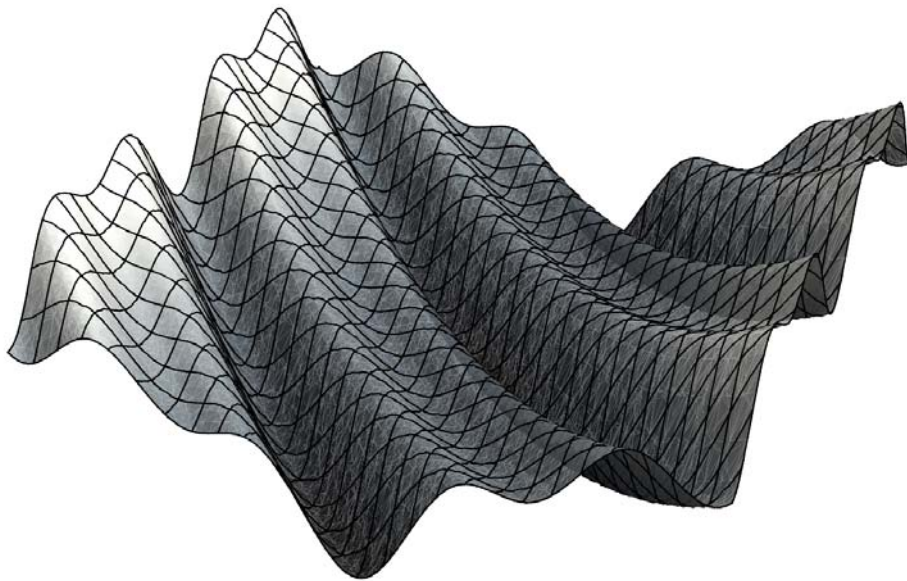


Vacuum structure of the strong interaction with a Peccei-Quinn symmetry

Wilco J. den Dunnen

November 27, 2008



Master's thesis
Supervisor: Dr. Daniël Boer
Second reader: Dr. Rick Bethlem
Theoretical Physics Group
Department of Physics and Astronomy
Faculty of Sciences

vrije Universiteit amsterdam



Abstract

In the context of the observed matter – anti-matter asymmetry in the universe it is important to explore all possible sources of CP -violation. In the strong interaction there is no CP -violation observed and the lack of an explanation for this is known as the strong CP -problem. The only known, reasonable solution to the strong CP -problem – that is not ruled out – is through the Peccei-Quinn mechanism. This mechanism forces explicit and spontaneous CP -violation in QCD to vanish, irrespective of the value of the θ -parameter, whenever the theory has a specific Peccei-Quinn (PQ) symmetry. On the other hand, it is not excluded that this PQ mechanism still allows a CP -violating metastable vacuum state, which could cause CP -violation when the fields ‘hang’ temporarily in this false vacuum. By using a single-flavor low-energy effective model approximation, the vacuum structure of the strong interaction with a PQ symmetry is explored. It is shown that no metastable vacuum state appears for reasonable parameter choices. Nevertheless, it is argued on the basis of an explicit example in the literature that at high temperatures (near the chiral transition) a metastable state *will* appear, albeit a CP -conserving one. Such a state is nonetheless interesting, as it might yield extra experimental signatures of a PQ symmetry. This might open up another way to test whether nature has a PQ symmetry, besides the search for the axion.

Contents

Abstract	iii
1 Introduction	1
1.1 Outline	2
2 Introduction to QCD	5
2.1 Free fermions	5
2.2 Non-Abelian gauge theory	6
2.3 Running of the coupling constant	9
2.4 Symmetries in QCD	11
2.4.1 Discrete symmetries	11
2.4.2 Continuous Symmetries	11
3 Instanton solutions	13
3.1 Homotopy classes	13
3.1.1 Definition of a homotopy class	13
3.1.2 Example: $S^1 \rightarrow S^1$	13
3.1.3 Example: $S^3 \rightarrow SU(n)$	14
3.1.4 Composition of functions	14
3.2 Non-trivial finite Euclidean action gauge fields	14
3.3 Winding-number vacua	16
3.4 Instantons	18
3.5 Example: instanton with winding number one	20
4 Consequences of instantons	23
4.1 Chirality non-conservation	23
4.2 θ -vacua	25
4.3 The strong CP -problem	27
4.4 Energy of the θ -vacua	27
5 The Peccei-Quinn mechanism	31
5.1 Basics of the Peccei-Quinn mechanism	32
5.2 Example: the Peccei-Quinn mechanism in a toy model	33
5.3 Axions in the toy model	38
5.4 Problems with the standard Peccei-Quinn model	41
5.5 Alternative models with a Peccei-Quinn symmetry	42
6 Low-energy effective models for QCD	45
6.1 The $SU(3) \otimes SU(3)$ non-linear sigma model	46
6.2 The $U(3) \otimes U(3)$ non-linear sigma model	48
6.3 Dashen's phenomenon and metastable CP -violating states	48

7	Non-linear sigma model with a Peccei-Quinn symmetry	53
7.1	Single-flavor non-linear sigma model with a PQ symmetry	54
7.2	Next-to-leading order non-linear sigma model with a PQ symmetry	55
8	Discussion	59
8.1	Applicability of a single-flavor model	59
8.2	Expectations for other models	60
9	Conclusion	63
A	Conventions used	65
	Bibliography	67
	Acknowledgements	71

Chapter 1

Introduction

One of the most puzzling problems in contemporary physics is the asymmetry between particles and anti-particles; why is the visible universe made up solely of matter? On the microscopic scale, particles and anti-particles behave almost exactly the same way. So what causes this large discrepancy on the cosmological scale?

The symmetry of nature with respect to interchanging particles with anti-particles is called C symmetry. One essential ingredient to explain the observed asymmetry between particles and anti-particles (baryon asymmetry) is C violation. Both the electromagnetic and strong interaction are symmetric under C and P (parity) transformations. On the contrary, the weak interaction, responsible for e.g. the decay of nuclei, does *not* have a C symmetry.

However, a violation of C symmetry alone is *not enough* to explain the baryon asymmetry in the universe. Sakharov discovered already in 1967 [1] that a violation of CP symmetry is *also* necessary, where CP is a C transformation combined with a spatial reflection, P . The electromagnetic and strong interaction are both known so far to be CP symmetric. At the same time, the weak interaction has only a very small violation of CP symmetry, which was measured by Cronin and Fitch (Nobel prize 1980) and put into theory by Kobayashi and Maskawa (Nobel prize 2008). Nevertheless, this small amount is not enough to explain the baryon asymmetry. The baryon asymmetry is one of the most important reasons for the ongoing interest of both experimental and theoretical physicists in the phenomenon of CP violation.

One important theoretical problem is the *absence* of CP violation in the strong interaction. The current theory used to describe the strong interaction is quantum chromodynamics (QCD). Unfortunately, it is notoriously difficult to do any calculation involving low ($< 1\text{GeV}$) energies in this theory. Because of that, not all aspects of QCD are yet fully understood. One thing that has become known is the fact that QCD is, in principle, not CP symmetric. There is one parameter, θ , characterizing the ‘amount’ of CP violation in QCD. One would expect this parameter to be of order one, but experiments¹ showing the absence of CP violation in the strong interaction imply an upper limit of $\theta < 10^{-10}$. The explanation that nature chose such an extremely small value by accident is very *unnatural*. It seems that there should be some reason for the absence of CP violation in the strong interaction, i.e. for θ to be zero.

The unnatural smallness of θ is usually referred to as the *strong CP problem*. So far no satisfying solution has been found. The only solution that has been put forward and has not been ruled out is through the Peccei-Quinn (PQ) mechanism². This mechanism makes CP conservation automatic if the theory has a specific extra symmetry, often called accordingly a Peccei-Quinn symmetry. One important side-effect of adding a PQ symmetry to QCD is the prediction of a new light ($m < 1\text{MeV}$) boson, called the *axion*. The hypothetical axion has been searched for since 1978, but has not been found yet. Restrictions on the possible existence of the axion have been derived from experiments, astrophysical and cosmological data. The common understanding is that axion masses outside the window $10\text{meV} - 10\mu\text{eV}$ are excluded, but this exclusion range depends heavily on the specific implementation of the PQ

¹See e.g. measurements of the electric dipole moment of the neutron in [2], which can be related to bounds on θ by using the relation derived in [3].

²See for the original articles by Peccei and Quinn [4, 5]. An up-to-date review can be found in [6].

symmetry under consideration. The main problem excluding the axion is the fact that models can be created with almost arbitrarily low mass and couplings to ordinary particles. Thus far, it is undecided whether nature has a PQ symmetry or not.

The axion aspect of the PQ mechanism is investigated exhaustively, but there might be more to it. Peccei and Quinn showed that QCD with a PQ symmetry does not have explicit CP -violation and, besides that, the ground state or vacuum is CP symmetric, which means that there is no spontaneous CP -violation either. But the PQ mechanism does not exclude metastable states, which might violate CP or are in any other way different from the true ground state.

A metastable state is a local minimum in the quantum effective potential for the fields (see also fig. 1.1). It is theoretically possible for the vacuum to ‘hang’ in a metastable state for some finite amount of time. Particles, which are nothing more than field excitations around the vacuum field, can have different masses and couplings when the vacuum is trapped into a metastable state. In practice, a quantum field might get trapped into a metastable state when it cools down after it has been heated to some high temperature. Such a heating happens e.g. in a high-energy heavy-ion collision in which quark-gluon plasma’s are created with temperatures $T \sim 200\text{MeV}$ ($\sim 2.3 \cdot 10^{12}\text{K}$). These collisions are performed at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory and in the future at the Large Hadron Collider (LHC) at CERN.

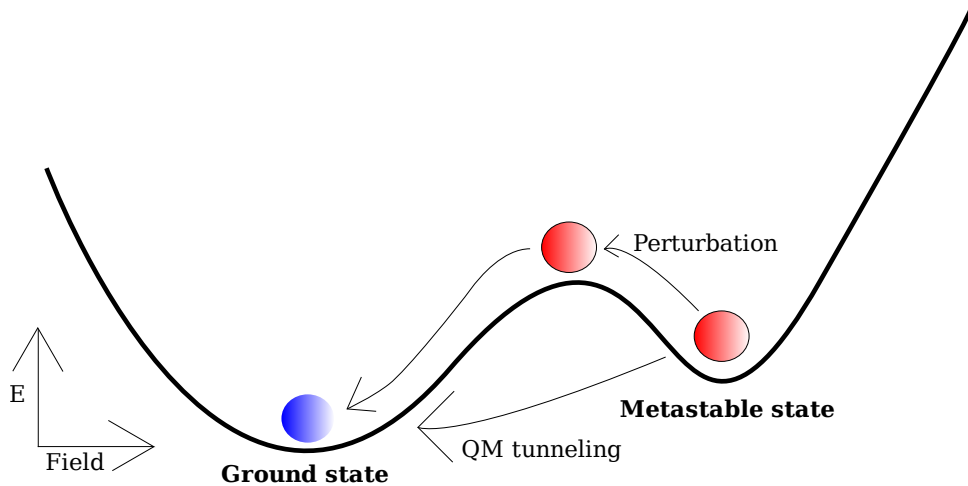


Figure 1.1: Cartoon displaying the principle of a metastable state. Transition to the true ground state can be caused by a perturbation, which tips the field over the potential barrier, or by quantum mechanical tunneling through the barrier.

It is interesting to know whether the PQ mechanism causes a metastable state to appear. It could thus give an experimental signature of a PQ symmetry in a heavy-ion collision. And apart from that, it is from the viewpoint of the baryon asymmetry important to investigate all possible sources of CP violation, as it might be that some of those metastable states violate CP symmetry.

It is this possibility of a metastable vacuum state in the strong interaction due to a PQ mechanism that is the subject of this thesis. The main question will be whether the PQ mechanism allows the possibility of a (CP violating) metastable state. This will be investigated by using a low-energy effective model approximation to QCD equipped with a PQ symmetry.

1.1 Outline

In chapter 2 an introduction to quantum chromodynamics will be given. Then in chapter 3 a special type of field configuration, called instanton, will be explained. These instantons are at the heart of the possibility of QCD to violate CP invariance. In chapter 4 the consequences of instantons will be discussed, of which the most important is the strong CP -problem. The proposed solution to the strong

CP -problem by Peccei and Quinn will be explained in chapter 5. The low-energy effective model that is used to approximate QCD is treated in chapter 6. Then in chapter 7 this effective model is equipped with a Peccei-Quinn symmetry and the effective potential is plotted. These results will be discussed in chapter 8 and, finally, conclusions will be drawn in chapter 9.

Chapter 2

Introduction to QCD

Quantum Chromodynamics (QCD) is the part of the standard model that describes the strong interaction. The strong interaction is (at low energies) the strongest of all fundamental forces. It is responsible for binding *quarks* into *mesons* and *hadrons*, of which the most well known examples are the *proton* and the *neutron*.

QCD is a *quantum field theory*, in which a number of *fermion fields*, N_f , interact through eight *gauge fields*. The quanta of the fermion fields are called quarks, of which six different flavors have yet been found ($N_f = 6$). The quanta associated with the gauge fields are called *gluons*. The gauge fields arise by imposing a local $SU(3)$ symmetry on the fermion fields, such a theory is called a *non-Abelian gauge theory*.

2.1 Free fermions

Basic familiarity of the reader with quantum field theory will be assumed. Only a quick reminder of the required knowledge will be given here. For a very good introduction to quantum field theory I advise the book by Peskin & Schroeder [7].

The classical Lagrangian density describing a free fermion (spin $\frac{1}{2}$) field with mass m is given, in e.g. section 3.2 of [7], by

$$\mathcal{L} = \bar{\Psi}(i\not{\partial} - m)\Psi \equiv \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi, \quad (2.1)$$

in which Ψ is a 4-component spinor and $\bar{\Psi} \equiv \Psi^\dagger\gamma^0$. The γ^μ 's are the 4×4 Dirac matrices, obeying the anticommutation relations $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. This Lagrangian density fully specifies the classical field theory for fermions.

Still this theory needs to be quantized in order to describe nature well. The quanta of this field will then be spin $\frac{1}{2}$ particles, with mass m , and so should obey *Fermi-Dirac statistics*. The path integral formalism provides a direct way of calculating quantum mechanical amplitudes from the classical Lagrangian density.

The path integral formalism, see e.g. chapter 9 in [7], says that a quantum mechanical amplitude can be expressed in terms of an integral over all possible field configurations, e.g. the *two-point correlation function* can be written as

$$\langle 0|T\hat{\Psi}(x_1)\hat{\bar{\Psi}}(x_2)|0\rangle = \frac{\lim_{T \rightarrow \infty(1-i\epsilon)} \int [d\bar{\Psi}][d\Psi] \Psi(x_1)\bar{\Psi}(x_2) \exp \left[i \int_{-T}^T d^4x \mathcal{L} \right]}{\lim_{T \rightarrow \infty(1-i\epsilon)} \int [d\bar{\Psi}][d\Psi] \exp \left[i \int_{-T}^T d^4x \mathcal{L} \right]}. \quad (2.2)$$

In the expression for the two-point correlation function, eq. (2.2), the left-hand side is the quantum mechanical expectation value for the time ordered product of the two field *operators* $\hat{\Psi}$ and $\hat{\bar{\Psi}}$ in the

ground state. Whereas in the right-hand side Ψ and $\bar{\Psi}$ are not operators, but anti-commuting complex numbers (complex Grassmann variables). The integral $\int [d\bar{\Psi}][d\Psi]$ is a path integral, i.e. an integration over all possible field configurations.

For the free fermion Lagrangian density, eq. (2.1), the path integral in eq. (2.2) can be calculated exactly. This gives a direct expression for the two-point correlation function,

$$\langle 0 | T \hat{\Psi}(x_1) \hat{\bar{\Psi}}(x_2) | 0 \rangle = S_F(x_1 - x_2) = \int \frac{d^4 k}{(2\pi)^4} \frac{i e^{-ik \cdot (x_1 - x_2)}}{\not{k} - m + i\epsilon}, \quad (2.3)$$

which is called the Feynman propagator. More complicated correlation functions can be calculated by inserting more operators in the left hand side of eq. (2.2) and correspondingly more Grassmann variables in the right hand side.

2.2 Non-Abelian gauge theory

A gauge theory is defined as a theory which has a Lagrangian that is invariant under some *local* and *continuous* set of transformations on the fields (so called *gauge* transformations). If one demands a Lagrangian, e.g. one containing some fermion fields, to be invariant under a set of local transformations, then one has to add extra fields, called *gauge fields*. This can be made clear with an example.

Consider a theory with three different fermion fields, all having mass m . The Lagrangian density can be written down in terms of a column vector containing these fields,

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi, \quad \text{where} \quad \Psi = \begin{pmatrix} \Psi_1(x) \\ \Psi_2(x) \\ \Psi_3(x) \end{pmatrix}. \quad (2.4)$$

This theory is symmetric under *global* unitary transformations on the column vector Ψ . The word global here means the same transformation on every point in space-time. Such a transformation can be written as

$$\Psi \rightarrow U\Psi, \quad \bar{\Psi} \rightarrow \bar{\Psi}U^\dagger, \quad (2.5)$$

where U is a 3×3 unitary matrix.

