delta_{\Delta n, \Delta j} \,. \tag{C.53}$$

Now we use the identity

$$\delta_{ab} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{i\theta(a-b)} \tag{C.54}$$

with $a = \Delta n = n - \tilde{n}$ and $b = \Delta j$ to obtain

$$\begin{aligned}
\langle j_+|e^{-HT}|j_-\rangle &= \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \sum_n \sum_{\tilde{n}} \frac{\left(Ke^{-S_0}T\right)^{n+\tilde{n}}}{n!\tilde{n}!} \int \frac{d\theta}{2\pi} e^{i\theta(n-\tilde{n}-\Delta j)} \\
&= \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \int \frac{d\theta}{2\pi} e^{i\theta\Delta j} \sum_n \frac{\left(Ke^{-S_0}e^{i\theta}T\right)^n}{n!} \sum_{\tilde{n}} \frac{\left(Ke^{-S_0}e^{-i\theta}T\right)^{\tilde{n}}}{\tilde{n}!} \\
&= \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \int \frac{d\theta}{2\pi} e^{i\theta\Delta j} \exp\left(Ke^{-S_0}e^{i\theta}T\right) \exp\left(Ke^{-S_0}e^{-i\theta}T\right) \\
&= \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \int \frac{d\theta}{2\pi} e^{i\theta\Delta j} \exp\left(Ke^{-S_0}T\left(e^{i\theta} + e^{-i\theta}\right)\right) \\
&= \left(\frac{\omega}{\pi}\right)^{\frac{1}{2}} e^{-\frac{\omega T}{2}} \int \frac{d\theta}{2\pi} e^{i\theta\Delta j} \exp\left(Ke^{-S_0}T \cdot 2\cos(\theta)\right) .
\end{aligned} \tag{C.55}$$

Again, by comparing with (C.46) we can read off the energy eigenvalues to be

$$E(\theta) = \frac{1}{2}\omega + 2K\cos(\theta)e^{-S_0} \tag{C.56}$$

what gives

$$E(\theta) \sim \cos(\theta)e^{-S_0} \,, \tag{C.57}$$

which is exactly the fifth property (C.5) we wanted to show. This justifies retroactively the notation of the energy eigenstate $|\theta\rangle$.

Appendix D

The thermal axion abundance

We now want to compute the thermal abundance of axions relative to the thermal photon abundance. We will loosely follow the order of argumentation presented in the book of Kolb and Turner[28]. However, we will start by deriving the so-called Boltzmann-equation and then modify it properly as a first approximation since we assume the axions to be in thermal equilibrium with the thermal bath. Later we will see that the era of inflation in fact gives a departure from this equilibrium, but we will see at the end that this does not matter effectively.

Let $f(x^\mu, p^\mu)$ be the particle's phase space distribution, then its evolution is described by the *Boltzmann-equation*

$$\hat{L}[f] = \hat{C}[f], \tag{D.1}$$

where

$$\hat{L} = p^\mu \partial_\mu - \Gamma^\mu_{\nu\rho} p^\nu p^\rho \partial_\mu \tag{D.2}$$

is the *Liouville-operator* and $\hat{C}$ is the *collision-operator*, which we will specify below. Due to the Christoffel symbols, $\Gamma^\mu_{\nu\rho}$, in $\hat{L}$ we take gravitational effects into account. We use the RW-metric

$$ds^2 = -dt^2 + a^2[dr^2 + r^2 d\Omega^2] \tag{D.3}$$

for a flat Universe. RW is short for *Robertson-Walker*, two of the many people who contributed significantly to arriving at this result. For the sake of shortness, we will always refer to it as the RW-metric, even though in the literature, it is typical that one finds FLRW for *Friedmann-Lemaitre-Robertson-Walker* or even more scientists mentioned. However, note that most importantly for us is that the RW-metric is the most general metric that obeys the cosmological principle. We explicitly state that we use it for a flat Universe because the RW-metric in general reads

$$ds^2 = -dt^2 + a^2[dr^2 + S_\kappa^2(r)d\Omega^2], \tag{D.4}$$

where $\kappa \in \{-1, 0, +1\}$ describes a negatively curved, flat or positively curved Universe, respectively and

$$S_\kappa(r) = \begin{cases} R\sin(r/R) & \kappa = +1 \\ r & \kappa = 0 \\ R\sinh(r/R) & \kappa = -1 \end{cases}, \tag{D.5}$$

where R is the Universe's radius of curvature. Back to the $\kappa = 0$ case, in these coordinates, f is spatially homogeneous and isotropic and thus can be written as

$$f(t, |\vec{p}|) \Leftrightarrow f(t, E). \tag{D.6}$$

Additionally, the non-vanishing Christoffel-symbol can be easily computed via the Lagrangian

$$L = \left(-g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right)^{\frac{1}{2}} \tag{D.7}$$

and the variational principle by comparing with the geodesic equation

$$\frac{d^2x^\mu}{d\lambda^2} = \Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}. \tag{D.8}$$

Since we only have to consider $\alpha = 0$ we are left with

$$\Gamma^0_{ij} = \frac{\dot{a}}{a}g_{ij} \tag{D.9}$$

as the only non-vanishing Christoffel-symbol of interest with which we get

$$\hat{L} = E\frac{\partial}{\partial t} - \frac{\dot{a}}{a}g_{ij}p^i p^j\frac{\partial}{\partial E} \Rightarrow \hat{L}[f] = E\frac{\partial f}{\partial t} - \frac{\dot{a}}{a}g_{ij}p^i p^j\frac{\partial f}{\partial E}. \tag{D.10}$$

The number density is generally given as

$$n(t) = \frac{g}{(2\pi)^3}\int d^3p f(E, t), \tag{D.11}$$

what we can derive with respect to t to get

$$\frac{dn(t)}{dt} = \frac{g}{(2\pi)^3} \int d^3p \frac{df(E,t)}{dt} \overset{E=\text{const.}}{=} \frac{g}{(2\pi)^3} \int d^3p \frac{\partial f(E,t)}{\partial t}. \tag{D.12}$$

Now we divide (D.10) by E to get

$$\frac{\hat{C}[f]}{E} = \frac{\hat{L}[f]}{E} = \frac{\partial f}{\partial t} - \frac{\dot{a}}{a}\frac{|\vec{p}|^2}{E}\frac{\partial f}{\partial E} \Rightarrow \frac{\partial f}{\partial t} = \frac{\hat{C}[f]}{E} + \frac{\dot{a}}{a}\frac{|\vec{p}|^2}{E}\frac{\partial f}{\partial E}, \tag{D.13}$$

what we plug into (D.12) and obtain

$$\frac{dn(t)}{dt} = \frac{g}{(2\pi)^3} \int \hat{C}[f]\frac{d^3p}{E} + \frac{g}{(2\pi)^3}\frac{\dot{a}}{a}\frac{1}{E} \int |\vec{p}|^2 \frac{\partial f}{\partial E} d^3p. \tag{D.14}$$

For the second term on the right-hand side we demand

$$\frac{g}{(2\pi)^3}\frac{\dot{a}}{a}\frac{1}{E} \int |\vec{p}|^2 \frac{\partial f}{\partial E} d^3p \overset{!}{=} -3\frac{\dot{a}}{a}\frac{g}{(2\pi)^3} \int d^3p f \tag{D.15}$$

$$\rightarrow \int \frac{|\vec{p}|^2}{E}\frac{\partial f}{\partial E} d^3p = -3 \int f d^3p. \tag{D.16}$$