It can be easily seen that the transformation law in eq. (2.5) leaves the Lagrangian in eq. (2.4) invariant. The matrix U commutes with the Dirac matrices, γ^μ , since they act in a different space and because U does not depend on x , it commutes with the derivative operator. It is a different situation, however, when the transformation matrix, U , *does* depend on x .

Quantum chromodynamics is a theory in which every quark field comes in three different *colors*. So instead of one *up-quark* there are three up-quarks, a *red*, *green* and a *blue* one. And, next to that, the theory is symmetric under local unitary transformations with determinant 1, which act on the vector of different quark colors in eq. (2.4).

The group of all these unitary matrices is called $SU(3)$. It is a non-Abelian Lie-group. The vector of the three quark colors transforms in the fundamental, or defining representation of this group. The elements of this representation can be parameterized by 8 real numbers, ξ_a , $a = 1, \dots, 8$. A local transformation then can be specified by 8 real fields, $\xi_a(x)$, through

$$\Psi \rightarrow U(x)\Psi, \quad \text{where} \quad U(x) = \exp(i\xi_a(x)T^a), \quad (2.6)$$

in which the T^a 's are the *generators* of the group.

The free theory example, in eq. (2.1), is not symmetric under local $SU(3)$ transformations,

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \rightarrow \mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi + i\bar{\Psi}U^\dagger(x)(\gamma^\mu \partial_\mu U(x))\Psi. \quad (2.7)$$

The problem is the derivative term. If the derivative working on the field vector Ψ would transform in the same way as the field vector Ψ itself, then the Lagrangian would be invariant. The derivative with this property is called the *covariant derivative* and denoted by D_μ .

The covariant derivative can explicitly be constructed in the following way,

$$D_\mu = \partial_\mu - igA_\mu^a(x)T^a, \quad (2.8)$$

where g is an arbitrary (coupling) constant. Eight new, vector valued fields ($A_\mu^a(x)$) are needed. Those fields should obey a specific transformation law under the $SU(3)$ transformations in order to let the covariant derivative transform in the right way. If the fields transform as

$$A_\mu^a T^a \rightarrow U(x) \left(A_\mu^a T^a + \frac{i}{g} \partial_\mu \right) U^\dagger(x), \quad (2.9)$$

where the derivative operator acts on $U^\dagger(x)$, then the covariant derivative acting on the field vector transforms as

$$\begin{aligned} D_\mu \Psi &\rightarrow \left[\partial_\mu - ig \left(U(x) A_\mu^a T^a U^\dagger(x) + U(x) (\partial_\mu U^\dagger(x)) \right) \right] U(x) \Psi \\ &= \partial_\mu (U(x) \Psi) + U(x) (\partial_\mu U^\dagger(x)) U(x) \Psi - ig U(x) A_\mu^a T^a \Psi \\ &= U(x) (\partial_\mu - ig A_\mu^a T^a) \Psi + \left[\partial_\mu U(x) + U(x) (\partial_\mu U^\dagger(x)) U(x) \right] \Psi. \end{aligned} \quad (2.10)$$

The covariant derivative thus transforms in the right way, i.e.

$$D_\mu \Psi \rightarrow U(x) D_\mu \Psi, \quad (2.11)$$

if the quantity in the square brackets in eq. (2.10) vanishes. The fact that this happens can readily be seen by realizing that U is unitary so $UU^\dagger = 1$ and consequently

$$\partial_\mu (UU^\dagger) = 0 \quad \Rightarrow \quad (\partial_\mu U) U^\dagger = -U (\partial_\mu U^\dagger) \quad \Rightarrow \quad \partial_\mu U = -U (\partial_\mu U^\dagger) U. \quad (2.12)$$

Accordingly the covariant derivative, defined by eq. (2.8), acting on the field vector *does* transform in the right way under local $SU(3)$ transformations, but only if the $A_\mu^a(x)$ fields are imposed to transform according to eq. (2.9).

If the ‘normal’ derivative in the Lagrangian for the three free fermions, eq. (2.4), is now replaced by the covariant derivative, i.e. $\partial_\mu \rightarrow D_\mu$, the Lagrangian will be invariant under the local $SU(3)$ transformations from eq. (2.6). But keep in mind that we had to add eight new, vector valued fields A_μ^a .

The common understanding on how to create a Lagrangian is that it should contain all terms that are invariant under the demanded symmetry. Next to that, it should not contain combinations of fields which have a dimension higher than four, in order to keep the theory renormalizable. Concerning terms which contain fermion fields, all possible combinations are just the two we already have: the covariant derivative term and the mass term.

Still, combinations with just the new A_μ^a fields can also be invariant under the local $SU(3)$ symmetry. Apart from that, they should, of course, also be *Lorentz invariant*. To construct a gauge invariant quantity, let us first look at transformation properties of the commutator of covariant derivatives operating on a field vector,

$$[D_\mu, D_\nu] \Psi \rightarrow U(x) [D_\mu, D_\nu] \Psi. \quad (2.13)$$

This follows directly from the transformation law of the covariant derivative.

With use of the definition, eq. (2.8), it turns out that the commutator of covariant derivatives is not a derivative but just a matrix,

$$[D_\mu, D_\nu] = -igF_{\mu\nu}, \quad (2.14)$$

where

$$F_{\mu\nu} = F_{\mu\nu}^a T^a = \partial_\mu A_\nu^a T^a - \partial_\nu A_\mu^a T^a - ig[A_\mu^a T^a, A_\nu^b T^b]. \quad (2.15)$$

From the transformation rule of the commutator of covariant derivatives, eq. (2.13), and the transformation rule for the field vector, eq. (2.6), the transformation law for the $F_{\mu\nu}$ matrix is easily determined to be

$$F_{\mu\nu} \rightarrow U(x)F_{\mu\nu}U^\dagger(x). \quad (2.16)$$

A gauge invariant quantity would thus be the trace of one or the product of multiple $F_{\mu\nu}$ matrices. To make a Lorentz invariant quantity the space-time indices should still be contracted. This can be done in two different ways,

$$\text{Tr}[F^{\mu\nu}F_{\mu\nu}] \quad \text{and} \quad \text{Tr}[\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}], \quad (2.17)$$

resulting in a scalar and a pseudoscalar respectively.

The pseudoscalar term has a special property. It can be written as a full derivative, i.e. the divergence of a vector field. It is not difficult to check that the vector field

$$K^\mu = 4\epsilon^{\mu\nu\rho\sigma}\text{Tr}[A_\nu\partial_\rho A_\sigma - \frac{2ig}{3}A_\nu A_\rho A_\sigma] \quad (2.18)$$

has a divergence

$$\partial_\mu K^\mu = \text{Tr}[\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}]. \quad (2.19)$$

This is important, because in Euclidean space, a volume integral over a divergence can be written as a flux through a surface integral using Gauss's theorem. So, if the Lagrangian density would contain a term as in eq. (2.19), then in Wick rotated space (see appendix A) it can be rewritten to a surface integral:

$$\int d^4x \text{Tr}[\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}] = \int_{S^3} d^4x \partial_\mu K^\mu = \int_{S^3} dA_\mu K^\mu. \quad (2.20)$$

The last integral is over the three dimensional sphere representing Euclidean infinity. If now $K^\mu(x) \rightarrow 0$ fast enough as $x \rightarrow \infty$, then this integral vanishes and the term could be left out of the Lagrangian.

The assumption that the potentials, A_μ , vanish at infinity seems very reasonable. For example, a term

$$- \int d^4x \text{Tr}[F^{\mu\nu}F_{\mu\nu}] \quad (2.21)$$

in the Lagrangian should be finite in order to contribute to the path integral (at least in the semiclassical approximation, see e.g. section 7.3.2 in [8]). To see that it has to be finite, consider the path integral expression for the two-point function, eq. (2.2). If we look at, e.g. the numerator, we see that the $F_{\mu\nu}F^{\mu\nu}$ term will contribute a factor

$$\lim_{T \rightarrow \infty(1-i\epsilon)} \exp \left[-i \int_{-T}^T d^4x \text{Tr}[F_{\mu\nu}F^{\mu\nu}] \right], \quad (2.22)$$

which can, by a Wick rotation (see appendix A), be written as

$$\exp \left[- \int d^4x \text{Tr}[F_{\mu\nu}F^{\mu\nu}] \right], \quad (2.23)$$

in which the integration is now over Euclidean \mathbb{R}^4 -space. All gauge field configurations for which the integral is infinite will contribute a factor zero to the path integral. The path integral, therefore, only needs to include those fields for which this integral is finite. These are so called finite Euclidean action field configurations.

The finiteness of the action means, for the fields, that they should go to zero, $F_{\mu\nu} \rightarrow 0$, as $x \rightarrow \infty$. This boundary condition on the fields is certainly satisfied when the potentials go to zero, $A_\mu \rightarrow 0$, when

$x \rightarrow 0$. All gauge transformations of these potentials (eq. (2.9)) are, of course, also fine, because they will not change the value of $\text{Tr}[F^{\mu\nu}F_{\mu\nu}]$, since it is gauge invariant.

It is therefore tempting to say that a term like eq. (2.20) will always yield zero and can therefore be left out of the Lagrangian. This is exactly what has been done, until the discovery, by Belavin, Polyakov, Schwartz and Tyupkin (1975) [9], of nontrivial $A_\mu(x)$ field configurations which do satisfy the boundary conditions, but also give a contribution to the surface integral in eq. (2.20). These field configurations, called instantons, and their consequences will be discussed in the next chapters. In the remainder of this chapter these instantons will be ignored.

The most general, renormalizable, $SU(3)$ gauge invariant Lagrangian density for a three fermion vector (eq. (2.4)) can now be given as

$$\mathcal{L} = \bar{\Psi}(i\not{D} - m)\Psi - \frac{1}{4}\text{Tr}[F^{\mu\nu}F_{\mu\nu}]. \quad (2.24)$$

A constant in front of the last term is not needed. It can always be absorbed into the coupling constant g , in the covariant derivative (eq. (2.8)), by a rescaling of the $A_\mu(x)$ fields. The specific choice $\frac{1}{4}$ is for convenience in calculations. The minus sign ensures that the field configuration with the lowest energy (at least classically) is one with zero strength. This is, of course, what one should want, because a non-zero vector field in the ground state will break Lorentz invariance. More fermion vectors can be added by repetitively inserting the first term for the extra fermions.

To calculate quantum mechanical amplitudes from this Lagrangian, again a path integral can be used. This can, however, not been done exactly. A way of approximating the amplitude is to use perturbation theory. In perturbation theory one expands the amplitude in a power series expansion of the coupling constant. This perturbative approach of calculating amplitudes can be summarized in a set of rules, called Feynman rules. Familiarity with perturbation theory is assumed and it will therefore not be explained here. A summary of the Feynman rules for non-Abelian gauge theories can be found in any good book on quantum field theory, e.g. appendix A.1 in [7] and appendix B in [10].

Another important technique one also has to use is renormalization. To make sense out of a perturbation series, the coupling constants in the Lagrangian, i.e. the *bare coupling constants*, (and also the scaling of the fields) have to be shifted by infinite amounts in order to explain any physical properties or coupling constants of the particles. This process introduces a renormalization scale, e.g. μ , at which the physical constants that are predicted by the theory are assigned a numerical value.

2.3 Running of the coupling constant

The idea behind an expansion in the coupling parameter (e.g. g) is that higher order diagrams contribute less and less. If the contribution of the diagrams drops off fast enough, then only a very small amount of diagrams need to be calculated. This is essential, as the complexity and the size of the calculations blow up as the order of the diagrams increase.

The order of a diagram is measured as the power of the coupling constant that is in it. So for a given expansion, the smaller the coupling the less important are the higher order diagrams and the faster it converges.

A direct expansion in the bare coupling constant, g , from eq. (2.8), is not possible as its size is infinite. A better expansion is in the effective (or renormalized) coupling constant, i.e. $\alpha_{s,\text{eff}} = g_{s,\text{eff}}^2/4\pi^2$. This is called renormalized perturbation theory.

The renormalized coupling, however, depends on the renormalization scale μ . The renormalized coupling could also have been fixed at a different scale μ' . At this new scale, μ' , exactly one combination of numerical assignments results in the same theory as the previous assignment at scale μ . A theory (i.e. a set of coupling constants) at one scale can thus be related to a theory at another scale. The equations governing this behavior are called the renormalization group equations. The specific equation describing the behavior of $\alpha_s(\mu)$ (dropping the $_{\text{eff}}$ label for clarity) is (see e.g. section 12.2 in [7])

$$\mu \frac{d\alpha_s(\mu)}{d\mu} = 2\beta(\alpha_s(\mu)), \quad (2.25)$$

in which $\beta(\alpha_s)$ is called the β -function. This function can be calculated in perturbation theory as a series expansion in α_s ,

$$2\beta(\alpha_s) = -\frac{\beta_0}{2\pi}\alpha_s^2 - \frac{\beta_1}{4\pi^2}\alpha_s^3 - \frac{\beta_2}{64\pi^3}\alpha_s^4 - \dots \quad (2.26)$$

The first three coefficients are, for QCD, known to be (see e.g. [11])

$$\begin{aligned} \beta_0 &= 11 - \frac{2}{3}n_f, \\ \beta_1 &= 51 - \frac{19}{3}n_f, \\ \beta_2 &= 2857 - \frac{5033}{9}n_f + \frac{325}{27}n_f^2. \end{aligned} \quad (2.27)$$

This results in a differential equation for $\alpha_s(\mu)$. One which is, of course, only valid for small α_s , because of the expansion of the β -function.

A numerical solution is easily obtained, if one specifies α_s at some scale μ_0 . As a reference point, the Z-boson mass is usually taken. According to the most recent experiments (Particle Data Group [11]), $\alpha_s(M_z) = 0.1176$. In figure 2.1 a solution is plotted with this boundary condition.

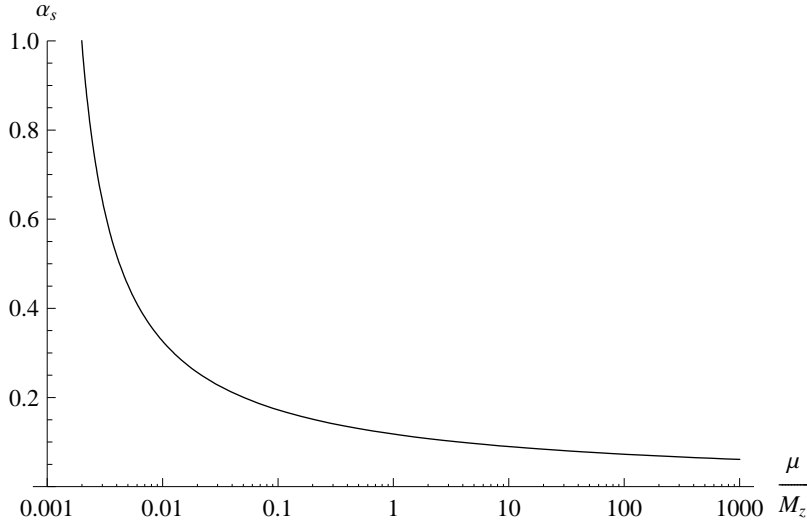


Figure 2.1: The dependence of α_s on the renormalization scale μ .

Two important things should be noted on the running of α_s . The first thing is that, if $\mu \rightarrow 0$, then $\alpha_s \rightarrow 0$, this is called asymptotic freedom. The second is that it grows very fast as the energy decreases. Below energies of 1GeV, we should doubt whether the perturbation expansion of the β -function still holds.

If one wants to compute a measurable quantity, e.g. a cross section, one best uses the value of $\alpha_s(\mu)$ at a typical momentum scale involved in the process. Otherwise the higher order corrections will automatically bring the coupling to the appropriate effective strength for that problem. The quarks will ‘feel’ the effective coupling at the scale at which they exchange momentum. It is clear that a perturbation expansion in α_s will fail hopelessly in situations where the momentum transfer is around, or less than 1GeV.

Perturbative QCD is only useful in a high energy limit ($> 1\text{GeV}$). Bound states of quarks such as pions, neutrons and protons cannot be calculated in this way. Neither can the QCD vacuum state be calculated perturbatively.

2.4 Symmetries in QCD

Apart from the local $SU(3)$ symmetry, QCD also possesses a number of global symmetries. These can be divided into discrete and continuous symmetries. Discrete symmetries consist of a finite or countably infinite set of transformations. Continuous symmetries, on the other hand, form an uncountable set of transformations. These elements, then, depend on one or more variables.

2.4.1 Discrete symmetries

Examples of discrete symmetries are parity inversion (P), time inversion (T) and charge conjugation (C). If a point in space time is given by $x = (t, \vec{x})$, then P transforms $x \rightarrow (t, -\vec{x})$ and T transforms $x \rightarrow (-t, \vec{x})$. From this it is immediately clear how a vector transforms, but it can also be derived how a spinor (fermion field) transforms. Charge conjugation is the operation that replaces particles by anti particles. The derivation of how different fermion-field bilinears transform under these transformations can be found in, e.g. section 3.6 in [7], and are summarized in table 2.1. In the table the shorthand $(-1)^\mu \equiv 1$ for $\mu = 0$ and $(-1)^\mu \equiv -1$ for $\mu = 1, 2, 3$ is used.

Table 2.1: C , P and T transformations

	$\bar{\Psi}\Psi$	$i\bar{\Psi}\gamma^5\Psi$	$\bar{\Psi}\gamma^\mu\Psi$	$\bar{\Psi}\gamma^\mu\gamma^5\Psi$	$\bar{\Psi}\sigma^{\mu\nu}\Psi$	∂_μ
P	+1	-1	$(-1)^\mu$	$-(-1)^\mu$	$(-1)^\mu(-1)^\nu$	$(-1)^\mu$
T	+1	-1	$(-1)^\mu$	$(-1)^\mu$	$-(-1)^\mu(-1)^\nu$	$-(-1)^\mu$
C	+1	+1	-1	+1	-1	+1

The C and P transformation properties of the $F_{\mu\nu}$ field are easily determined; since it has two indices it transforms as two vectors,

$$\begin{aligned} F_{\mu\nu} &\xrightarrow{P} (-1)^\mu(-1)^\nu F_{\mu\nu} \\ F_{\mu\nu} &\xrightarrow{T} -(-1)^\mu(-1)^\nu F_{\mu\nu} = (-1)^\mu(-1)^\nu F_{\mu\nu}. \end{aligned} \quad (2.28)$$

From this, it follows that PT leaves $F_{\mu\nu}$ invariant. The CPT theorem (see [12]) says that CPT is always a symmetry of a Lorentz invariant local field theory. Combining the last two statements, this shows that also C leaves $F_{\mu\nu}$ invariant.

The two possible terms in the QCD Lagrangian made out of $F_{\mu\nu}$ only transform as

$$\begin{aligned} F_{\mu\nu}F^{\mu\nu} &\xrightarrow{P,T} (1)^\mu(-1)_\mu(-1)^\nu(-1)_\nu F_{\mu\nu}F^{\mu\nu} = F_{\mu\nu}F^{\mu\nu} \\ \epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta} &\xrightarrow{P,T} \epsilon^{\mu\nu\alpha\beta}(-1)^\mu(-1)_\nu(-1)^\alpha(-1)_\beta F_{\mu\nu}F_{\alpha\beta} \\ &= -\epsilon^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}. \end{aligned} \quad (2.29)$$

Or equivalently, the first term is a scalar, whereas the second is a pseudoscalar. Under C , both are invariant and so the first term is CP even and the second CP odd.

We see that the standard QCD Lagrangian, eq. (2.24), is invariant under C , P and T separately (and therefore all combinations). It will be shown in the next two chapters how special field configurations can modify the ‘standard’ Lagrangian and can cause CP violation.

2.4.2 Continuous Symmetries

The full QCD Lagrangian, with six quark flavors: u, d, s, c, t, b , can be split into a light and heavy quark part. The light quark fields can be written as a vector in flavor space,

$$q(x) = \begin{pmatrix} u(x) \\ d(x) \\ s(x) \end{pmatrix}, \quad (2.30)$$

in which every quark field is again a vector of three different colors. The Lagrangian density can then be written as

$$\mathcal{L} = \bar{q}i\not{D} \otimes \mathbb{1}_3 q - \bar{q}Mq - \frac{1}{4}\text{Tr}[F^{\mu\nu}F_{\mu\nu}] + \text{heavy quarks}, \quad (2.31)$$

in which $\mathbb{1}_3$ is the 3×3 unit matrix (in flavor space) and M the mass matrix (also in flavor space), given by

$$M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}. \quad (2.32)$$

In flavor space a set of continuous transformation can be defined. Let us define a vector $SU(3)$ flavor transformation by

$$SU(3)_V : \quad q \rightarrow e^{i\alpha_a T^a} q, \quad (2.33)$$

where the α_a 's are real numbers and the T^a 's the generators of $SU(3)$ in the defining representation. Also transformations which treat the chiral left- and right-handed components of a spinor differently can be defined. This can be done with the use of the right/left projection operators,

$$P_{R,L} = \frac{1 \pm \gamma_5}{2}, \quad (2.34)$$

which project out the chiral left- and right-handed components of a spinor. A transformation on the right-handed components can then be written as

$$SU(3)_R : \quad q \rightarrow e^{i\alpha_a T^a P_R} q \quad (2.35)$$

and a transformation on the left-handed components as

$$SU(3)_L : \quad q \rightarrow e^{i\alpha_a T^a P_L} q. \quad (2.36)$$

Next to those, also a set of $U(1)$ transformation can be defined. For example a $U(1)$ transformation on a specific flavor, e.g. the up quark, by

$$U(1)_u : \quad u \rightarrow e^{i\delta} u, \quad (2.37)$$

the $U(1)$ ‘baryon-number’ transformation as an equal $U(1)$ transformation on every flavor, i.e.

$$U(1)_V : \quad q \rightarrow e^{i\delta} \otimes \mathbb{1}_3 q \quad (2.38)$$

and an axial ‘baryon-number’ $U(1)$ transformation as

$$U(1)_A : \quad q \rightarrow e^{i\delta\gamma_5} \otimes \mathbb{1}_3 q. \quad (2.39)$$

The parameter δ is again a real number and $\mathbb{1}_3$ is the unit matrix in flavor space.

From the Lagrangian density, eq. (2.31), it is clear that in the case that

$$m_u, m_d, m_s \text{ arbitrary} \Rightarrow U(1)_u \otimes U(1)_d \otimes U(1)_s \quad (2.40)$$

is the global symmetry group of the light quark Lagrangian. In the case that the light quark masses are degenerate, i.e.

$$m_u = m_d = m_s \Rightarrow SU(3)_V \otimes U(1)_B \quad (2.41)$$

is the global symmetry group. When all light quark masses are zero, i.e.

$$m_u = m_d = m_s = 0 \Rightarrow SU(3)_R \otimes SU(3)_L \otimes U(1)_B \otimes U(1)_A \quad (2.42)$$

transformations leave the Lagrangian invariant.

The last two cases are not purely academic. Since the light quarks are so light, these last two symmetries can be seen as approximate symmetries of the theory.

Chapter 3

Instanton solutions

This chapter will lean heavily on Coleman's explanation (see section 7.3 in [8]) of instantons.

In the previous chapter a specific assumption was made about the behavior of the gauge fields when they approach Euclidean infinity. It was assumed that the vector potentials $A_\mu^a(x) \rightarrow 0$ when $x \rightarrow \infty$. This assumption, however, is too strict.

The reason for looking at those field configurations was the requirement to have a finite Euclidean action. This Euclidean action can only be finite when the fields, $F_{\mu\nu}(x)$, go to zero fast enough as they approach Euclidean infinity, $x \rightarrow \infty$. This in turn was taken to mean that the potentials should go to zero fast enough. It turns out that there are also non-trivial potentials (i.e. not obtainable by a gauge transformation from a configuration with $A_\mu^a = 0$ at infinity), which do result in a finite Euclidean action.

3.1 Homotopy classes

To analyze those field configurations, an introduction to the mathematical concept of a homotopy class will be needed. This concept allows the distinction of functions, from one topological space to another, into different classes. Every function belongs to exactly one class and so can be classified accordingly.

3.1.1 Definition of a homotopy class

A homotopy class consists of all functions that are connected by a homotopy. A homotopy is, in words, a continuous transformation from one to the other. More rigorously[13], if f and g are two functions from a topological space X to another, Y , then a homotopy between them is a *continuous* map $G : X \otimes [0, 1] \rightarrow Y$, such that $G(x, 0) = f(x)$ and $G(x, 1) = g(x)$ in which $x \in X$.

3.1.2 Example: $S^1 \rightarrow S^1$

The simplest example (see e.g. section 16.1 in [10]) of a set of functions which can be divided into different homotopy classes is the set of functions from a circle to a circle, i.e. $S^1 \rightarrow S^1$; let us call the complete set $F_{S^1 \rightarrow S^1}$. Let the first circle be parameterized by $\Theta \in \mathbb{R} \bmod 2\pi$ and let the second circle be the set of unimodular complex numbers (which are topologically equivalent). Then a subset of the functions from one to the other is

$$f_a(\Theta) = \exp[i(\nu\Theta + a)] \in F_{S^1 \rightarrow S^1}, \quad (3.1)$$

for $\nu \in \mathbb{Z}$ and $a \in \mathbb{R}$. Those functions form a homotopy class for different values of a and fixed ν . This is because a homotopy $G(\Theta, t)$ can be constructed between e.g. f_{a_0} and f_{a_1} :

$$G(\Theta, t) = \exp[i(\nu\Theta + (1-t)a_0 + ta_1)]. \quad (3.2)$$

This is a homotopy because it is continuous and

$$\begin{aligned} G(\Theta, 0) &= \exp[i(\nu\Theta + a_0)] = f_{a_0}, \\ G(\Theta, 1) &= \exp[i(\nu\Theta + a_1)] = f_{a_1}. \end{aligned} \quad (3.3)$$

Every two functions with different ν are in different homotopy classes. To prove that this is a complete set of prototype functions for all classes, we should prove that *every* function from $S^1 \rightarrow S^1$ is homotopic to one of the prototype functions in eq. (3.1). This will not be done here, but it turns out that the set is complete. Thus every function from $S^1 \rightarrow S^1$ can be put in a homotopy class labeled by $\nu \in \mathbb{Z}$, which is called the *winding number*.

One more important property can be made clear by this simple example. It is the fact that the winding number for an arbitrary function, $g \in F_{S^1 \rightarrow S^1}$, can be expressed directly, with the use of an integral, in terms of g . This can be done through (see e.g. section 7.3.2 in [8])

$$\nu = \frac{-i}{2\pi} \int_0^{2\pi} d\Theta \frac{1}{g(\Theta)} \frac{dg(\Theta)}{d\Theta}. \quad (3.4)$$

3.1.3 Example: $S^3 \rightarrow SU(n)$

For the analysis of the finite action gauge fields later on, we will need to consider functions from S^3 to elements of the defining representation of $SU(3)$. I will just state the important facts and for proofs I will refer the reader to the literature, e.g. section 7.3.2 in [8].

It turns out that functions from $S^3 \rightarrow SU(n)$ can be classified in homotopy classes with just one integer label, the winding number ν . This winding number can be expressed as

$$\nu = -\frac{1}{24\pi^2} \int d\Theta_1 d\Theta_2 d\Theta_3 \epsilon^{ijk} \text{Tr} [g \partial_i g^{-1} g \partial_j g^{-1} g \partial_k g^{-1}], \quad (3.5)$$

in which Θ_1, Θ_2 and Θ_3 parameterize S^3 , the derivatives ∂_i work in this S^3 -space ($i = \Theta_1, \Theta_2, \Theta_3$) and g is a function from S^3 to the representation matrices of $SU(n)$.

3.1.4 Composition of functions

Another important property of the winding number is that a composition of two functions, i.e.

$$g(\Theta_1, \Theta_2, \Theta_3) \equiv g_1(\Theta_1, \Theta_2, \Theta_3) g_2(\Theta_1, \Theta_2, \Theta_3), \quad (3.6)$$

has as its winding number the sum of the two individual winding numbers:

$$\nu = \nu_1 + \nu_2. \quad (3.7)$$

From this it is also immediately clear that the inverse of a function has minus the winding number of that function.

3.2 Non-trivial finite Euclidean action gauge fields

The Euclidean action (see for the used conventions appendix A) for a pure gauge field theory is given by

$$S_E = \int d^4x \text{Tr} [F_{\mu\nu} F^{\mu\nu}] = \int_0^\infty dr r^3 \int d\Omega \text{Tr} [F_{\mu\nu}(r, \Omega) F^{\mu\nu}(r, \Omega)], \quad (3.8)$$

where r is the radial coordinate and Ω the set of angular coordinates in 4-dimensional polar coordinates. For this action to be finite, $F_{\mu\nu}$ must fall off faster than $1/r^2$ as r goes to infinity, so

$$F_{\mu\nu} = \mathcal{O}(1/r^3). \quad (3.9)$$

For the field to be of the form in eq. (3.9) it is not necessary that the potentials go to zero as $r \rightarrow \infty$. Every *pure gauge field*, i.e. a gauge transform of $A_\mu^a = 0$, also produces a zero field strength tensor $F_{\mu\nu}$. So the potentials can also be a pure gauge field at Euclidean infinity.

If a matrix valued vector potential is defined as

$$A_\mu \equiv A_\mu^a T^a, \quad (3.10)$$

then, to satisfy eq. (3.9), A_μ must be of the form

$$A_\mu = \frac{i}{g} G(\Omega) \partial_\mu G(\Omega)^{-1} + \mathcal{O}(1/r^2). \quad (3.11)$$

In this expression, $G(\Omega)$ is a continuous and differentiable function from the angular variables (S^3) to the elements of the gauge group in the defining representation.

Under a gauge transformation $U(x)$, the potential A_μ transforms according to eq. (2.9), i.e.

$$A_\mu \rightarrow U A_\mu U^{-1} + \frac{i}{g} U \partial_\mu U^{-1}. \quad (3.12)$$

So our finite Euclidean action potential, eq. (3.11), transforms as

$$G \rightarrow U G + \mathcal{O}(1/r^2). \quad (3.13)$$

If one could choose U equal to G^{-1} at infinity, then our finite action gauge field from eq. (3.11) would just be equivalent (up to a gauge transformation) to the trivial finite action field configuration $A_\mu = 0$ at infinity. In a path integral, the integration should be only over gauge inequivalent field configurations. So we could as well just integrate over field configurations in which $A_\mu \rightarrow 0$ as $x \rightarrow 0$. The fact is, however, that U *cannot* in general be chosen to equal G^{-1} at infinity.

The reason for this is that U should be a continuous function on the whole \mathbb{R}^4 . The gauge function, U , should be independent of the angular variables at the origin to be continuous. From there it is continuously deformable, by letting $r \rightarrow \infty$, to the configuration at infinity. Mathematically this means that $U(r, \Omega)$, for every fixed r , is in the same homotopy class as the trivial function $1(\Omega) \equiv 1$, which maps S^3 to the unit element of the group. Thus every gauge transformation $U(r, \Omega)$, at fixed r , belongs to the trivial homotopy class, i.e. $\nu = 0$.

It is only possible to choose U at infinity equal to G^{-1} if G happens to also be from the trivial homotopy class. However, G , being only a function of the radial variables, can be of any homotopy class. Therefore this is not possible in general.

The conclusion is that there are vector potentials, defined by eq. (3.11), which have a finite Euclidean action, but which *cannot* be obtained by a gauge transformation from a potential which has ‘trivial’ behavior at infinity, i.e. $A_\mu = \mathcal{O}(1/r^2)$.

These non-trivial potentials are characterized by a non-zero winding number, associated with the function G in eq. (3.11). This winding number can be expressed, by eq. (3.5), as an integral of $G(\Omega)$ over a 3-sphere. It is also possible to express the winding number, for those finite Euclidean action fields, directly in terms of the field.

To give a direct expression for the winding number, consider eq. (2.20), which is repeated here for convenience:

$$\int d^4x \text{Tr}[\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}] = \int_{S^3} dA_\mu K^\mu. \quad (3.14)$$

In the right hand side the integral is over the 3-sphere at Euclidean infinity. The measure dA_μ is the surface element of this sphere. The vector field K^μ , at infinity, can, with the use of eq. (2.18) and eq. (3.11), be expressed as

$$K^\mu = \frac{4}{3g^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} [(G \partial_\nu G^{-1}) (G \partial_\rho G^{-1}) (G \partial_\sigma G^{-1})]. \quad (3.15)$$

Substituting this equation into eq. (3.14) and comparing with expression (3.5) for the winding number it is not hard to see that

$$\int d^4x \text{Tr}[\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}] = -\frac{32\pi^2\nu}{g^2}. \quad (3.16)$$

With the use of eq. (3.16) every finite Euclidean action gauge field can be assigned an integer winding number ν . A zero winding number corresponding to the trivial potentials which behave (after an appropriate gauge transformation) as $\mathcal{O}(1/r^2)$ as $r \rightarrow \infty$. Non-zero winding number correspond to non-trivial potentials which cannot be gauged to zero at infinity.

3.3 Winding-number vacua

In this section we will look at the topological properties of the classical vacuum of non-Abelian gauge field theories. The classical vacuum of a field theory is the field which minimizes the classical potential energy. It will turn out that non-Abelian gauge field theories do not just have a single vacuum. Multiple topologically different vacuum states are possible.