For convenience, we write $|\vec{p}|^2 \equiv p^2$ and note that $d^3p = p^2 dp d\Omega$ and since $E^2 = p^2 + m^2$ holds, we get

$$EdE = pdp \Rightarrow \frac{\partial}{\partial E} = \frac{dp}{dE}\frac{\partial}{\partial p} = \frac{E}{p}\frac{\partial}{\partial p} \tag{D.17}$$

to further rewrite the previous integral as

$$\int \frac{p^2}{E}\frac{\partial f}{\partial E} d^3p = \int \frac{p^2}{E}\frac{E}{p}\frac{\partial f}{\partial p} p^2 dp d\Omega = \int \frac{\partial f}{\partial p} p^3 dp d\Omega \tag{D.18}$$

$$= -\int f \frac{\partial p^3}{\partial p} dp d\Omega = -3 \int f p^2 dp d\Omega = -3 \int f d^3p \equiv n(t), \tag{D.19}$$

where in third equality we integrated by parts with respect to p and in the fourth equality we recognized d^3p we expanded above. With this we can replace the second term in (D.14) and rearrange the terms to get

$$\frac{dn}{dt} + 3\frac{\dot{a}}{a}n = \frac{g}{(2\pi)^3} \int \hat{C}[f]\frac{d^3}{E}. \tag{D.20}$$

Let us now consider the general collision process

$$\psi + a + b + \ldots \leftrightarrow i + j + \ldots, \tag{D.21}$$

so that the collision term is given by

$$\frac{g}{(2\pi)^3}\int \hat{C}[f]\frac{d^3p_\psi}{E_\psi} = -\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdots$$
$$\cdot (2\pi)^4 \delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots)$$
$$\cdot [|\mathcal{M}|^2_{\psi+a+b+\ldots\to i+j+\ldots} f_a f_b \cdots f_\psi (1 \pm f_i)(1 \pm f_j)\cdots$$
$$- |\mathcal{M}|^2_{\psi+a+b+\ldots\leftarrow i+j+\ldots} f_i f_j \cdots (1 \pm f_a)(1 \pm f_b)\cdots(1 \pm f_\psi)], \tag{D.22}$$

where $f_{a,b,\ldots,i,j,\ldots}$ are the phase space densities of the species $a, b, \ldots, i, j, \ldots$, f_ψ is the phase space density of ψ, on whose evolution we are focusing on, $+$ and $-$ signs in the parentheses correspond to bosons and fermions, respectively,

$$d\Pi_i := \frac{g}{(2\pi)^3}\frac{d^3p_i}{2E_i}, \tag{D.23}$$

g are the internal degrees of freedom, the $\delta^{(4)}$-term enforces energy and momentum conservation and $|\mathcal{M}|^2$ are the matrix elements for the mentioned processes, averaged over the initial and final spins and includes the appropriate symmetry factors for identical particles in the initial and final state coming from Quantum Field Theory.

We now make the following assumptions. First, T-symmetry holds, i.e.

$$|\mathcal{M}|^2_{\psi+a+b+\ldots\to i+j+\ldots} = |\mathcal{M}|^2_{\psi+a+b+\ldots\leftarrow i+j+\ldots} =: |\mathcal{M}|^2, \tag{D.24}$$

and second, all species can be described by Maxwell-Boltzmann statistics instead of Fermi-Dirac and Bose-Einstein statistics for fermions and bosons, respectively. Note, the former one gets exact for $(m_i - \mu_i)/T \ll 1$ besides the fact that they are all three quite similiar anyways and much smaller than one for $p \sim p_{\text{peak}}$. However, in the absence of Bose condensation or Fermi degeneracy the blocking/stimulated emission factors, i.e. the $1 \pm f$ terms in parentheses, become

$$1 \pm f \approx 1 \text{ and } f_i(E_i) = \exp\left\{-\frac{E_i - \mu_i}{T}\right\} \tag{D.25}$$

for all species in kinetic equilibrium. The collision term then reads

$$\frac{g}{(2\pi)^3}\int \hat{C}[f]\frac{d^3p_\psi}{E_\psi} = -\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdots$$
$$\cdot (2\pi)^4 \delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots)$$
$$\cdot |\mathcal{M}|^2 [f_a f_b \cdots f_\psi \cdots - f_i f_j], \tag{D.26}$$

so that the Boltzmann equation (D.20) becomes

$$\begin{aligned}\dot{n}_\psi + 3Hn_\psi = -\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdots \\ \cdot (2\pi)^4 \delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots) \\ \cdot |\mathcal{M}|^2 [f_a f_b \cdots f_\psi \cdots - f_i f_j],\end{aligned} \tag{D.27}$$

where we replaced $H \equiv \dot{a}/a$ as usual. Note, the term $3Hn_\psi$ describes the dilution effect of the expansion of the Universe and the right-hand side describes the interactions that change the number of ψ-particles present. Thus, if the right-hand side vanishes, i.e. there are no interactions, then simply

$$\dot{n}_\psi + 3Hn_\psi = 0 \Rightarrow \dot{n}_\psi = -3Hn_\psi \Rightarrow \frac{dn_\psi}{n_\psi} = -3\frac{da}{dt}\frac{1}{a}dt = -3\frac{da}{a} \tag{D.28}$$

$$\Rightarrow \ln(n_\psi) = -3\ln(a) + \text{const.} = \ln\left(\frac{1}{a^3}\right) + \text{const.} \tag{D.29}$$

$$\Rightarrow n_\psi = \frac{1}{a^3} + \text{const.} \sim a^{-3}, \tag{D.30}$$

what is the expected behaviour for an expanding Universe since the total number $N_\psi = \text{const.}$ and the volume $V \sim a^3$, so that $n_\psi \sim V{-}1 \sim a^{-3}$.

Let us scale out the expansion effects by using the entropy density, s, to define

$$Y := \frac{n_\psi}{s}, \tag{D.31}$$

where sa^3 =const. is conserved in a comoving volume, so that

$$0 = \frac{d}{dt}(sa^3) = \dot{s}a^3 + 3sa^2\dot{a} = \dot{s}a^3 + 3sa^3H \Rightarrow \dot{s} = -3sH, \tag{D.32}$$

what we can use to rewrite the left-hand side of the Boltzmann equation (D.27) in the form

$$\begin{aligned}\dot{n}_\psi + 3Hn_\psi = \frac{d}{dt}(Ys) + 3HsY = \dot{Y}s + Y\dot{s} + 3HsY \\ = \dot{Y}s + Y(-3sH) + 3HsY = s\dot{Y}.\end{aligned} \tag{D.33}$$

The collision term in (D.27) depends on the temperature, not the time, so we introduce a new variable, $x := m/T$, that relates the temperature to the mass scale of the considered collision particle. Since inflation should have taken place during the radiation-dominated epoch, x and t are related. To find this relation, we consider the Friedmann equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} = \frac{8\pi G\rho_0}{3}a^{-3(1+w)} \overset{w=\frac{1}{3}}{=} \frac{8\pi G\rho_{\gamma,0}}{3}a^{-4}, \tag{D.34}$$

where the subscript γ denotes radiation and $w = 1/3$ is the density of state-parameter for the radiation-dominated epoch what we discussed in appendix A. Now, we use the density parameter

$$\Omega_{\gamma,0} := \frac{\rho_{\gamma,0}}{\rho_{\text{crit}}} \quad , \rho_{\text{crit}} = \frac{3H_0^2}{8\pi G} \tag{D.35}$$

to get

$$\left(\frac{\dot{a}}{a}\right)^2 = \Omega_{\gamma,0}H_0^2 a^{-4} \Rightarrow \dot{a}^2 = \Omega_{\gamma,0}H_0^2 a^{-2} \Rightarrow \dot{a} = \Omega_{\gamma,0}^{\frac{1}{2}} H_0 a^{-1}$$
$$\Rightarrow a da = \Omega_{\gamma,0}^{\frac{1}{2}} H_0 dt \Rightarrow \frac{1}{2}a^2 = \Omega_{\gamma,0}^{\frac{1}{2}} H_0 t, \tag{D.36}$$

where we used $a(t = 0) = 0$ to set the integration constant to zero. By setting $a = 1$ we find

$$\frac{1}{2} = \Omega_{\gamma,0}^{\frac{1}{2}} H_0 t_0 \Rightarrow t_0 = \left(2\Omega_{\gamma,0}^{\frac{1}{2}} H_0\right)^{-1} \tag{D.37}$$