Consider a pure non-Abelian gauge theory which is described by a Euclidean action

$$S_E = \int d^4x \text{Tr}[F_{\mu\nu} F^{\mu\nu}], \quad (3.17)$$

in which the field strength tensor is matrix valued and can be expressed in terms of matrix valued vector potentials as usual (eq. (2.15)). The Hamiltonian, at a given time x_0 , for such a theory is given by

$$H = \int d^3x (F_{\mu\nu}(\mathbf{x}))^2. \quad (3.18)$$

This makes it clear that the field configuration with the lowest energy is just the constant field

$$F_{\mu\nu}(x) = 0 \quad \forall x \in \mathbb{R}^4. \quad (3.19)$$

All *pure gauge* potentials, i.e.

$$A_\mu(x) = U(x) \partial_\mu U^{-1}(x) \quad \forall x \in \mathbb{R}^4, \quad (3.20)$$

will produce a constant zero field. The functions $U(x)$ are continuous functions from \mathbb{R}^4 to the representation matrices of the gauge group.

Not all different functions in eq. (3.20) correspond to different vacua. Let us first pick a specific gauge, e.g. the temporal gauge:

$$A_0(x) = 0 \quad \forall x \in \mathbb{R}^4. \quad (3.21)$$

All time-independent pure gauge potentials automatically satisfy the gauge fixing condition, because

$$A_0(x) = U(\mathbf{x}) \partial_0 U^{-1}(\mathbf{x}) = 0. \quad (3.22)$$

So in this specific gauge, all potentials given by

$$A_i(x) = A_i(\mathbf{x}) = \frac{i}{g} U(\mathbf{x}) \partial_i U^{-1}(\mathbf{x}) \quad (3.23)$$

result in a zero field strength tensor, i.e. a vacuum field. In this equation $U(\mathbf{x})$ is now a function from \mathbb{R}^3 to the representation matrices of the gauge group. To classify those functions, a continuous deformation of U can be made, since this will not alter the class of U .

Consider now the transformation

$$U \rightarrow H_1 U, \quad (3.24)$$

where

$$H_\epsilon(r, \Omega) = \frac{1}{1+r} + \frac{r}{1+r} U^{-1}(\epsilon r, \Omega), \quad \epsilon \in [0, 1]. \quad (3.25)$$

This function $H_\epsilon(r, \Omega)$ has the property that 1) it is continuous on \mathbb{R}^3 , 2) it is continuously obtainable from the unit function (by letting $\epsilon \rightarrow 0$) and 3) $H_1 U(r, \Omega) \rightarrow \mathbb{1}$ as $r \rightarrow \infty$. So the transformation in eq. (3.24) does not alter the class structure of U . It does, however, transform every U to a function which is constant at Euclidean infinity ($r \rightarrow \infty$).

To classify the different U 's, one can thus also classify $H_1 U$, i.e. only functions which map infinity to the same element have to be considered. The three dimensional Euclidean space with infinity identified is topologically equivalent to S^3 . So $U(\mathbf{x})$ is, topologically, a mapping from $S^3 \rightarrow SU(n)$ and it can, therefore, be divided into homotopy classes as discussed in section 3.1.3. The vacuum configurations will be labeled by a roman winding number, e.g. n , in contrast with the finite Euclidean action gauge field configurations, which are labeled by a Greek winding number, e.g. ν .

As for the finite Euclidean action fields, a direct expression (similar to eq. (3.5)) for the winding number in terms of the potential is available for vacuum configurations, given in e.g. section 16.1 in [10]. Given a vacuum potential as in eq. (3.23), the winding number for a vacuum field configuration is given by

$$n = -\frac{1}{24\pi^2} \int d^3x \epsilon^{ijk} \text{Tr} [A_i A_j A_k]. \quad (3.26)$$

An example, given in section 16.2 in [10], of a $n = 1$ vacuum state, in $SU(2)$ gauge theory, is eq. (3.23) with

$$U(\mathbf{x}) = U_1(\mathbf{x}) \equiv \exp \left[\frac{i\pi \boldsymbol{\sigma} \cdot \mathbf{x}}{\sqrt{|\mathbf{x}|^2 + \lambda^2}} \right], \quad (3.27)$$

in which $\boldsymbol{\sigma}$ is the 3-vector of Pauli matrices. Note that this $U(\mathbf{x})$ indeed approaches a constant value if $|\mathbf{x}| \rightarrow \infty$, because

$$\lim_{|\mathbf{x}| \rightarrow \infty} U_1(\mathbf{x}) = e^{i\pi \boldsymbol{\sigma} \cdot \hat{\mathbf{x}}} \quad (3.28)$$

and using the identity

$$e^{-ia(\hat{n} \cdot \boldsymbol{\sigma})} = \cos a + i(\hat{n} \cdot \boldsymbol{\sigma}) \sin a, \quad (3.29)$$

it is clear that

$$\lim_{|\mathbf{x}| \rightarrow \infty} U_1(\mathbf{x}) = -1. \quad (3.30)$$

Higher winding-number vacuum states can, for example, be constructed by taking higher powers of this U :

$$U_n(\mathbf{x}) \equiv U_1(\mathbf{x})^n = \exp \left[\frac{i\pi \boldsymbol{\sigma} \cdot \mathbf{x}}{\sqrt{|\mathbf{x}|^2 + \lambda^2}} n \right]. \quad (3.31)$$

The functions in eq. (3.31) give a complete set of prototype functions; there is one function for every class. All different functions $U(\mathbf{x})$ in the *same* homotopy class are obtainable from each other by *small gauge transformations*, which are defined to be continuously obtainable from the identity transformation, $V = \mathbb{1}$.

A gauge transformation, V , will transform a vacuum configuration (eq. (3.23)) through

$$\frac{i}{g} U \partial_\mu U^{-1} \rightarrow \frac{i}{g} V (U \partial_\mu U^{-1}) V^{-1} + \frac{i}{g} V \partial_\mu V^{-1} = \frac{i}{g} V U \partial_\mu (V U)^{-1}, \quad (3.32)$$

i.e.

$$U \rightarrow VU. \quad (3.33)$$

A small gauge transformation is by definition continuously obtainable from the identity gauge transformation, $V = 1$. Therefore, all small gauge transformations belong to the trivial homotopy class. So a small gauge transformation, V_{small} , will transform a vacuum configuration, with some winding number n , e.g. U_n , to another vacuum configuration $V_{\text{small}}U_n$, which has the same winding number n .

Large gauge transformations, on the contrary, are defined to be *not* continuously obtainable from the identity transformation. They are not in the trivial homotopy class and will, therefore, transform a vacuum configuration with winding number n to one with a different winding number $m \neq n$.

The conclusion is that the vacuum structure of a non-Abelian gauge theory is topologically non-trivial. Infinitely many different vacua exist, which *cannot* be transformed into each other by small gauge transformations. They can be labeled by the (gauge variant) winding number and are called *winding-number vacua*.

3.4 Instantons

Instantons are a special case of non-trivial finite Euclidean action field configurations; they do not only have a finite action, they are also a *local minimum* of the Euclidean action. Being a local minimum of the Euclidean action they are a solution to the classical Euclidean equations of motion.

One might question why it is important to investigate the minima of the (Euclidean) action when one wants to do a *quantum mechanical* computation. The fact is that, in a semi classical (i.e. small \hbar) or small coupling constant approximation, the path integral is dominated by those minima of the action, which are the solutions of the classical equations of motion.

The implicit statement that a semi classical expansion is equivalent to a small coupling expansion will be made clear by an example. Consider a non-Abelian gauge field theory with a Euclidean action

$$S_E[A_\mu] = \int d^4x \text{Tr}[F_{\mu\nu}F^{\mu\nu}]. \quad (3.34)$$

The partition function is given in terms of a path integral by

$$Z = \int [dA_\mu] e^{-S_E[A_\mu]/\hbar}, \quad (3.35)$$

in which the factor of \hbar is restored. This allows a semi classical approximation, i.e. a small \hbar approximation. In the limit $\hbar \rightarrow 0$, the classical result should be obtained, so terms of order $\mathcal{O}(\hbar^0)$ represent the classical result. Higher order terms are quantum mechanical corrections, where, in the limit of small \hbar , the higher the order the less important the term is.

The dependence of the action on the coupling constant, g , is hidden in the definition of the field $F_{\mu\nu}$ in eq. (2.15). But with a rescaling of the fields

$$A_\mu \rightarrow \frac{1}{g}A_\mu, \quad F_{\mu\nu} \rightarrow \frac{1}{g}F_{\mu\nu}, \quad (3.36)$$

the dependence is moved from within $F_{\mu\nu}$ to the same location as where \hbar is. The partition function is then

$$Z = \int \left[\frac{dA_\mu}{g} \right] e^{-S_E[A_\mu]/g^2\hbar}, \quad (3.37)$$

in which the Euclidean action is the same as in eq. (3.34), but now in terms of the new fields

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]. \quad (3.38)$$

From this expression for the partition it is clear that a small \hbar approximation corresponds to a small g approximation. The conclusion is that a semiclassical approximation, i.e. a small \hbar approximation, is the same as a small coupling constant approximation.

From the expression for the path integral, eq. (3.37), it is also clear that, in the limit $\hbar \rightarrow 0$ or $g \rightarrow 0$, the integral is dominated by the minimum/minima of the action, A_μ^c :

$$\left. \frac{\delta S_E[A_\mu]}{\delta A_\mu} \right|_{A_\mu=A_\mu^c} = 0. \quad (3.39)$$

This fact is used in the method of steepest descent/stationary phases, see e.g. [14] chap. 1.2. It expands the action around these (local) minima.

The method of steepest descent can be used on the path integral in eq. (3.37) [15] by first making the change of variables

$$A_\mu = A_\mu^c + g A'_\mu \quad (3.40)$$

and then expanding the action through

$$\begin{aligned} \frac{1}{g^2} S_E[A_\mu] &= \frac{1}{g^2} S_E[A_\mu^c] + \int dx dy \frac{1}{2!} \left. \frac{\delta^2 S_E[A_\mu]}{\delta A_\mu(x) \delta A_\mu(y)} \right|_{A_\mu^c} A'_\mu(x) A'_\mu(y) \\ &\quad + \sum_{k=3}^{\infty} \frac{g^{k-2}}{k!} \int dx_1 \cdots dx_k \left. \frac{\delta^k S_E[A_\mu]}{\delta A_\mu(x_1) \cdots \delta A_\mu(x_k)} \right|_{A_\mu^c} A'_\mu(x_1) \cdots A'_\mu(x_k). \end{aligned} \quad (3.41)$$

The integral then becomes (working again in units where $\hbar = 1$)

$$\begin{aligned} Z &= e^{-S[A_\mu^c]/g^2} \int [dA'_\mu] \exp \left[-\frac{1}{2!} \int dx dy \left. \frac{\delta^2 S_E[A_\mu]}{\delta A_\mu(x) \delta A_\mu(y)} \right|_{A_\mu^c} A'_\mu(x) A'_\mu(y) - R[A'_\mu] \right] \\ R[A'_\mu] &= \sum_{k=3}^{\infty} \frac{g^{k-2}}{k!} \int dx_1 \cdots dx_k \left. \frac{\delta^k S_E[A_\mu]}{\delta A_\mu(x_1) \cdots \delta A_\mu(x_k)} \right|_{A_\mu^c} A'_\mu(x_1) \cdots A'_\mu(x_k). \end{aligned} \quad (3.42)$$

The integrand can then be expanded in powers of g . At each order the integrand is a polynomial times a gaussian and can be integrated by the standard rules of functional integration. If the action contains more than one minimum, then a summation over all minima should be made. The result of the functional integration will be of the form

$$Z = \sum_{A_\mu^c} e^{-S[A_\mu^c]/g^2} \det \left[\left. \frac{\delta^2 S_E[A_\mu]}{\delta A_\mu^2} \right|_{A_\mu^c} \right] [1 + \mathcal{O}(g)]. \quad (3.43)$$

To make use of the method of steepest descent, all minima of the Euclidean action should be known. In a ‘simple’ theory the Euclidean action has just one global minimum. It will now be shown that in a non-Abelian gauge theory, the Euclidean action has multiple local minima.

The minima of the action will definitely belong to fields which have a finite action. We already know that those field configuration can be divided into different homotopy classes. In fact, it will turn out that every homotopy class has, at least one, field configuration in it which corresponds to a local minimum of the action.

The existence of multiple minima of the Euclidean action can be shown from the fact that, in Euclidean space,

$$\text{Tr} \int d^4x \left(F_{\mu\nu} \pm \tilde{F}_{\mu\nu} \right)^2 \geq 0, \quad (3.44)$$

where

$$\tilde{F}_{\mu\nu} \equiv \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma} \quad (3.45)$$

is the dual of $F_{\mu\nu}$. The integrand in eq. (3.44) can be written as

$$\left(F_{\mu\nu} \pm \tilde{F}_{\mu\nu}\right)^2 = 2\left(F_{\mu\nu}F^{\mu\nu} \pm F_{\mu\nu}\tilde{F}^{\mu\nu}\right), \quad (3.46)$$

such that the inequality

$$\text{Tr} \int d^4x F_{\mu\nu}F^{\mu\nu} \geq \left| \text{Tr} \int d^4x F_{\mu\nu}\tilde{F}^{\mu\nu} \right| = 16\pi^2|\nu| \quad (3.47)$$

is obtained. The last equal sign is established by using eq. (3.16) and realizing that the fields are rescaled according to eq. (3.36). From this inequality a lower bound on the Euclidean action follows:

$$S_E[A_\mu] \geq 16\pi^2|\nu|. \quad (3.48)$$

So there is a different lower bound on the action in every homotopy class of the finite Euclidean action field configurations.

A property of a (local) minimum of the action is that the action should not decrease under infinitesimal changes of the field. These infinitesimal changes are continuously obtainable from the original field and so they keep the field in the same homotopy class. For the field configuration at which equality is achieved in eq. (3.48) every infinitesimal transformation will therefore not lower the action. This establishes the fact that there is at least one local minimum of the action in every homotopy class.

From the general form in eq. (3.43) it is seen that, if there are multiple inequivalent minima of the action, then the exponential in front is not just an overall factor which cancels out in the calculation of correlation functions. And from the form of this exponential it is also clear that in ‘ordinary’ perturbation theory this result could never be obtained. This is because a function

$$f(g) = e^{-1/g^2}, \quad (3.49)$$

has no power (Taylor) expansion around $g = 0$, because of the singularity at that point.

One could question why the method of steepest descent and ordinary perturbation theory do not agree and which one should be trusted more. The method of steepest descent is based on the assumption that fluctuations of the potential field around the classical solutions is small instead of the much simpler assumption, of ordinary perturbation theory, that the fluctuations around zero are small. The last assumption is clearly less valid since we already discovered multiple vacua with nonzero potential.

The next question is how one should interpret these instanton field configurations. From the path integral formulation of quantum mechanics (see e.g. section 7.2 in [8]), it is clear that non-trivial solutions to the classical Euclidean equations of motion, i.e. the instanton solutions, correspond to quantum mechanical tunneling. It adds the possibility of tunneling from one vacuum state (at $t_E = -\infty$) to another (at $t_E = \infty$) to the path integral.

In the next section it will be shown, by an example, that in non-Abelian gauge field theories instantons correspond to a tunneling from one winding-number vacuum to another. In fact, *an instanton with winding number ν connects two winding-number vacua, n , at $t_E = \infty$, and m , at $t_E = -\infty$, which differ by ν units, i.e. $\nu = n - m$.*

3.5 Example: instanton with winding number one

Finding instanton solutions is not trivial. However, checking whether a given field configuration is an instanton or not is in principle straightforward. If the action integral satisfies the equality in eq. (3.47) for a nonzero ν , then the field configuration is a local minimum of the action, i.e. an instanton.

Belavin *et al.* [9] were the first to constructed an instanton with $\nu = 1$ solution in $SU(2)$ gauge theory. The vector potentials have the form

$$A_\mu(x) = \left(\frac{x^2}{x^2 + \lambda^2} \right) U(x) \partial_\mu U^{-1}(x), \quad (3.50)$$

where λ is an arbitrary parameter, often called the *instanton size*, and

$$U(x) = \frac{x_0 + i\mathbf{x} \cdot \boldsymbol{\sigma}}{|x|}. \quad (3.51)$$

Far away from the center of the instanton (at $x_\mu = 0$ in this case), i.e. $|x| \gg \lambda$, the potentials behave as

$$A_\mu(x) \rightarrow U(x)\partial_\mu U^{-1}(x), \quad (3.52)$$

as is required by the finiteness of the action. Also one can check that the value of the action integral equals $16\pi^2$ as it should be for a $\nu = 1$ instanton.