as the current time. During radiation-domination, however, we can use this equation for all times and simply drop the subscript 0 to get

$$t = \left(2\Omega_{\gamma}^{\frac{1}{2}} H\right)^{-1}. \tag{D.38}$$

Recall (D.35) to get

$$t^2 = \frac{1}{4\Omega_\gamma H^2} = \frac{1}{4H^2}\frac{3H^2}{8\pi G\rho_\gamma} = \frac{3}{32\pi G}\frac{1}{\rho_\gamma} \Rightarrow \rho_\gamma(t) = \frac{3}{32\pi G}\frac{1}{t^2} \tag{D.39}$$

as the t-dependent density of the Universe during radiation-domination. Let us now find another expression for $\rho(t)$. We will reinstate $c, \hbar$ and k_B solely because it will help us in the end to identify m_{pl}. We can write the energy density as

$$c^2\rho(t) = \frac{g}{(2\pi\hbar)^3}\int f(p,t)E(p)d^3p = \frac{g}{2\pi^2\hbar^3}\int f(p,t)(p^2c^2 + m^2c^4)^{\frac{1}{2}}p^2 dp, \tag{D.40}$$

where we used $E^2(p) = p^2c^2 + m^2c^4$ as usual. As discussed earlier, we consider the particle species in thermal equilibrium, so that we can use

$$f(p,t) = \left[\exp\left\{\frac{E(p)-\mu}{k_BT}\right\} \pm 1\right]^{-1} \tag{D.41}$$

with + or − for fermions or bosons, respectively. We neglect the chemical potential, μ, to get

$$c^2\rho(t) = \frac{g}{2\pi^2\hbar^3}\int_0^\infty dp \frac{p^2(p^2c^2+m^2c^4)^{\frac{1}{2}}}{\left[\exp\left\{\frac{E(p)-\mu}{k_BT}\right\} \pm 1\right]}. \tag{D.42}$$

During radiation-domination we can safely assume the particle species to be relativistic, so we take the relativistic limit $pc \gg mc^2$ above to simplify

$$c^2\rho(t) \approx \frac{g}{2\pi^2\hbar^3}\int_0^\infty dp \frac{p^3c}{\exp\left[\frac{pc}{k_BT}\right] \pm 1} \tag{D.43}$$

and define

$$x := \frac{pc}{k_BT} \Leftrightarrow p = \frac{k_BTx}{c} \Rightarrow dp = \frac{k_BT}{c}dx \tag{D.44}$$

to continue with

$$\begin{aligned} c^2\rho(t) &\approx \frac{g}{2\pi^2\hbar^3}\int_0^\infty \frac{k_BTdx}{c}\frac{\left(\frac{k_BT}{c}\right)^3 x^3c}{\exp(x)\pm 1} \\ &= \frac{gc}{2\pi^2\hbar^3}\left(\frac{k_BT}{c}\right)^4\int_0^\infty dx\frac{x^3}{\exp(x)\pm 1} \\ &= \frac{gc}{2\pi^2\hbar^3}\left(\frac{k_BT}{c}\right)^4 \cdot \begin{cases} \frac{\pi^4}{15} & \text{bosons} \\ \frac{7}{8}\frac{\pi^4}{15} & \text{fermions} \end{cases}. \end{aligned} \tag{D.45}$$

Since the total energy density is a sum of its constituents, we can split the solution explicitly in a part for bosons and one for fermions as

$$\begin{aligned} c^2\rho_{\text{tot}} &= \frac{\pi^2}{30(\hbar c)^3}\left[\sum_{i\in\text{bosons}} g_i(k_BT_i)^4 + \frac{7}{8}\sum_{i\in\text{fermions}} g_i(k_BT_i)^4\right] \\ &= \frac{\pi^2}{30(\hbar c)^3}(k_BT)^4\left[\sum_{i\in\text{bosons}} g_i\left(\frac{T_i}{T}\right)^4 + \frac{7}{8}\sum_{i\in\text{fermions}} g_i\left(\frac{T_i}{T}\right)^4\right] \\ &=: \frac{\pi^2}{30(\hbar c)^3}g_*(T), \end{aligned} \tag{D.46}$$

where in the last line we defined $g_*(T) \equiv g_*$ to be the *effective number of degrees of freedom*. We can use this expression for the density to replace it in the time equation (D.39), what becomes

$$\frac{1}{c^2}\frac{\pi^2}{30(\hbar c)^3}g_*(k_BT)^4 = \frac{3}{32\pi G}\frac{1}{t^2}. \tag{D.47}$$

Solving for t gives

$$t^2 = \frac{90\hbar^3c^5}{32\pi^3 Gk_B^4}\frac{1}{g_*}\frac{1}{T^4} = \frac{45\hbar^2c^4}{16\pi^3k_B^4}g_*^{-1}\frac{m_{\rm pl}^2}{T^4} = \frac{45}{16\pi^3}g_*\frac{m_{\rm pl}}{T^2}, \tag{D.48}$$

where in the second equality we recognized $m_{\rm pl}^2 = \hbar c/G$ and in the third equality we reinstated $c = \hbar = k_B = 1$. Finally, after taking the square root and using our transformation $x = m/T$ we get

$$t = \left(\frac{45}{16\pi^3}\right)^{\frac{1}{2}} g_*^{-\frac{1}{2}}\frac{m_{\rm pl}}{T^2} \approx 0.301 g_*^{-\frac{1}{2}}\frac{m_{\rm pl}}{T^2} = 0.301 g_*^{-\frac{1}{2}}\frac{m_{\rm pl}}{m^2}x^2. \tag{D.49}$$

Recall $\dot{n}_\psi + 3Hn_\psi = s\dot{Y}$ and use

$$\frac{dt}{dx} = 0.602 g_*^{-\frac{1}{2}}\frac{m_{\rm pl}}{m^2}x \tag{D.50}$$

to get

$$\begin{aligned}
&\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdot (2\pi)^4\delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots)\\
&\cdot |\mathcal{M}|^2[f_af_b\cdots f_\psi\cdots - f_if_j] = \dot{n}_\psi + 3Hn_\psi = s\frac{dY}{dt} = s\frac{dY}{dx}\frac{g_*^{\frac{1}{2}}}{0.602}\frac{m^2}{m_{\rm pl}}\frac{1}{x}\\
&\Rightarrow \frac{dY}{dx} = -0.602 g_*^{-\frac{1}{2}}\frac{m_{\rm pl}}{m^2}\frac{1}{sx}\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdot \cdots (2\pi)^4\\
&\cdot \delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots)|\mathcal{M}|^2[f_af_b\cdots f_\psi\cdots - f_if_j]
\end{aligned} \tag{D.51}$$

and thus at the final form of the Boltzmann-equation

$$\begin{aligned}
&\frac{dY}{dx} =: -\frac{x}{H(m)s}\int d\Pi_\psi d\Pi_a d\Pi_b \cdots d\Pi_i d\Pi_j \cdot \cdots (2\pi)^4 \cdot\\
&\delta^{(4)}(p_\psi + p_a + p_b + \ldots - p_i - p_j - \ldots)|\mathcal{M}|^2[f_af_b\cdots f_\psi\cdots - f_if_j],
\end{aligned} \tag{D.52}$$

where we defined $H(m) := 1.67 g_*^{1/2} m^2/m_{\text{pl}}$ and further $H(m) := x^2 H(x)$. With the Boltzmann equation at hand we can finally compute the freeze-out of certain particle species. First, we consider a $2 \leftrightarrow 2$process generally and later plug in the specific, for us very interesting, case of photons. Afterwards we redo the calculation for axions to give the number density of axions relative to those of photons since the latter one is well-measured by CMB observation.