To compare the instanton solution with the winding-number vacua, they should be in the same gauge. The winding-number vacua were calculated in the temporal gauge, i.e. $A_0(x) = 0$. To transform the instanton solution, eq. (3.50), to this gauge, use the transformation (see section 16.2 in [10])

$$A_\mu(x) \rightarrow A'_\mu(x) = V(x)A_\mu(x)V^{-1}(x) + V\partial_\mu V^{-1}(x). \quad (3.53)$$

The condition $A'_0(x) = 0$ implies

$$\begin{aligned} \frac{\partial}{\partial x_0} V^{-1}(x) &= -A_0(x)V^{-1}(x) \\ &= \frac{-i\mathbf{x} \cdot \boldsymbol{\sigma}}{x_0^2 + \mathbf{x}^2 + \lambda^2} V^{-1}(x), \end{aligned} \quad (3.54)$$

where in the last line the specific form of the instanton in eq. (3.50) is used. The condition on $V^{-1}(x)$ is sufficient to determine its form up to an integration constant Θ_0 :

$$V^{-1}(x) = \exp \left[\frac{-i\mathbf{x} \cdot \boldsymbol{\sigma}}{\sqrt{\mathbf{x}^2 + \lambda^2}} \left(\arctan \left\{ \frac{x_0}{\sqrt{\mathbf{x}^2 + \lambda^2}} \right\} + \Theta_0 \right) \right]. \quad (3.55)$$

The integration constant can be fixed by boundary conditions at $x_0 = \pm\infty$. The choice

$$\Theta_0 = (n + \frac{1}{2})\pi \quad (3.56)$$

will turn out to be the right one to connect different winding-number vacua. This can be seen by realizing that

$$A_\mu(x) \rightarrow 0 \quad \text{as} \quad x_0 \rightarrow \pm\infty, \quad (3.57)$$

so

$$A'_i \rightarrow V(x)\partial_i V^{-1}(x) \quad \text{as} \quad x_0 \rightarrow \pm\infty, \quad (3.58)$$

with

$$V(x_0 = -\infty) = \exp \left[i\pi \frac{\mathbf{x} \cdot \boldsymbol{\sigma}}{\sqrt{\mathbf{x}^2 + \lambda^2}} n \right] \quad (3.59)$$

and

$$V(x_0 = \infty) = \exp \left[i\pi \frac{\mathbf{x} \cdot \boldsymbol{\sigma}}{\sqrt{\mathbf{x}^2 + \lambda^2}} (n + 1) \right]. \quad (3.60)$$

Thus a $\nu = 1$ instanton solution in eq. (3.50) connects winding-number vacua (eq. (3.31)) which differ by one unit of winding number.

Chapter 4

Consequences of instantons

In the previous chapter we discussed non-trivial gauge field configurations. Those field configurations significantly influence the predictions of the theory. The effects of instantons are not yet fully understood. Some striking consequences have, however, been derived. In this chapter a couple of those consequences will be discussed.

4.1 Chirality non-conservation

One of those consequences is the non-conservation of chirality, even in the massless limit, through the axial anomaly. An anomalous symmetry is a symmetry on the classical level, which is broken in the quantum theory. This means that, in the classical theory, the current belonging to the symmetry transformation is conserved, but in the quantum theory the divergence of the current is non-zero.

In section 2.4 two different ‘baryon-number’ transformations were introduced. These were a normal and an *axial* (or *chiral*) transformation. In the limit of massless fermions, the chiral transformation is a symmetry of the Lagrangian.

To see that a chiral transformation leaves the classical Lagrangian invariant, consider the Lagrangian density of a simple model consisting of one free fermion, which is gauge invariant under some non-Abelian gauge group,

$$\mathcal{L} = \bar{\Psi} (i\gamma^\mu D_\mu - m) \Psi - \frac{1}{4} \text{Tr} [F_{\mu\nu} F^{\mu\nu}]. \quad (4.1)$$

The chiral transformation acts on the fermion as

$$\begin{aligned} \Psi &\rightarrow e^{i\alpha\gamma^5} \Psi, \\ \bar{\Psi} &\rightarrow \left(e^{i\alpha\gamma^5} \Psi \right)^\dagger \gamma^0 = \bar{\Psi} e^{i\alpha\gamma^5}. \end{aligned} \quad (4.2)$$

In the last line it has been used that $\{\gamma^5, \gamma^\mu\} = 0$. Inserting the transformed fields in the Lagrangian density in eq. (4.1) and using the anti-commutator relation shows the Lagrangian density transforms as

$$\mathcal{L} \rightarrow \bar{\Psi} \left(i\gamma^\mu D_\mu - e^{2i\alpha\gamma^5} m \right) \Psi - \frac{1}{4} \text{Tr} [F_{\mu\nu} F^{\mu\nu}]. \quad (4.3)$$

This makes it clear that in the massless case the Lagrangian is invariant under a chiral transformation.

According to Noether’s theorem, see e.g. [7], to every symmetry of the Lagrangian corresponds a conserved current. It follows from the fact that, when the equations of motion hold, the Lagrangian should be stationary under every infinitesimal transformation of the fields.

When the transformation in eq. (4.2) is made local (i.e. $\alpha \rightarrow \alpha(x)$), the Lagrangian density from eq. (4.1), with $m = 0$, transforms as

$$\int d^4x \mathcal{L} \rightarrow \int d^4x [\mathcal{L} - \partial_\mu \alpha(x) \bar{\Psi} \gamma^\mu \gamma^5 \Psi] = \int d^4x [\mathcal{L} + \alpha(x) \partial_\mu (\bar{\Psi} \gamma^\mu \gamma^5 \Psi)]. \quad (4.4)$$

The Lagrangian should be stationary under variations of $\alpha(x)$, when the fields obey the equations of motion, so

$$\frac{\delta L}{\delta \alpha(x)} = \partial_\mu (\bar{\Psi} \gamma^\mu \gamma^5 \Psi) = 0. \quad (4.5)$$

This stationarity of the Lagrangian thus leads to the conserved chiral current

$$j^{\mu 5} = \bar{\Psi} \gamma^\mu \gamma^5 \Psi. \quad (4.6)$$

Written in chiral right- and left-handed components of a spinor,

$$\Psi_{R,L} = \frac{1 \pm \gamma_5}{2} \Psi, \quad (4.7)$$

the zeroth component of the chiral current in eq. (4.6) is

$$j^{05} = \Psi_R^\dagger \Psi_R - \Psi_L^\dagger \Psi_L, \quad (4.8)$$

which is nothing but the number density of chiral right-handed particles minus the number density of chiral left-handed particles. So the conserved charge, belonging to the chiral current,

$$Q^5 \equiv \int d^3x j^{05}, \quad (4.9)$$

is the difference in number of chiral right- and left-handed particles.

To see whether the previously derived conservation law still holds in the quantized theory, the path integral formalism can be used. A conservation law in the quantum theory is usually derived by using the fact that the functional integral,

$$Z = \int [d\Psi][d\bar{\Psi}][dA_\mu] e^{i \int d^4x \mathcal{L}}, \quad (4.10)$$

should be stationary under coordinate transformations. Consider, again, the transformation in eq. (4.2), but with α dependent on x . The transformation of the action, under this transformation, is that of eq. (4.4). In the functional integral, there is one more quantity which could change under the chiral transformation and therefore spoil the symmetry: the functional measure.

The transformation properties of the functional measure will not be derived here. The derivation can be found in e.g. chapter 22.2 in [16] or chapter 19.2 in [7]. It turns out that the measure is *not* invariant under the local version of the chiral transformation in eq. (4.2). The measure transforms as

$$[d\Psi][d\bar{\Psi}] \rightarrow \exp \left[i \int d^4x \alpha(x) \mathcal{A}(x) \right] [d\Psi][d\bar{\Psi}], \quad (4.11)$$

where $\mathcal{A}(x)$ is called the *anomaly function*, it can be expressed in terms of the gauge fields as

$$\mathcal{A}(x) = -\frac{g^2}{16\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} [F_{\mu\nu} F_{\rho\sigma}]. \quad (4.12)$$

The effect of the chiral transformation, on the functional measure in the path integral expression eq. (4.10), is of a form such that it can also be written as an effective change in the Lagrangian density,

$$\mathcal{L}_{\text{eff}} \rightarrow \mathcal{L}'_{\text{eff}} + \alpha(x) \mathcal{A}(x). \quad (4.13)$$

Applying the transformation laws for the action (eq. (4.4)) and the functional measure (eq. (4.11)), the effect on the the functional integral, eq. (4.10), is determined to be

$$Z \rightarrow \int [d\Psi][d\bar{\Psi}][dA_\mu] e^{i \int d^4x \mathcal{L} + \alpha(x) \partial_\mu j^{\mu 5} + \alpha(x) \mathcal{A}(x)}. \quad (4.14)$$

Since the chiral transformation in the functional integral is nothing more than a change of variables, it should be stationary under arbitrary variation of $\alpha(x)$. This leads us to the equation

$$\left. \frac{\delta Z}{\delta \alpha(x)} \right|_{\alpha=0} = \int [d\Psi][d\bar{\Psi}][dA_\mu] i (\mathcal{A}(x) + \partial_\mu j^{\mu 5}) e^{i \int d^4 x \mathcal{L}} = 0. \quad (4.15)$$

The path integral in the last equation is an expression for the expectation value of the operator in the brackets. So in the quantum theory the expectation value for the divergence of the chiral current is

$$\langle \partial_\mu j^{\mu 5} \rangle = -\langle \mathcal{A}(x) \rangle = \frac{g^2}{16\pi^2} \langle \epsilon^{\mu\nu\rho\sigma} \text{Tr} [F_{\mu\nu} F_{\rho\sigma}] \rangle. \quad (4.16)$$

In the massless quantum theory, the amount of left- and right-handed particles is not separately conserved in any process. This is seen from the volume integral over the divergence of the current,

$$\begin{aligned} \int d^4 x \partial_\mu j^{\mu 5} &= \int dx_0 \partial_0 \int d^3 x j^{05} - \int dx_0 \int d^4 x \partial_i j^{i5} = \int dx_0 \partial_0 Q^5 - \int_{\partial \mathbb{R}^3} dS \hat{n}_i j^{i5} \\ &= Q_f^5 - Q_i^5, \end{aligned} \quad (4.17)$$

where $Q_{f,i}^5$ represent the final and initial difference in right- and left-handed particles respectively. In the last equation the assumption has been made that the flux of the current through the sphere at spatial infinity is zero. The left-hand side of the equation can be evaluated in a background gauge field $F_{\mu\nu}$, with the use of eq. (4.16):

$$Q_f^5 - Q_i^5 = \int d^4 x \frac{g^2}{16\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} [F_{\mu\nu} F_{\rho\sigma}]. \quad (4.18)$$

Comparing the right-hand side of this equation with eq. (3.16), it turns out that it is just an expression for the winding number of the background $F_{\mu\nu}$ field. So in the presence of a background instanton with winding number ν , the conservation of left- and right-handed particles is violated by the amount

$$Q_f^5 - Q_i^5 = -2\nu. \quad (4.19)$$

This means that instantons convert right-handed particles into left-handed particles and anti-instantons (with negative winding number) do exactly the opposite. This also shows that the expectation value of the divergence of the chiral current, eq. (4.16), would just be zero when one only includes trivial gauge field configurations (with zero winding number) in the path integral. Instantons cause a violation of the separate conservation of left- and right-handed fermions.

4.2 theta-vacua

In section 3.3 the vacuum structure of a classical non-Abelian gauge field theory was explored. It turned out there is a huge degeneracy of physically equivalent vacuum fields. They could be divided into topologically different homotopy classes, labeled by a winding number and were called winding-number vacua.

The quantum theory should be built on a single vacuum state. All classical vacua are physically equivalent. One might just pick one classical (winding number) vacuum with, e.g. winding number n , to be the ground state of the quantum theory: $|n\rangle$. It will be shown now that the right quantum mechanical vacuum, for a non-Abelian gauge theory, is not a winding-number vacuum, but a θ -vacuum.

Let us denote the transition amplitude from one winding-number vacuum to another by

$${}_{\text{out}} \langle n|m \rangle_{\text{in}} = Z_\nu, \quad \text{where } \nu = n - m. \quad (4.20)$$

The amplitude does not depend on the sum of the winding numbers, since large gauge transformations (eq. (3.32), in which V has a non-zero winding number) change the winding numbers of the in- and

out-vacuum states by the same amount. These gauge transformation should not change the amplitude, therefore it can only depend on the *difference* of the winding numbers.

The winding-number vacuum transition amplitude, eq. (4.20), is, in general, non-zero for $n \neq m$, because of the instanton field configurations, described in section 3.4. Those field configurations represent tunneling from one winding-number vacuum to another. Because of this overlap of the winding-number vacuum states, these *cannot* be used as the ground state of a quantum theory.

The quantum mechanical vacua should be labeled by a time independent quantity and states with different labels should not overlap. The eigenvalues of operators which commute with the Hamiltonian are time independent. The Hamiltonian is invariant under gauge transformations, so it commutes with the gauge transformation operator,

$$[T_m, H] = 0, \quad (4.21)$$

where T_m is a (large) gauge transformation with winding number m . It changes the winding number of a winding-number vacuum by the amount m ,

$$T_m|n\rangle = |n+m\rangle. \quad (4.22)$$

The eigenstates of these operators should have the right properties for a vacuum state. The eigenstates can be constructed as a sum over winding-number vacua:

$$|\theta\rangle = \sum_n e^{in\theta} |n\rangle. \quad (4.23)$$

These states are labeled by $\theta \in \mathbb{R} \bmod 2\pi$ and called θ -vacua. It is easily shown to be an eigenstate of T_m ,

$$T_m|\theta\rangle = \sum_n e^{in\theta} |n+m\rangle = \sum_{n'} e^{i(n'-m)\theta} |n'\rangle = e^{-im\theta} |\theta\rangle. \quad (4.24)$$

The different θ -vacua do not overlap,

$$\begin{aligned} {}_{\text{out}}\langle\theta'|\theta\rangle_{\text{in}} &= \sum_{n,m} e^{-im\theta'} e^{in\theta} {}_{\text{out}}\langle m|n\rangle_{\text{in}} \\ &= \sum_{\nu,k} e^{\frac{i}{2}(\theta'+\theta)\nu} e^{\frac{i}{2}(\theta'-\theta)k} {}_{\text{out}}\langle\nu|0\rangle_{\text{in}} = \delta(\theta' - \theta) \sum_{\nu} e^{i\theta\nu} {}_{\text{out}}\langle\nu|0\rangle_{\text{in}}. \end{aligned} \quad (4.25)$$

The vacuum to vacuum transition amplitude in the θ -vacuum can be written as a path integral through

$$\begin{aligned} {}_{\text{out}}\langle\theta|\theta\rangle_{\text{in}} &\propto \sum_{\nu} e^{i\theta\nu} {}_{\text{out}}\langle\nu|0\rangle_{\text{in}} = \sum_{\nu} \int [dA_{\mu}]_{\nu} e^{i\theta\nu} e^{\int d^4x \mathcal{L}^E} \\ &= \sum_{\nu} \int [dA_{\mu}]_{\nu} \exp \left\{ \int d^4x \left(\mathcal{L}^E - \frac{i\theta g^2}{32\pi^2} \text{Tr}[\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}] \right) \right\}, \end{aligned} \quad (4.26)$$

where $\int [dA_{\mu}]_{\nu}$ means integration over all finite action gauge field configurations with winding number ν . The winding number has been rewritten in terms of the gauge fields with the use of eq. (3.16).

From eq. (4.26), it can be concluded that picking the right vacuum is nothing more than replacing the Lagrangian density by a new effective Lagrangian density, which contains a ' θ -term':

$$\mathcal{L} \rightarrow \mathcal{L}_{\text{eff}} = \mathcal{L} + \frac{\theta g^2}{32\pi^2} \text{Tr}[\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}]. \quad (4.27)$$

This is exactly the term which we already considered in section 2.2. It could be left out of the Lagrangian in the absence of instantons, but in the presence of instantons it has to be there. The value of θ could be anything, it is a property of the vacuum, but it *cannot* change in time.

4.3 The strong CP -problem

The effects of the θ -vacuum are taken into account by replacing the Lagrangian density with the effective one in eq. (4.27). The new term, however, can have phenomenological consequences. The term transforms as

$$\epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \xrightarrow{P,T} -\epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \quad (4.28)$$

under P and T transformations (see eq. (2.29)). This implies that the theory is only P -, T - and CP -invariant if $\theta = -\theta$, i.e. $\theta \bmod \pi = 0$ (because of the periodicity of θ). The value of θ could be anything and therefore QCD does not naturally conserve CP .

In the presence of a massless fermion the theory, however, *is* CP -conserving. The extra θ -term can be ‘rotated away’ by a chiral transformation. This is possible, because a chiral rotation on the massless fermion field, under an angle α , (eq. (4.2)) has the effect of, according to eq. (4.13), changing $\theta \rightarrow \theta - 2\alpha$. Since there is no mass term in the Lagrangian, nothing else is changing under the chiral rotation. A change of variables (the chiral rotation) in the path integral should not be observable. So the predictions of the theory cannot depend on the value of θ and therefore it is equivalent to one with $\theta = 0$, i.e. a CP -conserving one.

The same transformation could, of course, also be performed when the fermion is massive. The mass term in the Lagrangian density is, however, not invariant under a chiral transformation. If the mass term is written as

$$m \bar{\Psi}_L \Psi_R + m^* \bar{\Psi}_R \Psi_L, \quad (4.29)$$

then under a chiral transformation, with $\alpha = \theta/2$, the θ -term is removed, but the mass becomes complex, i.e.

$$\theta \rightarrow 0, \quad \text{but} \quad m \rightarrow e^{i\theta} m. \quad (4.30)$$

The new mass term violates CP and so we can move the θ -term into the mass, but not eliminate CP -violation. So the actual condition, whether the theory is CP -conserving or not, should be that

$$CP \text{ invariant if } \theta = 0 \bmod \pi \text{ when } m \in \mathbb{R}. \quad (4.31)$$

The experimental bounds on CP -violation in the strong interaction are very strong. The strongest bound comes from measurements of the electric dipole moment of the neutron. It can be calculated, see e.g. [3], that a non-zero θ -parameter induces an electric dipole moment of the neutron with the size of

$$d_n \sim 5.2 \cdot 10^{-16} |\theta| e \text{ cm}, \quad (4.32)$$

in which is assumed that θ is small. Measurements of d_n , see [2], have set an upper limit of

$$d_n < 0.29 \cdot 10^{-25} e \text{ cm}. \quad (4.33)$$

This leads to a strict limit on θ , being

$$|\theta| < 5.6 \cdot 10^{-11}. \quad (4.34)$$

In general one would expect θ to be of order 1. The extremely small experimental value is considered to be *unnatural*. It suggests that there is some underlying reason for θ to be this small. This problem is called the *strong CP -problem*.

4.4 Energy of the theta-vacua

The Vafa-Witten theorem [17] says that, in every arbitrary parity-conserving vectorlike gauge field theory, parity is *not spontaneously broken*. This is done by proving that no parity-odd operator can have a non-zero vacuum expectation value.

The argument that is used in the proof of the Vafa-Witten theorem can also be applied to the θ -dependence of the vacuum energy density for non-Abelian gauge field theories. The energy density of the different θ -vacua in a pure non-Abelian $SU(N)$ gauge field theory can be expressed in terms of a path integral through

$$e^{-\mathcal{V}_4 E(\theta)} = \sum_{\nu} \int [dA_{\mu}]_{\nu} e^{-S_{E,\theta}^{\text{eff}}[A_{\mu}]}, \quad (4.35)$$

where $S_{E,\theta}^{\text{eff}}$ is the Euclidean effective action in a θ -vacuum, $\int [dA_{\mu}]_{\nu}$ means integration over all gauge fields with winding number ν and \mathcal{V}_4 is the volume of 4-dimensional Euclidean space. The effective action can be read off from eq. (4.26), resulting in

$$e^{-\mathcal{V}_4 E(\theta)} = \sum_{\nu} \int [dA_{\mu}]_{\nu} \exp \left[- \int d^4x \left(\frac{1}{4} \text{Tr} [F_{\mu\nu} F^{\mu\nu}] + \frac{i\theta g^2}{32\pi^2} \epsilon^{\mu\nu\rho\sigma} \text{Tr} [F_{\mu\nu} F_{\rho\sigma}] \right) \right]. \quad (4.36)$$

The traces over the field matrices can be expressed in terms of the N real-valued fields, $F_{\mu\nu}^a$, by

$$\text{Tr} [F_{\mu\nu} F_{\rho\sigma}] = F_{\mu\nu}^a F_{\rho\sigma}^b \text{Tr} [A^a A^b] \propto F_{\mu\nu}^a F_{\rho\sigma}^b \delta^{ab} = F_{\mu\nu}^a F_{\rho\sigma}^a, \quad (4.37)$$

where A^a are the generators of the group. The fields $F_{\mu\nu}^a$ are real, both in Minkowski and Euclidean space, because the integration bounds on these fields do not change under a Wick rotation (see appendix A).

For $\theta = 0$, the integrand in eq. (4.36) is real. For non-zero θ , the last term in the exponential only contributes a phase factor and can, therefore, only *lower* the integral. This implies, for the energy of the θ -vacuum, that

$$E(\theta) \geq E(0). \quad (4.38)$$

From this, it can be concluded that the CP conserving vacuum state, i.e. $\theta = 0$, has minimal energy. One can also see that, if there were no instantons, i.e. the integration in eq. (4.36) is only done for $\nu = 0$, then all θ -vacua would be degenerate. Instantons cause the energy splitting of the θ -vacua.

The θ dependence can, in fact, be written down straightforwardly if one only includes $\nu = \pm 1$ instantons, by taking the logarithm of eq. (4.36) and then restricting the path integral to $\nu = 0, \pm 1$ fields only. Taking the logarithm has, on the right hand side, the effect of only including connected diagrams, such that

$$-\mathcal{V}_4 E(\theta) \simeq \int_{\mathcal{C}} [dA_{\mu}]_0 e^{-S_E[A_{\mu}]} + e^{-i\theta} \int_{\mathcal{C}} [dA_{\mu}]_1 e^{-S_E[A_{\mu}]} + e^{i\theta} \int_{\mathcal{C}} [dA_{\mu}]_{-1} e^{-S_E[A_{\mu}]}. \quad (4.39)$$

From eq. (3.48) it is known¹ that the Euclidean action, for a field with winding number ν , has a lower bound,

$$S_E[A_{\mu}] \geq \frac{16\pi^2 |\nu|}{g^2}, \quad (4.40)$$

where equality is obtained for the instanton configuration. This can be used in the expression for the approximate energy of the θ -vacuum in eq. (4.39) to obtain a lower bound on the energy:

$$\begin{aligned} -\mathcal{V}_4 E(\theta) &\leq \int [dA_{\mu}]_0 e^0 + \cos(\theta) \int [dA_{\mu}]_1 e^{-16\pi^2/g^2} \\ \Rightarrow \mathcal{V}_4 E(\theta) &\geq -C_0 - \cos(\theta) e^{-16\pi^2/g^2} C_1, \end{aligned} \quad (4.41)$$

where $C_{0,1}$ are positive constants *independent* of θ . If quantum fluctuations of the field around the instanton solution are neglected, i.e. the path integral in eq. (4.39) is restricted to only instanton configurations, then equality is obtained in eq. (4.40) and, therefore, also in the last equation.

¹Remember that eq. (3.48) is written in terms of the rescaled fields in eq. (3.36), the partition function in that representation is given by eq. (3.37) instead of eq. (4.35), which is used here.

The fact that the lowest energy θ -vacuum is the one with $\theta = 0$ does, however, not mean that the system will always relax into the CP -conserving $\theta = 0$ ground state. This is because θ is not a dynamical variable, i.e. it *cannot* change in time (see eq. (4.25)), it is just a parameter which has to be supplied before any calculation can be done. The vacuum energy density dependence on θ is, however, used in a particular (proposed) solution to the strong CP -problem: the Peccei-Quinn mechanism, which will be discussed in the next chapter.

Chapter 5

The Peccei-Quinn mechanism

In section 4.3 the strong CP -problem was discussed. It was mentioned that a zero quark mass would make the theory CP conserving for every value of θ . Therefore a vanishing mass of one of the quarks would resolve the strong CP -problem.

It is widely believed, however, that a non-zero quark mass is inconsistent with experimental results. For example, in the review on quark masses [18], the ratio of the lightest quarks (the up- and down-quark) is determined to be $m_d/m_u = 1.76 \pm 0.13$, making a zero-mass up-quark highly improbable.

A theoretical framework is needed in order to extract quark masses from experiments since quarks cannot be ‘weighed’ directly. Leaving open the option of differing theoretical models, in [19], it is concluded that *“while a vanishing up-quark mass is not rigorously ruled out, it is unattractive from the standpoint of the currently consistent phenomenology of hadronic symmetry breaking.”*

Also theoretical problems with zero quark mass have been brought forward in e.g. [20] and [21]. The most striking argument (by Creutz) is the one that m_u is not renormalization group invariant. This means that the condition $m_u = 0$ is renormalization scheme dependent, hence non-physical, since physical observables should not depend on the renormalization used. A non physical particle mass might sound strange, but single quarks are not observed and meson and baryon masses are only dependent on quark mass ratios. For a zero quark mass to be a solution to the strong CP -problem it should be zero at all energy scales and in every renormalization scheme. Setting $m_u = 0$ in some scheme at some energy does not imply that it is always zero and so it is not a solution to the strong CP -problem. Altogether, a vanishing quark mass is not considered to be a solution to the strong CP -problem.

Another possible solution to the strong CP -problem is the Peccei-Quinn mechanism. It was proposed in 1977 by R. D. Peccei and Helen R. Quinn, see [4]. They showed that, if the QCD Lagrangian possesses an axial $U(1)$ symmetry, which is broken only by the axial anomaly, then CP -conservation is automatically satisfied. This type of additional symmetry has become known as a Peccei-Quinn symmetry.

A zero quark mass would actually be an example of a Peccei-Quinn symmetry; an axial rotation on the massless quark field is a symmetry of the Lagrangian, which is only broken by the axial anomaly. There are, however, more ways to realize a Peccei-Quinn symmetry. A better option (considering the previously discussed problems with vanishing quark mass) is a realization of the Peccei-Quinn symmetry in Goldstone mode, i.e. a spontaneously broken symmetry.

One consequence of a spontaneously broken Peccei-Quinn symmetry is the prediction of (pseudo) Goldstone bosons, called *axions*. Since the Peccei-Quinn symmetry is not exact (due to the anomaly), the resulting pseudo Goldstone bosons will not be exactly massless. Depending on the specific implementation of the Peccei-Quinn symmetry, these axions should have been seen already in experiments, or could have escaped detection so far.

In the next section the basics of the Peccei-Quinn mechanism will be described. In section 5.2 it will be explicitly shown that the Peccei-Quinn mechanism realized in a simple toy model is CP -conserving. Section 5.3 will discuss the axion that is predicted by the toy model. The remaining sections will cover the problems with the standard Peccei-Quinn model and possible solutions in the form of *invisible axions*.

5.1 Basics of the Peccei-Quinn mechanism

In a QCD-like non-Abelian gauge field theory, e.g.

$$\mathcal{L} = -\frac{1}{4}\text{Tr}[F_{\mu\nu}F^{\mu\nu}] + i\bar{\Psi}\not{D}\Psi - \bar{\Psi}_L M \Psi_R - \bar{\Psi}_R M^\dagger \Psi_L + \frac{\theta g^2}{16\pi^2}\text{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}], \quad (5.1)$$

where Ψ is a vector of different fermion flavors, all terms except the mass and θ -term are P and CP invariant. The mass term transforms as

$$-\bar{\Psi}_L M \Psi_R - \bar{\Psi}_R M^\dagger \Psi_L \xrightarrow{P,CP} -\bar{\Psi}_R M \Psi_L - \bar{\Psi}_L M^\dagger \Psi_R, \quad (5.2)$$

which is CP invariant if $M = M^\dagger$. The θ -term transforms as

$$\frac{\theta g^2}{16\pi^2}\text{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}] \xrightarrow{P,CP} -\frac{\theta g^2}{16\pi^2}\text{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}], \quad (5.3)$$

which is CP invariant if $\theta = 0 \pmod{\pi}$ (remember that θ is a periodic parameter, with period 2π).

All CP violation can be put into an effective θ term. To do this, one first has to work in the appropriate flavor basis, such that the mass matrix is diagonal, i.e. $M \rightarrow M_D$, and then perform an axial rotation on every flavor, such that every mass in M_D is real. The axial transformations will change the value of θ (see section 4.3) to

$$\theta \rightarrow \theta_{\text{eff}} = \theta + \log \det M. \quad (5.4)$$

A theory, in which the only source of CP violation is an effective θ -term, can be rendered CP conserving by introducing a Peccei-Quinn (PQ) symmetry. This PQ symmetry is an additional axial $U(1)$ symmetry, which has to be broken by the axial anomaly.

By making a axial rotation belonging to the PQ symmetry, the value of θ_{eff} will change due to the axial anomaly (the anomalous breaking of the PQ symmetry is thus essential for the mechanism to work). A mere change of variables will not alter the predictions of the theory. So with this symmetry, $\theta_{\text{eff}} \neq 0$ is physically equivalent to $\theta_{\text{eff}} = 0$, i.e. the theory has no explicit CP violation. The PQ symmetry could, in principle, be spontaneously broken. To assure that the theory is then still CP conserving, one has to check that no spontaneous CP violation occurs, much in the same way as the Vafa-Witten theorem.

If the PQ symmetry is spontaneously broken, then the low energy effective theory (see for low-energy effective field theories e.g. 19.5 and 19.6 in [16]) has the form

$$\mathcal{L}_a = (\partial_\mu a) J^\mu + \left(\frac{a}{F_a} + \theta_{\text{eff}} \right) \frac{g^2}{16\pi^2} \text{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}], \quad (5.5)$$

where a is a low-energy degree of freedom, J^μ the current belonging to the PQ transformation and F_a is the breaking scale of the PQ symmetry. The field a behaves under a PQ transformation, with angle α , as $a \rightarrow a + \alpha F_a$. The first term is invariant under such transformations, while the second represents the breaking of the PQ symmetry. From this form it can be seen that this theory has an 'effective θ field', defined by

$$\Theta(x) = a/F_a + \theta_{\text{eff}}. \quad (5.6)$$

This model does not, as shown before, violate P and CP explicitly. From that fact, it can be concluded that under these transformations the Θ field transforms as

$$\Theta = \left(\frac{a}{F_a} + \theta_{\text{eff}} \right) \xrightarrow{P,CP} -\Theta = -\left(\frac{a}{F_a} + \theta_{\text{eff}} \right), \quad (5.7)$$

since θ_{eff} is just a parameter, this implies that the transformation of the a field is

$$a \xrightarrow{P,CP} -a - 2\theta_{\text{eff}}F_a. \quad (5.8)$$

The condition for a non CP -violating vacuum expectation value is thus

$$\langle \Theta \rangle = 0, \quad \text{or equivalently} \quad \langle a \rangle = -\theta_{\text{eff}} F_A. \quad (5.9)$$

This is like the CP -conservation condition $\theta_{\text{eff}} = 0$ in QCD, except that Θ is a field which can change in time.

From the Vafa-Witten theorem (in section 4.4) it is known that in a pure non-Abelian gauge theory the lowest energy θ -vacuum is the one with $\theta = 0$. We could, therefore, expect that in this theory, with an ‘effective θ field’, $\Theta(x)$, the vacuum state will be such that $\langle \Theta(x) \rangle = 0$, which is exactly the CP conserving ground state.

In the next section it will be explicitly shown, in first approximation, that for a toy model with a Peccei-Quinn symmetry the ground state of the theory is CP conserving.

5.2 Example: the Peccei-Quinn mechanism in a toy model

This example consists of a $SU(3)$ non-Abelian gauge field theory with a single fermion flavor, transforming in the defining representation of the gauge group. There is no ‘standard’ mass-term for the fermion. Instead, it will be coupled to a scalar field (invariant under $SU(3)$), which develops a non-zero vacuum-expectation-value. This is the same way how the standard model particles get their mass, except that there is no gauge symmetry that is spontaneously broken, which would lead to the Higgs mechanism¹.

The generating functional, in a θ -vacuum, for this model can be written as

$$Z_\theta = \sum_{\nu=-\infty}^{\infty} \int [dA_\mu]_n [d\phi][d\phi^*][d\Psi][d\bar{\Psi}] e^{\int d^4x \mathcal{L}_E[A_\mu, \phi, \phi^*, \Psi, \bar{\Psi}]} e^{i\nu\theta}, \quad (5.10)$$

in which the Euclidean Lagrangian density, \mathcal{L}_E , is given by

$$\begin{aligned} \mathcal{L}_E = & -\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi + \bar{\Psi} \left[G\phi \left(\frac{1+\gamma_5}{2} \right) + G^* \phi^* \left(\frac{1-\gamma_5}{2} \right) \right] \Psi \\ & - |\partial_\mu \phi|^2 - \mu^2 |\phi|^2 - h |\phi|^4. \end{aligned} \quad (5.11)$$

The parameter $h > 0$ and $\mu^2 < 0$, in order to let the ϕ field develop a vacuum expectation value.

The Lagrangian density in eq. (5.11) is invariant under the axial $U(1)$ transformation

$$\begin{aligned} \Psi_R & \rightarrow e^{-i\alpha} \Psi_R \\ \Psi_L & \rightarrow e^{i\alpha} \Psi_L \\ \phi & \rightarrow e^{2i\alpha} \phi. \end{aligned} \quad (5.12)$$

This transformation involves an axial rotation on the Ψ field. Since the Ψ field is charged under a non-Abelian gauge group, the axial rotation is anomalous (see section 4.1). The Lagrangian thus has a axial $U(1)$ symmetry which is broken by the axial anomaly, i.e. a Peccei-Quinn symmetry. This theory should then, according to Peccei and Quinn [4], be CP -conserving.