Now, a general $2 \leftrightarrow 2$ processs is

$$\psi\bar{\psi} \leftrightarrow X\bar{X}. \tag{D.53}$$

Let us assume that the species is stable (or long-lived compared to the age of the Universe), so that this is in fact its only mentionable production mechanism. Let us further assume that all the species $X, \bar{X}$ follow a thermal distribution with vanishing chemical potential and that $X\bar{X}$ interact stronger than $X\psi$ for example, so that the $X\bar{X}$ can be thought to be in equilibrium. A typical exmaple for such a process is

$$\nu + \bar{\nu} \leftrightarrow e^- + e^+, \tag{D.54}$$

where the electron-positron pair interacts electromagnetically whilst the neutrinos interact weakly. However, we can write down the Boltzmann-equation (D.52)

$$\begin{aligned}\frac{dY}{dx} = -0.602 g_*^{-\frac{1}{2}} \frac{m_{\text{pl}}}{m^2} \frac{1}{sx} \int d\Pi_\psi d\Pi_{\bar{\psi}} d\Pi_X d\Pi_{\bar{X}} (2\pi)^4 |\mathcal{M}|^2 \\ \cdot \delta^{(4)}(p_\psi + p_{\bar{\psi}} - p_X - p_{\bar{X}})[f_\psi f_{\bar{\psi}} - f_X f_{\bar{X}}].\end{aligned} \tag{D.55}$$

Note, that the $\delta^{(4)}$-term implies $E_\psi + E_{\bar{\psi}} = E_X + E_{\bar{X}}$ and further note that we assumed

$$f_i = \exp\left\{-\frac{E_i - \mu_i}{T_i}\right\} \tag{D.56}$$

with $\mu_i = 0$ and $T_i \equiv T$ for all species i in thermal equilibrium, so that

$$f_\psi^{\text{EQ}} f_{\bar{\psi}}^{\text{EQ}} = \exp\left\{-\frac{E_\psi + E_{\bar{\psi}}}{T}\right\} = \exp\left\{-\frac{E_X + E_{\bar{X}}}{T}\right\} = f_X f_{\bar{X}}, \tag{D.57}$$

so that

$$f_\psi f_{\bar{\psi}} - f_X f_{\bar{X}} = f_\psi f_{\bar{\psi}} - f_\psi^{\text{EQ}} f_{\bar{\psi}}^{\text{EQ}}. \tag{D.58}$$

The above equation only holds for ψ's and $\bar{\psi}$'s in thermal equilibrium because otherwise $T_i \not\equiv T$ for all i. With this results we can write

$$\int d\Pi_\psi d\Pi_{\bar{\psi}} d\Pi_X d\Pi_{\bar{X}} (2\pi)^4 |\mathcal{M}|^2 \cdot \delta^{(4)}(p_\psi + p_{\bar{\psi}} - p_X - p_{\bar{X}})$$
$$\cdot \exp\left\{ -\frac{E_\psi^{\mathrm{EQ}} + E_{\bar{\psi}}^{\mathrm{EQ}}}{T} \right\} =: \langle \sigma_{\psi\bar{\psi}\to X\bar{X}} |v| \rangle \left(n_\psi^{\mathrm{EQ}}\right)^2 \tag{D.59}$$

and likewise for the ψ's and $\bar{\psi}$'s out of equilibrium by simply dropping the superscript EQ in the above equation. We can combine both equations to replace the integral in the Boltzmann equation (D.55) to get

$$\frac{dY}{dx} = -\frac{x}{H(m)s}\left[\langle \sigma_{\psi\bar{\psi}\to X\bar{X}} |v| \rangle (sY)^2 - \langle \sigma_{\psi\bar{\psi}\to X\bar{X}} |v| \rangle (sY_{\mathrm{EQ}})^2\right]$$
$$= -\frac{x\langle \sigma_{\psi\bar{\psi}\to X\bar{X}} |v| \rangle s}{H(m)} \left(Y^2 - Y_{\mathrm{EQ}}^2\right). \tag{D.60}$$

Note, that if there are more annihilation channels, say to a final state F, we get additional terms in the equation, which are all of the form above but with $\langle \sigma_{\psi\bar{\psi}\to F} |v| \rangle$ instead, so that after summation over all of them we can describe the whole process simply by the total annihilation cross-section $\langle \sigma_{\mathrm{tot}} |v| \rangle$, which gives

$$\frac{dY}{dx} = -\frac{x\langle \sigma_{\mathrm{tot}} |v| \rangle s}{H(m)} \left(Y^2 - Y_{\mathrm{EQ}}^2\right). \tag{D.61}$$

Recall that $x = m/T$, so that the relativistic limit is $x \ll 3$, where we explicitly use 3 instead of 1 in order to even overcome the effects of expansion in the general Boltzmann equation (D.27). $x \ll 3$ now gives $\frac{dY}{dx}$ ~const., what implies that for a certain freeze-out x_f we can say $Y(x \lesssim x_f) \approx Y_{\mathrm{EQ}}$, i.e. freeze-out happens while the species is still relativistic, what we call *hot relics* since $xT \gg m$ for $x \ll 3$. However, in the relativistic limit we simply get

$$Y_{\mathrm{EQ}} = \frac{n_\psi^{\mathrm{EQ}}}{s} = \mathrm{const.}. \tag{D.62}$$

Let us now compute n_ψ and s in the relativist limit in general. We start with

$$n = \frac{g}{(2\pi)^3} = \int f(p,t) d^3p = \frac{g}{2\pi^2}\int_0^\infty f(p,t) p^2 dp, \tag{D.63}$$

where we integrated out the angles to get a factor 4π since our initial assumption is that the cosmological principle holds. As always, $\mu = 0$, so that

$$f(p,t) = \left[\exp\left\{\frac{E(p)}{T}\right\} \pm 1\right]^{-1} = \left[\exp\left\{\frac{\sqrt{p^2c^2 + m^2c^4}}{T}\right\} \pm 1\right]^{-1}$$
$$=: \left[\exp\{\sqrt{a^2 + b^2}\} \pm 1\right]^{-1}, \tag{D.64}$$

with

$$a := \frac{m}{T} \text{ and } b := \frac{p}{T}. \tag{D.65}$$

With $dp = Tdb$ we get

$$n = \frac{g}{2\pi^2}\int_0^\infty T^3 b^2 \left[\exp\{\sqrt{a^2+b^2}\} \pm 1\right]^{-1} db$$
$$= \frac{gT^3}{2\pi^2}\int_0^\infty \frac{b^2}{\exp\{\sqrt{a^2+b^2}\} \pm 1} = \frac{gT^3}{2\pi^2} \cdot \begin{cases} 2\zeta(3) & \text{bosons} \\ \frac{3}{4}\cdot 2\zeta(3) & \text{fermions} \end{cases}$$
$$= \frac{\zeta(3) g T^3}{\pi^2} \cdot \begin{cases} 1 & \text{bosons} \\ \frac{3}{4} & \text{fermions} \end{cases}, \tag{D.66}$$

where in the second equality we used the relativistic limit to evaluate the integral. Now, we want to compute the entropy density, s. Therefore, we start with the first law of thermodynamics

$$dU = TdS - PdV + \mu dN \overset{\mu=0}{\Rightarrow} dS = \frac{1}{T}(dU + PdV)$$
$$= \frac{1}{T}(d(\rho V) + PdV) = \frac{1}{T}(d((\rho + P)dV) - VdP), \tag{D.67}$$

where we used $U = \rho V$ and the second equality in the second line added $VdP - VdP = 0$. To derive an expression for dP we start with the first equality in the second line of the above equation

$$dS = \frac{1}{T}(d(\rho V) + PdV) = \frac{1}{T}\left(V\frac{d\rho}{dT}dT + (\rho + P)dV\right). \tag{D.68}$$