It will first be shown that this theory does not explicitly violate CP . With the use of an axial transformation on the fermion field, the θ -term can be put into the mass term (see section 4.3):

$$\mathcal{L}_E^{\text{mass}} = G e^{i\theta} \phi \bar{\Psi} \frac{1+\gamma_5}{2} \Psi + G^* e^{-i\theta} \phi^* \bar{\Psi} \frac{1-\gamma_5}{2} \Psi. \quad (5.13)$$

The $F_{\mu\nu}^a F_{\mu\nu}^a$ and fermion derivative term are P and CP invariant. The mass term is, as in the last section, invariant if

$$G e^{i\theta} \phi \xrightarrow{P, CP} G^* e^{-i\theta} \phi^*, \quad (5.14)$$

¹The Higgs mechanism is the spontaneous breaking of gauge symmetries, giving mass to the vector bosons. It is well explained in e.g. chapter 20 of [7] and chapter 21 of [16].

which is true if we define the P transformation on ϕ to be

$$\phi \xrightarrow{P, CP} e^{-2i(\theta + \arg G)} \phi^*. \quad (5.15)$$

This choice is valid, since all ϕ -only terms are also invariant under this transformation.

Knowing the P and CP transformation rules for ϕ , it is straightforward to see that the condition for no spontaneous CP violation is

$$\langle \phi \rangle \xrightarrow{P, CP} \langle \phi \rangle \quad \text{if} \quad \langle \phi \rangle = C e^{-i(\theta + \arg G)}, \quad C \in \mathbb{R}, \quad C \geq 0. \quad (5.16)$$

In the rest of this section it will be shown that the VEV of ϕ indeed takes this P and CP conserving minimum.

The generating functional in eq. (5.10), with the fermion-scalar interaction expanded out, can be written as

$$Z_\theta = \sum_{\nu=-\infty}^{\infty} \int [dA_\mu]_\nu [d\phi][d\phi^*][d\Psi][d\bar{\Psi}] e^{i\nu\theta} \exp \left[\int d^4x \left(-\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi \right) \right] \times \\ \exp \left[\int d^4x \mathcal{L}_E^\phi(\phi, \phi^*) \right] \sum_{p,q} \frac{1}{p!q!} \left[\int d^4x \bar{\Psi} G \phi \frac{1+\gamma_5}{2} \Psi \right]^p \left[\int d^4y \bar{\Psi} G^* \phi^* \frac{1-\gamma_5}{2} \Psi \right]^q, \quad (5.17)$$

where the part of the Lagrangian density that is dependent on ϕ only is written as $\mathcal{L}_E^\phi(\phi, \phi^*)$, and is defined by

$$\mathcal{L}_E^\phi(\phi, \phi^*) \equiv -|\partial_\mu \phi|^2 - \mu^2 |\phi|^2 - h|\phi|^4. \quad (5.18)$$

In the expression for the generating functional, eq. (5.17), one of the sums can be eliminated. An operator like $\bar{\Psi}(\frac{1-\gamma_5}{2})\Psi = \bar{\Psi}_R \Psi_L$ changes a left-handed particle into a right-handed particle, so it changes chirality by two units. In section 4.1, it was shown that the change in chirality in the presence of an instanton, with winding number ν , is equal to $-\nu$ (eq. (4.19)). So a term in the sum over p and q only contributes if $p - q = \nu$. The generating functional can, therefore, be written as

$$Z_\theta = \sum_{\nu=-\infty}^{\infty} \int [dA_\mu]_\nu [d\phi][d\phi^*][d\Psi][d\bar{\Psi}] e^{i\nu\theta} \exp \left[\int d^4x \left(-\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi \right) \right] \times \\ \exp \left[\int d^4x \mathcal{L}_E^\phi(\phi, \phi^*) \right] \sum_q \frac{[\int d^4x \bar{\Psi} G \phi \frac{1}{2}(1+\gamma_5) \Psi]^{q+\nu}}{(q+\nu)!} \frac{[\int d^4y \bar{\Psi} G^* \phi^* \frac{1}{2}(1-\gamma_5) \Psi]^q}{(q)!} \\ = \int [d\phi][d\phi^*] \sum_{\nu=-\infty}^{\infty} e^{i\nu\theta} \sum_q \left(\prod_{i=1}^{q+\nu} \int d^4x_i G \phi(x_i) \right) \left(\prod_{j=1}^q \int d^4y_j G^* \phi^*(y_j) \right) \times \\ \frac{c_q^\nu(x_1, \dots, x_{q+\nu}, y_1, \dots, y_q)}{q!(q+\nu)!} \exp \left[\int d^4x \mathcal{L}_E^\phi(\phi, \phi^*) \right] \\ = \int [d\phi][d\phi^*] \sum_{\nu=-\infty}^{\infty} \int \cdots \int A_\nu[\phi, \phi^*, x_1, \dots, x_\nu] G \phi(x_1) \cdots G \phi(x_\nu) d^4x_1 \cdots d^4x_\nu. \quad (5.19)$$

Where $A_\nu[\phi, \phi^*, x_1, \dots, x_\nu]$ is defined through

$$A_\nu[\phi, \phi^*, x_1, \dots, x_\nu] \equiv \sum_q \left(\prod_{i=\nu+1}^{q+\nu} \int d^4x_i G \phi(x_i) \right) \left(\prod_{j=1}^q \int d^4y_j G^* \phi^*(y_j) \right) \times \\ \frac{c_q^\nu(x_1, \dots, x_{q+\nu}, y_1, \dots, y_q)}{q!(q+\nu)!} \exp \left[\int d^4x \mathcal{L}_E^\phi(\phi, \phi^*) \right]. \quad (5.20)$$

And $c_q^\nu(x_1, \dots, x_{q+\nu}, y_1, \dots, y_q)$ is defined as

$$c_q^\nu(x_1, \dots, x_{q+\nu}, y_1, \dots, y_q) \equiv \int [dA_\mu]_\nu [d\Psi][d\bar{\Psi}] e^{\int d^4x (-\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi)} \frac{(\bar{\Psi}(x_1)^{\frac{1+\gamma_5}{2}} \Psi(x_1)) \cdots (\bar{\Psi}(x_{q+\nu})^{\frac{1+\gamma_5}{2}} \Psi(x_{q+\nu}))}{(q+\nu)!} \frac{(\bar{\Psi}(y_1)^{\frac{1-\gamma_5}{2}} \Psi(y_1)) \cdots (\bar{\Psi}(y_q)^{\frac{1-\gamma_5}{2}} \Psi(y_q))}{(q)!}. \quad (5.21)$$

Using the property

$$A_{-\nu}[\phi, \phi^*, x_1, \dots, x_\nu] = A_\nu^*[\phi, \phi^*, x_1, \dots, x_\nu], \quad (5.22)$$

the generating functional in terms of A_ν , eq. (5.19), can be written as

$$Z_\theta = \int [d\phi][d\phi^*] \left\{ A_0[\phi, \phi^*] + \sum_{\nu=1}^{\infty} \left\{ e^{i\nu\theta} \iint A_\nu[\phi, \phi^*, x_1, \dots, x_\nu] G\phi(x_1) \cdots G\phi(x_\nu) d^4x_1 \cdots d^4x_\nu \right. \right. \\ \left. \left. + e^{-i\nu\theta} \iint A_\nu^*[\phi, \phi^*, x_1, \dots, x_\nu] G^*\phi^*(x_1) \cdots G^*\phi^*(x_\nu) d^4x_1 \cdots d^4x_\nu \right\} \right\}. \quad (5.23)$$

In a classical theory, the field will relax to a value such that the potential is minimized. In a quantum theory, the vacuum expectation value of the fields will be such that the *quantum effective potential* is minimized. The concept of the quantum effective potential and the *quantum effective action* are well explained in e.g. chapter 16 in [16]. The quantum effective potential (QEP) for a field ϕ , with an x -independent expectation value ϕ_0 , is defined in terms of the quantum effective action (QEA) through

$$V(\phi_0) \equiv -\frac{1}{\mathcal{V}_4} \Gamma[\phi_0], \quad (5.24)$$

in which \mathcal{V}_4 is the volume of 4-dimensional Euclidean space. The QEP, $V(\phi_0)$, is a real function of the variable ϕ_0 . The QEA, $\Gamma[\phi]$, is a functional of the field ϕ .

One method, explained in e.g. section 16.1 in [16], to calculate the quantum effective action for a field ϕ , is to replace, in the integrand of the generating functional, the fields according to $\phi \rightarrow \phi_0 + \phi$ and restrict the path integral to include only connected, one-particle irreducible (1PI) graphs. Graphs are 1PI if and only if they *cannot* be made disconnected by cutting a single propagator line. The generating functional then becomes a function of ϕ_0 and equals the quantum effective action.

To determine the vacuum expectation value of the field ϕ , in the theory described by the generating functional in eq. (5.10), the quantum effective action as a function of ϕ only is needed. So only for the fields ϕ and ϕ^* a substitution in the integrand needs to be made.

Whether the theory is P and CP conserving depends on the complex phase of the vacuum expectation of the ϕ field. So it will be useful to parameterize the expectation value, ϕ_0 , by its absolute value and complex phase separately. This leads to a substitution, in the integrand of the generating functional in eq. (5.23), which can be written down as

$$\begin{aligned} \phi(x) &\rightarrow \phi(x) + \kappa e^{i\delta}, \\ \phi^*(x) &\rightarrow \phi^*(x) + \kappa e^{-i\delta}. \end{aligned} \quad (5.25)$$

This will give the quantum effective action as a function of κ and δ ,

$$\Gamma[\kappa, \delta] = \int_{1\text{PI, connected}} [d\phi][d\phi^*] \left\{ A_0[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}] + \sum_{\nu=1}^{\infty} \left[e^{i\nu\theta} \iint d^4x_1 \cdots d^4x_\nu A_\nu[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1, \dots, x_\nu] G(\phi(x_1) + \kappa e^{i\delta}) \cdots G(\phi(x_\nu) + \kappa e^{i\delta}) \right. \right. \\ \left. \left. + e^{-i\nu\theta} \iint d^4x_1 \cdots d^4x_\nu A_\nu^*[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1, \dots, x_\nu] \times \right. \right. \\ \left. \left. G^*(\phi^*(x_1) + \kappa e^{-i\delta}) \cdots G^*(\phi^*(x_\nu) + \kappa e^{-i\delta}) \right] \right\}. \quad (5.26)$$

The expression for the QEA has the QCD-related part, the path integrals over the gauge and fermion field, hidden in the c_q^ν 's, which are again hidden in the A_ν 's. The problem with these quantities is that we do not have a way to evaluate them. Perturbative QCD is valid only in a high energy regime, but the c_q^ν 's are quantities which include all energy scales; not only the very short distance properties of these functions are needed. The only way to say something over the QEP is to get rid off all but one of these c_q^ν 's. By inspection of eq. (5.19) to (5.21) we see that every c_q^ν comes with a power of $G^{2q+\nu}$. So the QEP expanded to first order in G might still be calculable.

The QEP in eq. (5.26) expanded to 1st order in G is

$$\begin{aligned}
\Gamma[\kappa, \delta] &= \int_{\text{1PI, connected}} [d\phi][d\phi^*] \left\{ A_0^{(1)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}] + \right. \\
&\quad e^{i\theta} \int d^4x_1 A_1^{(0)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] G \left(\phi(x_1) + \kappa(x_1) e^{i\delta(x_1)} \right) + \\
&\quad \left. e^{-i\theta} \int d^4x_1 A_1^{(0)*}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] G^* \left(\phi^*(x_1) + \kappa(x_1) e^{-i\delta(x_1)} \right) \right\} + \mathcal{O}(G^2) \\
&= \int_{\text{1PI, connected}} [d\phi][d\phi^*] \left\{ A_0^{(1)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}] \right. \\
&\quad + \int d^4x_1 \text{Re} \left[|G| \kappa(x_1) A_1^{(0)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] \right] 2 \cos(\theta + \delta(x_1) + \arg(G)) \\
&\quad - \int d^4x_1 \text{Im} \left[|G| \kappa(x_1) A_1^{(0)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] \right] 2 \sin(\theta + \delta(x_1) + \arg(G)) \\
&\quad \left. + 2 \text{Re} \left[\int d^4x_1 e^{i\theta} G \phi(x_1) A_1^{(0)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] \right] \right\} + \mathcal{O}(G^2), \tag{5.27}
\end{aligned}$$

where $A_\nu^{(i)}$ is A_ν to order i . Explicitly writing out the $A_\nu^{(i)}$'s in eq. (5.27),

$$A_0^{(1)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}] = c_0^0 e^{\int d^4x \mathcal{L}_E^\phi(\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta})}, \tag{5.28}$$

$$A_1^{(0)}[\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}, x_1] = c_0^1(x_1) e^{\int d^4x \mathcal{L}_E^\phi(\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta})}, \tag{5.29}$$

confirms that indeed only two c_q^ν 's are left.

Both of these expressions depend on the ϕ part of the Lagrangian density, \mathcal{L}_E^ϕ , as a function of the shifted fields. The ϕ part of the Lagrangian density is defined in eq. (5.18). As a function of the shifted fields it is given by

$$\begin{aligned}
\mathcal{L}_E^\phi(\phi + \kappa e^{i\delta}, \phi^* + \kappa e^{-i\delta}) &= -|\partial_\mu \phi + (\partial_\mu \kappa) e^{i\delta} + \kappa e^{i\delta} (\partial_\mu \delta)|^2 \\
&\quad - \mu^2 (\phi^* + \kappa e^{-i\delta}) (\phi + \kappa e^{i\delta}) - h [(\phi^* + \kappa e^{-i\delta}) (\phi + \kappa e^{i\delta})]^2. \tag{5.30}
\end{aligned}$$

This expression can be simplified by using the fact that κ and δ are x -independent. The path integral has to be performed over the complex ϕ field. This is most easily done by treating it as two real fields $\phi_r = \text{Re}(\phi)$ and $\phi_i = \text{Im}(\phi)$. The Lagrangian density, in terms of these fields, is then given by

$$\begin{aligned}
\mathcal{L}_E^\phi(\phi_r, \phi_i, \kappa, \delta) &= -[\partial_\mu \phi_r]^2 - [\partial_\mu \phi_i]^2 - (\mu^2 + 2h\kappa^2) 2\kappa [\phi_r \cos(\delta) + \phi_i \sin(\delta)] \\
&\quad - h\kappa^2 4 [\phi_r^2 \cos^2(\delta) + \phi_i^2 \sin^2(\delta) + 2\phi_r \phi_i \cos(\delta) \sin(\delta)] - (\mu^2 + 2h\kappa^2) [\phi_r^2 + \phi_i^2] \\
&\quad - 2h\kappa^2 [\phi_r^2 + \phi_i^2] [\phi_r \cos(\delta) + \phi_i \sin(\delta)] - h [\phi_r^4 + \phi_i^4 + 2\phi_r^2 \phi_i^2] - \mu^2 \kappa^2 - h\kappa^4. \tag{5.31}
\end{aligned}$$

The expression for the quantum effective action, eq. (5.27), contains a path integral over the complex ϕ field. This path integral can be expanded in Feynman diagrams with different number of loops. The Feynman rules for these diagrams can be read from the Lagrangian density in eq. (5.31). The quantum

effective action will be evaluated to zeroth order in scalar loops only. To this order, the 4-, 3-, 2- and 1-point vertices can be discarded, since they will either produce loops or be connected to a 1-point vertex, but such a diagram is not 1PI. Also lines between 1-point vertices are not 1PI.

The first term in the quantum effective action in eq. (5.27) is

$$\int_{1\text{PI, Connected}} [d\phi_r][d\phi_i] c_0^0 e^{\int d^4x \mathcal{L}_E^\phi(\phi_r, \phi_i, \kappa, \delta)}. \quad (5.32)$$

The factor c_0^0 , defined in eq. (5.21), can also be written as a diagram expansion. The whole expression in eq. (5.32) needs to be connected. So if c_0^0 is a diagram, then the exponential in eq. (5.32) cannot produce a diagram and so the whole expression does, in that case, not depend on κ or δ . That would just be a shift of the quantum effective potential and thus does not influence the minimum. Therefore, c_0^0 can be taken equal to 1. The only diagram contributing is just the 0-point vertex from the Lagrangian density in eq. (5.31), integrated over space:

$$\int_{1\text{PI, Connected}} [d\phi_r][d\phi_i] c_0^0 e^{\int d^4x \mathcal{L}_E^\phi(\phi_r, \phi_i, \kappa, \delta)} = \int d^4x [-\mu^2 \kappa^2 - h \kappa^4] + \text{loop corrections}. \quad (5.33)$$

Since κ does not depend on the x , the integral is infinite. The infinite factor in front is the volume of 4-dimensional Euclidean space, \mathcal{V}_4 .

The second and third term in the the quantum effective action, eq. (5.27), both contain the expression

$$\int_{1\text{PI, Connected}} [d\phi_r][d\phi_i] \left(\int d^4x_1 c_0^1(x_1) \right) e^{\int d^4x \mathcal{L}_E^\phi(\phi_r, \phi_i, \kappa, \delta)}. \quad (5.34)$$

Diagrammatically it can be written as

$$\text{Diagram} \times \left(1 + \bullet + \bigcirc + \dots \right), \quad (5.35)$$

where the first factor in the diagram is $(\int d^4x_1 c_0^1(x_1))$. In words, it is a 0-point scalar, 2-point fermion vertex, with the two fermion lines connected to an instanton. The second factor in the diagram is the exponential expanded out. It is the sum of all connected, 1PI diagrams, made with scalar lines and vertices. For the computation of the quantum effective action only the connected diagrams need to be kept. Since the second factor cannot be connected to the first, only the first term in the expansion need to be kept, i.e. 1.