By comparison with the chain rule for S

$$dS(V,T) = \left(\frac{\partial S}{\partial T}\right)_V + \left(\frac{\partial S}{\partial V}\right)_T dV, \tag{D.69}$$

we can read off two equations

$$\left(\frac{\partial S}{\partial T}\right)_V = \frac{1}{T}\frac{d\rho}{dT}V \Rightarrow \frac{\partial^2 S}{\partial V \partial T} = \left(\frac{\partial}{\partial V}\right)_T \left(\frac{\partial S}{\partial T}\right)_V = \frac{1}{T}\frac{d\rho}{dT} \text{ and} \tag{D.70}$$

$$\begin{aligned}\left(\frac{\partial S}{\partial V}\right)_T = \frac{1}{T}(\rho + P) \Rightarrow \frac{\partial^2}{\partial T \partial V} &= \left(\frac{\partial}{\partial T}\right)_V \left(\frac{\partial S}{\partial V}\right)_T \\ &= \frac{1}{T}\left[\frac{d\rho}{dT} + \frac{dP}{dT}\right] - \frac{1}{T^2}(\rho + P),\end{aligned} \tag{D.71}$$

which we can set equal to get

$$\frac{1}{T}\frac{d\rho}{dT} = \frac{1}{T}\left[\frac{d\rho}{dT} + \frac{dP}{dT}\right] - \frac{1}{T^2}(\rho + P) \Rightarrow dP = \frac{\rho + P}{T}dT. \tag{D.72}$$

Plugging this back in (D.67) gives

$$dS = \frac{1}{T}d\left[(\rho + P)V\right] - \frac{V}{T^2}(\rho + P)dT = d\left[\frac{\rho + P}{T}V\right]. \tag{D.73}$$

Recall that the entropy density is apparently defined as $s := S/V$, so that $dS = d(sV)$ and we can read s off directly as

$$s = \frac{\rho + P}{T}. \tag{D.74}$$

In the radiation-dominated epoch we have $P = \frac{1}{3}\rho$, so that $s = \frac{4}{3}\frac{\rho}{T}$. Let us compute the energy density ρ in the relativistic limit, which by now has become clear how this is done by starting at (D.11) and plugging in $\mu = 0$ along with b as defined in (D.65) and then performing the limit . However, we will arrive at

$$\rho = \frac{gT^4}{2\pi^2}\int_0^\infty \frac{b^3}{e^b \pm 1} = \frac{g\pi^2 T^4}{30}\begin{cases}1 & \text{bosons} \\ 7/8 & \text{fermions}\end{cases}. \tag{D.75}$$

Back to s, by considering multiple species with individual energy densities we get

$$s = \frac{4}{3}\frac{\pi^2}{30}\left[\sum_{i \in \text{bosons}} g_i T_i^3 + \frac{7}{8}\sum_{i \in \text{fermions}} g_i T_i^3\right] =: \frac{4}{3}\frac{\pi^2}{30}g_{*,s}T^3. \tag{D.76}$$

With n_ψ^{EQ} and s at hand we can compute

$$Y_{\mathrm{EQ}} = \frac{n_\psi^{\mathrm{EQ}}}{s} = \frac{\frac{\zeta(3)g}{\pi^2}T^3}{\frac{4}{3}\frac{\pi^2}{30}T^3 g_{*,s}} \cdot \begin{cases}1 & \text{bosons} \\ \frac{3}{4} & \text{fermions}\end{cases} = \frac{45\zeta(3)}{2\pi^4}\frac{g}{g_{*,s)}} \cdot \begin{cases}1 \\ \frac{3}{4}\end{cases}. \tag{D.77}$$

For photons (γ) that only can have two different polarizations we have $g = 2$ and can use $T = T_{\text{CMB},0}$ along with the fact that they are bosons to get

$$Y_{\text{EQ}}^{\gamma} = \frac{45\zeta(3)}{2\pi^4} \frac{2}{g_{*,s}(T_{\text{CMB}},0}. \tag{D.78}$$

By knowing $\zeta(3) \approx 1.20$ we get

$$Y_{\text{EQ}}^{\gamma} \approx 0.278 \frac{2}{g_{*,s}(T_{\text{CMB},0})}. \tag{D.79}$$

For axions (a) we have to consider a slightly different process since they are produced in collisions $a + 1 \leftrightarrow 2 + 3$ rather than $a + \bar{a} \leftrightarrow 2 + 3$. The main consequence is that in our derivation of the Boltzmann equation instead of Y^2 we get $Y \cdot Y_{\text{EQ}}$ since species 1 is also assumed to be in thermal equilibrium. Recalling

$$\frac{dY}{dx} = -\frac{\langle \sigma_{\text{tot}} |v| \rangle s}{x H(x)} (Y Y_{\text{EQ}} - Y_{\text{EQ}}^2) = -\frac{\Gamma_{\text{tot}}}{xH}(Y - Y_{\text{EQ}}), \tag{D.80}$$

where in the first equality we used $H(m) = x^2 H(x)$ and in the second equality we used $\Gamma := \langle \sigma_{\text{tot}} |v| \rangle \cdot n$ of species 1 with $n = Ys$ and $H(x) \equiv H$. we can now solve this differential equation to obtain

$$\int_{Y(0)}^{Y_\infty} \frac{dY'}{Y' - Y_{\text{EQ}}} = -\int_0^x \frac{\Gamma_{\text{tot}}}{x'H} \Rightarrow \ln(Y' - Y_{\text{EQ}})|_{Y(0)}^{Y_\infty} = -\int_0^x \frac{\Gamma_{\text{tot}}}{x_H} dx'$$
$$\Rightarrow \frac{Y(x) - Y_{\text{EQ}}}{Y(0) - Y_{\text{EQ}}} = -\int_0^x \frac{\Gamma_{\text{tot}}}{x'H} dx' \Rightarrow \frac{1 - \frac{Y(x)}{Y_{\text{EQ}}}}{1 - \frac{Y(0)}{Y_{\text{EQ}}}} = -\int_0^x \frac{\Gamma_{\text{tot}}}{x'H} dx'. \tag{D.81}$$

We assume that initially there are no axions, so that $Y(0) = 0$ and thus, $1_Y(0)/Y_{\text{EQ}} = 1$, so that

$$1 - \frac{Y(x)}{Y_{\text{EQ}}} = -\int_0^x \frac{\Gamma_{\text{tot}}}{x'H} dx'$$
$$\Rightarrow Y(x) = Y_{\text{EQ}} \left[1 - \exp\left\{ -\int_0^x \frac{\Gamma_{\text{tot}}}{x'H} dx' \right\} \right]. \tag{D.82}$$

Again, we consider the relativistic limit, $x \ll 1$, so that $\exp(-1/x) \ll 1$ and $Y(x) \approx Y_{\text{EQ}}$, which is again given by

$$Y_{\text{EQ}}^a = \frac{n_a^{\text{EQ}}}{s} \approx 0.278 \frac{g}{g_{*,s}} \begin{cases} 1 & \text{bosons} \\ \frac{3}{4} & \text{fermions} \end{cases}. \tag{D.83}$$

Axions are pseudo Goldstone bosons and thus, the axion field has only one degree of freedom, namely the value of θ, so that $g = 1$ and we have

$$Y^a_{\rm EQ} \approx \frac{0.278}{g_{*,s}}. \tag{D.84}$$

Denote the temperature at axion decoupling as $T^a_{\rm dec}$, then with $n_a^{\rm EQ} = Y^a_{\rm EQ} s$ we get

$$n_a^{\rm EQ} = \frac{0.278}{g_{*,s}(T^a_{\rm dec})}. \tag{D.85}$$

With $n_\gamma^{\rm EQ} \equiv n_\gamma$ and $n_a^{\rm EQ} \equiv n_a$ we finally arrive at

$$\frac{n_a}{n_\gamma} = \frac{1}{2}\frac{g_{*,s}(T_{\rm CMB,0})}{g_{*,s}(T^a_{\rm dec})} \tag{D.86}$$

as we sought after in the beginning of this appendix

Appendix E

Brief review of the theory of inflation

This appendix is fully devoted to a very brief review of the most important phenomenology and basic equations of nowadays's well-accepted *slow-roll inflation* as this is the kind of inflation we are considering in this work. Historically seen, inflation was invented independently by Starobinsky[43] and Guth[44][1]. Both of them and many others, attended a conference in 1982, organized by Hawkins and Gibbons on the very early Universe, what Guth describes in [46]. During this conference, Wilczek is quoted to have said "The idea is so simple, and yet it provides a qualitative understanding of some of the deepest puzzles of cosmology!"[46]. Let me note that by now, the original inflation models had to be reviewed several times in order to fit into the well-accepted picture of the Universe's evolution and properties. However, as stated in the beginning, we are not interested in the full story, for which one could work through the comprehensive review by Langlois[47].

First of all, why should we think about an additional era in the Universe's history? The ΛCDM-model described in appendix A suffers three major problems we would like to discuss now. First, there are strong evidences

[1] Vilenkin made the following, interesting historical note on that. "In fact, Starobinsky's paper appeared before Guth [...] suggested the standard version of inflation, although it was Guth who fully explained the advantages of inflationary scenarios."[45]

pointing to the Universe being flat[48], so that

$$1 = \sum_i \Omega_{i,0} = \Omega_{\gamma,0} + \Omega_{m,0} + \Omega_{\Lambda,0} = \Omega_{\text{tot}}, \tag{E.1}$$

where the subscript stands for total, must be satisfied. Recall the Friedmann equation (A.3)

$$H^2 = \frac{8\pi G \rho_{\text{tot}}}{3} - \frac{\kappa}{a^2}, \tag{E.2}$$

where we have used the definition of the Hubble parameter, $H := \dot{a}/a$, and the usual convention $a_0 = 1$ for convenience. Now let us replace the density by the density parameter via the critical density, $\rho_{\text{crit}} = \frac{3H^2}{8\pi G}$,

$$H^2 = H^2 \Omega_{\text{tot}} - \frac{\kappa}{a^2}, \tag{E.3}$$

what gives

$$\Omega_{\text{tot}} - 1 = \frac{\kappa}{a^2 H^2}. \tag{E.4}$$

As stated above, we must fulfill a flat geometry, so $\kappa \stackrel{!}{=} 0$, what leads to the condition, that $\Omega_{\text{tot}} = 1$ precisely. In fact, this solution appears to be unstable, for what Liddle[38] gave a very short argument, namely the Friedmann equations give the time dependencies

$$(a^2 H^2)^{-1} \sim t \quad \text{and} \quad (a^2 H^2)^{-1} \sim t^{\frac{2}{3}} \tag{E.5}$$

for a radiation- and matter-dominated Universe, respectively. As derived above, $(a^2 H^2)^{-1} \sim \Omega_{\text{tot}} - 1$, so that we see that the deviation of Ω_{tot} from one is increasing with time, simply meaning, that either $\Omega_{\text{tot}} = 1$ precisely or the Universe cannot be observed to be flat today. The question is how Ω_{tot} can be fine-tuned so precisely.

Second, without any doubt, the speed of light is finite, so that a photon emitted at the Big Bang can only have travelled a finite distance till today since the age of the Universe is finite as well. It is well-measured that the CMB can be considered isotropic and the CMB-photons all have the same temperature[38]T_0. Analogue to the argument we made in subsection 2.2, every patch of the observable Universe must have been in causal contact in order to achieve thermal equilibrium, so that the temperature is overall the same. This breaks down when we observe photons coming from opposite sides of the sky, since their corresponding patches clearly cannot have been in causal contact, since their emitted photons are reaching us today, not even reaching the other side of the sky to exchange any information[38]. The question is how these patches were able to be in causal contact in the early Universe, although today they appear not to be.

Third, in the very early Universe, when Grand-Unified-Theories (short: GUT) seem to be applicable, there is no reason to exclude magnetic monopoles from the particle picture as we are used to do. We do not need to go into detail here, but one can show that

$$\Omega_{\text{mono}} \sim 10^{13}, \tag{E.6}$$

what is disastrous as this is not only exceeding $\Omega_{\text{tot}} = 1$ by far, but the question arises where all these monopoles are since we were not able to detect them yet. This question was initially the purpose for Guth to invent his theory of inflation[38].

Let us now turn to the question how inflation is able to elegantly fix these problems. Suppose, there is an era in the radiation-dominated epoch, that expands the Universe exponentially, i.e. the expansion is accelerated. Hence, recall the acceleration equation (A.4)

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \overset{G=m_{\text{pl}}^{-2}}{=} -\frac{4\pi}{m_{\text{pl}}^2}(\rho + 3P), \tag{E.7}$$

where we have replaced G by the corresponding Planck-mass as we assume inflation to be in the very early Universe. In order to describe an acceleration, we demand $\ddot{a} > 0$ and since the prefactor of the parentheses is constant and negative and the prefactor of $\ddot{a}$ is strictly positive, we demand $\rho + 3P < 0$, what immediately gives

$$\rho + 3P < 0 \Leftrightarrow \rho < -3P \overset{P=w\rho}{=} -3w\rho \overset{:(-3\rho)}{\Rightarrow} -\frac{1}{3} > w \tag{E.8}$$

as the equation of state parameter for an accelerated period of expansion. As before, we could use $w = -1$ as this gives a constant energy density, which is convenient for the computations. Let this expansion be expressed by the exponential scaling law

$$a_f = a_i \cdot e^N, \tag{E.9}$$

where a_i and a_f are the initial and final scale factor, respectively and we call N the *number of e-foldings* in the inflationary expansion. We assume the temperature of the Universe at the start of inflation, T_i, to be the same as at the end of inflation, T_f, because although we know that $a \sim T^{-1}$ leads to an exponential reduction in the Universe's temperature during inflation, the kinetic energy to drive inflation is then converted back into heat, so that the temperature should be the same. This process is called *reheating*. What is meant by "kinetic energy to drive inflation" will become clearer later in this appendix. However, recall (E.4) and note that since $w = -1$ gives a constant energy density for the era of inflation, the Friedmann equation (A.3) implies that H^2 =const., so that we get the scaling

$$\Omega_{\text{tot}} - 1 = \frac{\kappa}{a^2 H^2} \sim a^{-2}. \tag{E.10}$$

Now, we take $\Omega_{\text{tot}} = \Omega_i$ at the time, when inflation occurs and $\Omega_{\text{tot}} = \Omega_f$ at the time, when inflation ends to get

$$\frac{\Omega_i - 1}{\Omega_f - 1} = \frac{a_f^2}{a_i^2} \overset{(E.9)}{=} e^{2N}. \tag{E.11}$$

Along with the observational constraint[48] $|\Omega_0 - 1| < 10^{-2}$ one can compute $|\Omega_f - 1| \lesssim 10^{-53} \approx e^{-122}$ and thus with the above equation we obtain

$$\Omega_i - 1 = e^{2N}(\Omega_f - 1) \lesssim e^{2N} \cdot e^{-122}, \tag{E.12}$$

so that $N = 61$ is needed as we only need to consider the absolute values, so that $\Omega_i \approx 0$ drops out. This means, that we need at least 61 e-foldings to solve the flatness problem. The next problem is the question how different patches of the Universe could have been in causal contact, even though they appear not to be so today. By using the RW-metric (D.3) for a flat Universe and considering a light ray, i.e. a radially propagating photon with $ds^2 = 0$ and $d\Omega = 0$

$$0 = ds^2 = -dt^2 + a^2 dr^2 \Leftrightarrow dt = -rdr, \tag{E.13}$$

where we use the negative solution to describe outgoing light rays, we can compute the comoving horizon size at the beginning of inflation, t_i, to be

$$r_H = \int_{r_0}^{r_i} dr \overset{(E.13)}{=} \int_0^{t_i} \frac{dt}{a} = \int_0^{t_i} \frac{dt}{a_i \left(\frac{t}{t_i}\right)^{1/2}} = \frac{2t_i}{a_i} \overset{(E.9)}{=} \frac{2t_i}{a_f} e^N. \tag{E.14}$$

Note, that we used $a \sim t^{-1/2}$, what holds for the radiation-dominated epoch. We can approximate $t_i \sim t_{\text{GUT}}$, but do not know a_f yet. Above, we explained that due to reheating, we assume $T_i = T_f$ and analogue to the time-approximation we can then approximate $T_f \sim T_{\mathrm{GUT}}$, so that along with $a \sim T^{-1}$ we can rewrite a_f as

$$a_f \overset{a_0=1}{=} \frac{a_f}{a_0} = \frac{T_0}{T_f} \sim \frac{T_0}{T_{\mathrm{GUT}}}, \tag{E.15}$$

what after plugging in numbers gives

$$a_f \sim 3 \cdot 10^{-28}, \tag{E.16}$$

so that we can approximate the horizon size at the beginning of inflation to be

$$r_H(t_i) \sim \frac{2 t_{\mathrm{GUT}} T_{\mathrm{GUT}}}{T_0} e^N \sim 2 \cdot 10^2 \cdot e^N \,\mathrm{cm}. \tag{E.17}$$

From the CMB power spectrum we can infer the comoving scale, λ, that corresponds to the diameter of the last-scattering surface centered on us to be

$$\lambda \sim 10^{29} \,\mathrm{cm}. \tag{E.18}$$

To be in causal contact, these scales have to be inside the horizon, i.e.

$$10^{29} \,\mathrm{cm} \sim \lambda \overset{!}{<} r_H(t_i) \sim 2 \cdot 10^2 \cdot e^N \,\mathrm{cm} \Rightarrow e^N \gtrsim 10^{27} \approx e^{62}, \tag{E.19}$$

so that we need at least 62 e-foldings, in order for all patches of the horizon to be in causal contact in the early Universe. Due to the exponential expansion, these patches get largely separated an thus, they do not have to be causal contact necessarily, what is exactly what we observe today. Finally, by taking the slightly stronger constraint, $N > 62$, we can compute the density reduction of magnetic monopoles to be

$$\frac{\rho_i^{\mathrm{mono}}}{\rho_f^{\mathrm{mono}}} \sim \frac{V_f}{V_i} \sim \frac{a_f^3}{a_i^3} = e^{3N} > e^{3 \cdot 62} = e^{186}. \tag{E.20}$$

This underlines very clearly how small ρ_f^{mono} is and explains why no magnetic monopoles have been detected yet. If they exist, they are simply too rare. Note, that this argument assumes the number of magnetic monopoles to be constant. In 1982, Linde[49] presented, based on Guth's initial inflation model a new model that was able to solve not only the monopole but also the flatness and horizon problem. Shortly after, but independent of, Linde, the same problems were solved by Albrecht and Steinhardt[50] by considering first order phase-transitions in GUT.

Now, that we saw how inflation can solve the big problems we faced in the beginning, let us finally turn to the question how we could model such an era of rapid acceleration. Guth[44] and Starobinsky[43] proposed a single scalar field, ϕ, with the inflation potential $V(\phi)$. When ϕ is deflected from the true vacuum state, i.e. the potential minimum, it simply propagates back to it. Imagine for simplicity a $V(\phi) = \phi^2$ potential and ϕ as a ball, then if $\phi \neq 0$ the ball will roll down the potential hill back to $\phi = 0$. In general the EOM for an arbitrary $V(\phi)$ can be written[37]

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV(\phi)}{d\phi} = 0. \tag{E.21}$$

Note, that one could add a term $-\nabla^2\phi$, where ∇ is taken with respect to proper spatial coordinates. In inflationary models, this term is considered negligible since one considers a small region of the Universe that is then inflated, what typically wipes out all small-scale fluctuations, so that the derivations in spatial directions basically do matter and we can assume that ϕ is the same everywhere[37]. Analogue to (2.68) we can write

$$P = \frac{1}{2}\dot{\phi} - V(\phi) \quad \text{and} \quad \rho = \frac{1}{2}\dot{\phi}^2 + V(\phi). \tag{E.22}$$

Recall, that we argued above that $w < -\frac{1}{3}$ or, equivalently, $3P < -\rho$ must hold for an accelerated era of expansion. From the equations above we then interfere

$$\frac{3}{2}\dot{\phi}^2 - 3V(\phi) < -\frac{1}{2}\dot{\phi}^2 - V(\phi) \Rightarrow \dot{\phi}^2 < V(\phi), \tag{E.23}$$

so the potential-term has to dominate over the kinetic term initially. Now, consider an infinitesimal time interval, δt, over which $\delta\phi$ is the corresponding deviation from ϕ, then we can associate to this deviation a kinetic term of the form $\phi^2/(\delta t)^2$. Since $V(\phi)$ is bigger than the kinetic term, we get

$$V(\phi) > \frac{\phi^2}{\delta t^2} \Rightarrow \frac{dV(\phi)}{d\phi} > \frac{2\phi}{\delta t^2} \sim \mathcal{O}\left(\ddot{\phi}\right), \tag{E.24}$$

so that one can neglect the $\ddot{\phi}$-term in the EOM (E.21) leading to

$$3H\dot{\phi} = -\frac{dV(\phi)}{d\phi} \tag{E.25}$$

is the so-called *slow-roll approximation* and the corresponding model of inflation is what we call *slow-roll inflation*. This simple argument given by [37] is rather a motivation than an actual proof, but it is sufficient for our purpose. Recall the Friedmann equation (A.3), where we again replace $G = m_{\mathrm{pl}}^{-2}$, with a constant energy density, corresponding to $w = -1$, which we replace now by the dominating potential term to approximate

$$H^2 \approx \frac{8\pi}{3m_{\mathrm{pl}}^2} V(\phi). \tag{E.26}$$

With this approximation and the above results, one can show, that the slow-roll approximation can be expressed by two dimensionless slow-roll parameters[37]

$$\epsilon_{\mathrm{inf}} = -\frac{\dot{H}}{H^2} = \frac{m_{\mathrm{pl}}^2}{16\pi}\left(\frac{V'}{V}\right)^2 \ll 1 \quad \text{and} \quad \eta_{\mathrm{inf}} = \frac{\ddot{\phi}}{H\dot{\phi}} = \frac{m_{\mathrm{pl}}^2}{8\pi}\left(\frac{V''}{V}\right) \ll 1, \tag{E.27}$$

whose derivation is not interesting for us[2]. Note, that a prime denotes the derivative with respect to ϕ. In fact, the dimensionless parameters basically say, that $V(\phi)$ has to be very flat to ensure the slow-roll aspect and very shallow to ensure a small duration of inflation. Note additionally, that different inflation models assume different potential terms leading to different ways to arrive at and interpret the outcomes of inflation. However, in an intuitive version, think about a plateau that is very slightly tilted and at some point has a deep potential well with the true vacuum as its minimum. Now think again of ϕ as a ball, slowly rolling down the plateau, picking up kinetic energy[3] along the way and finally falling down the hill into the potential minimum, overshooting it and then begin to oscillate around it. The latter fits perfectly well to the EOM (E.21) because it is of the same form as a harmonic oscillator with an additional friction term corresponding to the Universe's expansion.

One outcome of inflation is the production of Gravitational Waves (short: GW). We do not go into the production mechanism, but refer to [51] for a comprehensive introduction to and discussion of the topic. In 1977, Lyth pointed out that if GW are an outcome of inflation, their detection could

[2]Note, that Marsh[1] defines the slow-roll parameters with an additional factor 2π, which is just a rescaling that does not influence the behavior of $V(\phi)$.

[3]This is what was meant by the kinetic energy that is used to drive inflation and that is given back afterwards.

be used to constrain the inflation potential for instance, but the contemporary observational bounds on the tensor-to-scalar ratio, r_T, imply a potential depth several orders of magnitude larger than one should expect and additionally the connection to GW is only possible if the fluctuations $\delta\phi$ are of order of the Planck-scale, which would imply that inflation cannot be an ordinary extension of the standard model[25]. The bound on $\delta\phi$ has become known as the *Lyth-bound*. In 2021 Cai et al.[52] presented a new mechanism in order to produce sufficiently large GW that produce observable tensor modes within the Lyth-bound by proposing that there is the massive ϕ field that interacts with a massless scalar χ field. While the fluctuations of the ϕ field provides the necessary source of the GW, the χ field generates ”the observed nearly scale-invariant power spectrum for curvature perturbations, which are shown to be within current observational bounds”[52]. With this mechanism it is indeed possible to use GW-observations to put constraints on the inflation potential[52].

References

[1] D. J. Marsh, *Axion cosmology*, Physics Reports **643**, 1 (2016).

[2] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger, *Theory of cosmological perturbations*, Physics reports **215**, 203 (1992).

[3] G. Münster, *Von der Quantenfeldtheorie zum Standardmodell*, de Gruyter, 1st edition, 2019.

[4] S. L. Adler, *Axial-Vector Vertex in Spinor Electrodynamics*, Phys. Rev. **177**, 2426 (1969).

[5] W. A. Bardeen, *Anomalous Ward Identities in Spinor Field Theories*, Phys. Rev. **184**, 1848 (1969).

[6] J. S. Bell and R. W. Jackiw, *A PCAC puzzle: $\pi^0 \to \gamma\gamma$ in the σ-model*, Nuovo cimento **60**, 47 (1969).

[7] R. D. Peccei, *The Strong CP Problem and Axions*, in *Lecture Notes in Physics*, pages 3–17, Springer Berlin Heidelberg, 2008.

[8] G. 't Hooft, *Symmetry Breaking through Bell-Jackiw Anomalies*, Phys. Rev. Lett. **37**, 8 (1976).

[9] S. Weinberg, *A new light boson?*, Physical Review Letters **40**, 223 (1978).

[10] F. Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, Physical Review Letters **40**, 279 (1978).

[11] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38**, 1440 (1977).

[12] M. Peskin and D. Schröder, *An Introduction to Quantum Field Theory*, CRC Press, 2018.

[13] V. Rubakov and D. Gorbunov, *Introduction to the Theory of the early Universe - Hot Big Bang Theory*, World Scientific Publishing Co. Pte. Ltd., 2nd edition, 2018.

[14] R. D. Peccei and H. R. Quinn, *Constraints imposed by* CP *conservation in the presence of pseudoparticles*, Phys. Rev. D **16**, 1791 (1977).

[15] M. Srednicki, *Axion couplings to matter:(I). CP-conserving parts*, Nuclear Physics B **260**, 689 (1985).

[16] J. E. Kim, *Light pseudoscalars, particle physics and cosmology*, Physics Reports **150**, 1 (1987).

[17] R. Peccei and H. R. Quinn, *Some aspects of instantons*, Il Nuovo Cimento A (1965-1970) **41**, 309 (1977).

[18] G. Hooft, *How instantons solve the U (1) problem.*, Physics Reports **142**, 357 (1986).

[19] K. Mimasu and V. Sanz, *ALPs at colliders*, Journal of High Energy Physics **2015**, 1 (2015).

[20] L. Edelhäuser and A. Knochel, *Tutorium Quantenfeldtheorie*, Springer Spektrum, 1st edition, 2016.

[21] G. W. Gibbons and S. W. Hawking, *Cosmological event horizons, thermodynamics, and particle creation*, Physical Review D **15**, 2738 (1977).

[22] M. Tristram et al., *Improved limits on the tensor-to-scalar ratio using BICEP and P l a n c k data*, Physical Review D **105**, 083524 (2022).

[23] A. Wagner et al., *Search for hidden sector photons with the ADMX detector*, Physical review letters **105**, 171801 (2010).

[24] R. Cervantes et al., *ADMX-Orpheus first search for 70 μeV dark photon dark matter: Detailed design, operations, and analysis*, Physical Review D **106**, 102002 (2022).

[25] D. H. Lyth, *What would we learn by detecting a gravitational wave signal in the cosmic microwave background anisotropy?*, Physical Review Letters **78**, 1861 (1997).

[26] E. Pajer and M. Peloso, *A review of Axion Inflation in the era of Planck*, Classical and Quantum Gravity **30**, 214002 (2013).

[27] M. Schumann, *Direct detection of WIMP dark matter: concepts and status*, Journal of Physics G: Nuclear and Particle Physics **46**, 103003 (2019).

[28] E. W. Kolb and M. S. Turner, *The early universe*, CRC press, 2018.

[29] S. M. Carroll, *An Introduction to General Relativity - Spacetime and Geometry*, Cambridge University Press, 2020.

[30] R. Hlozek, D. Grin, D. J. Marsh, and P. G. Ferreira, *A search for ultralight axions using precision cosmological data*, Physical Review D **91**, 103512 (2015).

[31] D. J. Marsh and P. G. Ferreira, *Ultralight scalar fields and the growth of structure in the Universe*, Physical Review D **82**, 103528 (2010).

[32] D. H. Lyth and D. Wands, *Generating the curvature perturbation without an inflaton*, Physics Letters B **524**, 5 (2002).

[33] Z. Hou, R. Keisler, L. Knox, M. Millea, and C. Reichardt, *How massless neutrinos affect the cosmic microwave background damping tail*, Physical Review D **87**, 083008 (2013).

[34] C. Hogan and M. Rees, *Axion miniclusters*, Physics Letters B **205**, 228 (1988).

[35] C.-P. Ma and E. Bertschinger, *Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges*, The Astrophysical Journal **455** (1995).

[36] H. Mo, F. van den Bosch, and S. White, *Galaxy Formation and Evolution*, Cambridge University Press, 7th edition, 2021.

[37] S. Serjeant, *Observational Cosmology*, Cambridge University Press, 1st edition, 2010.

[38] A. Liddle, *An Introduction to Modern Cosmology*, Wiley, 2nd edition, 2008.

[39] B. Ryden, *Introduction to Cosmology*, Cambridge University Press, 2nd edition, 2017.

[40] S. Coleman, *Aspects of Symmetry*, Cambridge University Press, New York, United States of America, 1985.

[41] G. Arfken and H. Weber, *Mathematical Methods for Physicists*, Elsevier Academic Press, 6th edition, 2005.

[42] R. Bott, *An application of the Morse theory to the topology of Lie-groups*, Bulletin de la Société Mathématique de France **84**, 251 (1956).

[43] A. A. Starobinsky, *A new type of isotropic cosmological models without singularity*, Physics Letters B **91**, 99 (1980).

[44] A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, Physical Review D **23**, 347 (1981).

[45] A. Vilenkin, *Classical and quantum cosmology of the Starobinsky inflationary model*, Physical Review D **32**, 2511 (1985).

[46] A. H. Guth, *The inflationary universe : the quest for a new theory of cosmic origins*, Addison-Wesley Publishing, Reading, Mass. [etc.], 1997.

[47] D. Langlois, *Inflation and cosmological perturbations*, in *Lectures on Cosmology: Accelerated Expansion of the Universe*, pages 1–57, Springer, 2010.

[48] G. Efstathiou and S. Gratton, *The evidence for a spatially flat Universe*, Monthly Notices of the Royal Astronomical Society: Letters **496**, L91 (2020).

[49] A. D. Linde, *A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Physics Letters B **108**, 389 (1982).

[50] A. Albrecht and P. J. Steinhardt, *Cosmology for grand unified theories with radiatively induced symmetry breaking*, Physical Review Letters **48**, 1220 (1982).

[51] H. An, K.-F. Lyu, L.-T. Wang, and S. Zhou, *Gravitational waves from an inflation triggered first-order phase transition*, Journal of High Energy Physics **2022**, 1 (2022).

[52] Y.-F. Cai, J. Jiang, M. Sasaki, V. Vardanyan, and Z. Zhou, *Beating the Lyth bound by parametric resonance during inflation*, Physical Review Letters **127**, 251301 (2021).