The fourth and last term in the quantum effective action contains the expression

$$\int_{1\text{PI, Connected}} [d\phi_r][d\phi_i] \left(\int d^4x_1 c_0^1(x_1) \phi(x_1) \right) e^{\int d^4x \mathcal{L}_E^\phi(\phi_r, \phi_i, \kappa, \delta)}. \quad (5.36)$$

It differs from the previous expression in the sense that the 2-point fermion vertex now also has an attachment point for a scalar line. It needs to be attached to another vertex, coming from the expansion of the exponential. Diagrammatically this looks like

$$\text{Diagram} \text{---} \bullet + \text{Diagram} \text{---} \bigcirc + \dots \quad (5.37)$$

These diagrams are, however, not one-particle irreducible. They, therefore, do not need to be taken into account for the quantum effective action.

The expression for the quantum effective action, eq. (5.27), can now be assembled to be

$$\Gamma[\kappa, \delta] = \mathcal{V}_4(-\mu^2\kappa^2 - h\kappa^4) + \int d^4x_1 \left[c_0^1(x_1) \right]_C 2\kappa|G| \cos(\theta + \delta + \arg(G)) \\ + \text{scalar loop corrections} + \mathcal{O}(G^2), \quad (5.38)$$

where has been used that

$$\left[c_0^1(x_1) \right]_C \equiv \int_{\text{Connected}} [dA_\mu]_1 [d\Psi][d\bar{\Psi}] \left(\bar{\Psi}(x_1) \frac{1 + \gamma_5}{2} \Psi(x_1) \right) e^{\int d^4x \left(-\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi \right)} \quad (5.39)$$

is real. The Lagrangian in the expression for $c_0^1(x_1)$ is translational invariant and thus also the function $c_0^1(x_1)$. Therefore, the integration over x_1 produces a factor \mathcal{V}_4 . With the use of the definition of the quantum effective potential, eq. (5.24), it can be written down as

$$V(\kappa, \delta) = \mu^2\kappa^2 + h\kappa^4 - 2K|G|\kappa \cos(\theta + \delta + \arg(G)) + \text{scalar loop corrections} + \mathcal{O}(G^2), \quad (5.40)$$

in which the constant K is defined by

$$K \equiv \int_{\text{Connected}} [dA_\mu]_1 [d\Psi][d\bar{\Psi}] (\bar{\Psi}_L \Psi_R) e^{\int d^4x \left(-\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i\bar{\Psi} \not{D} \Psi \right)}. \quad (5.41)$$

The quantum effective potential is plotted in figure 5.1 for two different values of $K|G|$ to give an impression of what this potential looks like.

The minimum of the QEP in eq. (5.40) determines the vacuum expectation value of the ϕ field. This minimum will be denoted by ‘barred’ quantities, e.g. $\bar{\kappa}$ and $\bar{\delta}$, the VEV of the ϕ field can then be written as

$$\bar{\phi} = \bar{\kappa} e^{i\bar{\delta}}. \quad (5.42)$$

With this notation, the condition of CP -conservation, eq. (5.16), can be written as

$$\bar{\kappa} = 0 \quad \text{or} \quad \bar{\delta} = -\theta - \arg(G). \quad (5.43)$$

From the form of the QEP in eq. (5.40) or the plot of the potential in figure 5.1, it is immediately clear that, at the minimum of the potential, this condition holds. The conclusion is that this theory is indeed CP conserving; irrespective of the value of θ and $\arg(G)$ the ϕ field will always relax to the CP conserving value.

5.3 Axions in the toy model

A Peccei-Quinn symmetry can be implemented in many different ways. The previous example is just a toy model to show the properties of the Peccei-Quinn mechanism. It also shows nicely the important consequence of a spontaneously broken Peccei-Quinn symmetry: *axions*.

Goldstone’s theorem [22] says that for every spontaneously broken, global, continuous symmetry, the theory contains a massless scalar particle, which are called *Goldstone bosons*. If the symmetry is not exact, but explicitly violated by a ‘small’ term in the Lagrangian, then one can still speak of *pseudo Goldstone bosons*. The pseudo Goldstone bosons have a ‘small’ mass, depending on the amount of explicit symmetry breaking.

The easiest way to calculate the mass of the Goldstone boson, i.e. the axion, is to write the ϕ field in the Lagrangian density not as a real and a complex part, but instead as

$$\phi(x) = \kappa(x) e^{ia(x)/F_a}, \quad (5.44)$$

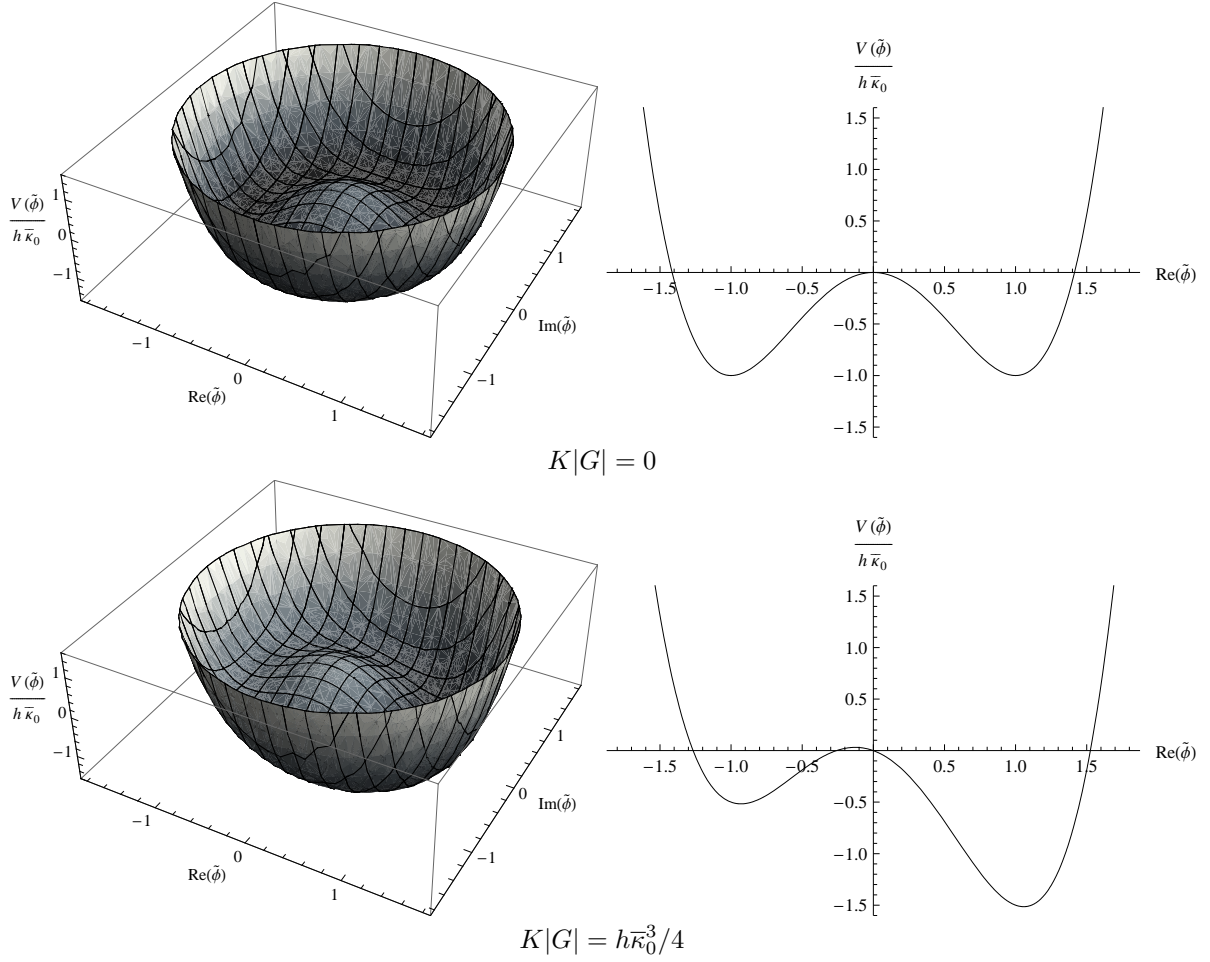


Figure 5.1: Plot of the effective potential in the toy model, eq. (5.40), (to first order in G and no scalar loops) as a function of $\tilde{\phi} = \tilde{\kappa}e^{i\tilde{\delta}}$, where $\tilde{\kappa} = \kappa/\bar{\kappa}_0$ and $\tilde{\delta} = \theta + \delta + \arg G$, in which $\bar{\kappa}_0$ is the value of κ at the minimum of the potential when $G = 0$, i.e. $\bar{\kappa}_0 = \sqrt{-\mu^2/2h}$.

where F_a is an arbitrary constant with units of energy. This gives, for the ϕ part of the Lagrangian density,

$$\mathcal{L}_E^\phi(\kappa, a) = -(\partial_\mu \kappa)(\partial^\mu \kappa) - \frac{\kappa^2}{F_a^2}(\partial_\mu a)(\partial^\mu a) - \mu^2 \kappa^2 - h\kappa^4. \quad (5.45)$$

The functional measure does not transform trivially under this transformation,

$$[d\phi][d\phi^*] \rightarrow \left[\frac{da}{F_a}\right][d\kappa 2\pi\kappa]. \quad (5.46)$$

This transformation, therefore, only works in the limit $\kappa/F_a \rightarrow 1$. In the low energy limit, the field κ will relax to its vacuum expectation value. The minimum of the potential part of the Lagrangian density in eq. (5.45) is, for $\mu^2 < 0$, at a non-zero value of κ ,

$$\bar{\kappa}_0 = \sqrt{\frac{-\mu^2}{2h}}. \quad (5.47)$$

Taking the constant F_a equal to $\bar{\kappa}$, it is seen that the transformation in eq. (5.44) is valid only if the fluctuations of the κ field are small compared to its VEV, $\bar{\kappa}$.

The Lagrangian density can then be written in terms of field fluctuations around the minimum by making the substitution

$$\kappa(x) = \bar{\kappa} + H(x). \quad (5.48)$$

This leads to a Lagrangian density, as a function of H and a ,

$$\mathcal{L}_E^\phi(H, a) = -(\partial_\mu H)(\partial^\mu H) - \left(1 + \frac{2h}{\bar{\kappa}} + \frac{H^2}{\bar{\kappa}^2}\right)(\partial_\mu a)(\partial^\mu a) + 2\mu^2 H^2 - 2\sqrt{-2h\mu^2}h^3 - hH^4. \quad (5.49)$$

In this way a Lagrangian density describing two quantum fields a and H is obtained.

The Lagrangian density in eq. (5.49) describes two particles. One particle corresponds to the H field and has a mass $m_H^2 = -2\mu^2$. The other particle corresponds to the a field, for which there is no bare mass term, only interaction terms with the H field. This last particle is the massless Goldstone boson.

The analysis so far only considered the scalar part of the Lagrangian, \mathcal{L}_E^ϕ . To include quantum effects, one has to make the same substitution as in eq. (5.44), but then in the quantum effective potential instead of the bare Lagrangian. After that, an expansion around the minimum has to be made. The quadratic term then is the mass term for the new fields and the coefficient is the mass squared. The QEP is, however, already given in terms of a radial and angular part. The only things that remain to be done to find the masses of the two bosons, is to expand the QEP, eq. (5.40), around the minimum and make the identification

$$\delta = \frac{a}{F} = \frac{a}{\bar{\kappa}}. \quad (5.50)$$

This then leads to an expression for the masses,

$$\begin{aligned} m_H^2 &= \frac{1}{2} \left. \frac{\partial^2 V(\kappa, \delta)}{\partial \kappa^2} \right|_{\min}, \\ m_a^2 &= \frac{1}{2} \left. \frac{\partial^2 V(\kappa, \delta)}{\partial a^2} \right|_{\min} = \frac{1}{2} \left. \frac{\partial^2 V(\kappa, \delta)}{\partial \delta^2} \right|_{\min} \left(\frac{\partial \delta}{\partial a} \right)^2 = \frac{1}{2\bar{\kappa}^2} \left. \frac{\partial^2 V(\kappa, \delta)}{\partial \delta^2} \right|_{\min}. \end{aligned} \quad (5.51)$$

With the use of the quantum effective potential (first order in G and zeroth order in scalar loops) in eq. (5.40), these can be evaluated to be

$$\begin{aligned} m_H^2 &= -2\mu^2, \\ m_a^2 &= \frac{K|G|}{\bar{\kappa}} = \frac{K|G|}{|\langle \phi \rangle|}. \end{aligned} \quad (5.52)$$

In the expression for the axion mass, eq. (5.52), two important and general features of the axion mass are present. The first feature is that the mass would be zero in the absence of instantons, this is because K would be zero then. So instanton interactions provide the mass of the axion. The second feature is that the mass is inversely proportional to the size of the vacuum expectation value that breaks the Peccei-Quinn symmetry, $|\langle\phi\rangle|$.

5.4 Problems with the standard Peccei-Quinn model

The toy model that was considered in the previous sections shows the main qualitative features of the Peccei-Quinn mechanism. To obtain more quantitative predictions the model should at least include more quarks and also the electroweak interaction.

A more realistic model was proposed in a subsequent paper [5] by Peccei and Quinn. It involves a 4-quark model, in which the electroweak interactions are modeled by a $SU(2)_L \times U(1)_Y$ gauge theory spontaneously broken down to $U(1)_{EM}$. It has, unlike the standard model, two scalar doublets. The Peccei-Quinn symmetry in this model is broken by the vacuum expectation values of the two scalar doublets, which also break the electroweak symmetry.

It was pointed out soon thereafter, by Weinberg [23] and Wilczek [24], that a spontaneously broken Peccei-Quinn symmetry should imply a ‘light’ pseudo Goldstone-boson, which is neutral and pseudoscalar, hence CP -odd. Predictions of the mass and the coupling to known particles of this axion were derived. It was immediately realized that this *standard axion* scenario was very unlikely, because it should have been observed already.

The most accurate prediction of the standard axion mass was done by Weinberg [23]. His analysis includes the mixing of the axion with the neutral, CP -odd π^0 and η^0 mesons. The mass of the standard axion should then be around

$$m_a \simeq \frac{Nm_\pi F_\pi}{2\sqrt{m_u + m_d}} \sqrt{\frac{m_u m_d m_s}{m_u m_d + m_d m_s + m_s m_u}} \frac{1}{v \sin 2\alpha}. \quad (5.53)$$

Here m_u , m_d and m_s are the quark masses, F_π is the pion decay constant, m_π the pion mass, N the number of quark flavors, α an unknown angle relating the VEV’s of the two scalar doublets and v is the scale at which the electroweak symmetry (and also the PQ symmetry) is spontaneously broken. The angle α is defined through

$$|\langle\phi_1\rangle| = v \sin \alpha \quad \text{and} \quad |\langle\phi_2\rangle| = v \cos \alpha. \quad (5.54)$$

The vacuum expectation values of the scalar doublets should not differ too much. Otherwise we are exchanging one fine-tuning problem (the strong CP -problem) with another one. So for numerical estimates $\alpha \approx \pi/4$.

The mass prediction in eq. (5.53) can be evaluated, with the use of quark mass ratios, $m_s/m_d = 20$ and $m_d/m_u = 1.8$, the number of quark flavors in this model $N = 4$, the pion mass $m_\pi = 135\text{MeV}$, the pion decay constant $F_\pi = 92\text{MeV}$ and the electroweak breaking scale $v = 250\text{GeV}$, to be

$$m_a \simeq \frac{m_\pi F_\pi}{v} \frac{1}{\sin 2\alpha} = \frac{50\text{keV}}{\sin 2\alpha}. \quad (5.55)$$

At low energies, where only the light quarks are important, interactions of the axion with ordinary particles should happen mainly through mixing of the axion with the bare π^0 and η^0 , because direct coupling of the axion to the light quarks is suppressed by factors of m_u/v and m_d/v . These bare π^0 and η^0 are the strong interaction eigenstates, which make up almost completely the physical π^0 and η^0 , for which the couplings to ordinary matter are well known. The coupling to ordinary matter for the bare π^0 and η^0 are, due to the very small mixing, to good approximation equal to the couplings of the physical π^0 and η^0 .

Absorption or emission of an axion should then happen with an amplitude of the form $\xi_\pi A_\pi + \xi_\eta A_\eta$, where $A_{\pi,\eta}$ are the amplitudes for absorption or emission of a bare π^0 or η^0 . The coefficients $\xi_{\pi,\eta}$ are

the components of the physical axion along the bare π^0 and η^0 , given [23] for $N = 4$ and $m_s \gg m_{u,d}$ by

$$\begin{aligned}\xi_\pi &= \xi \left[\left(\frac{3m_d - m_u}{m_d + m_u} \right) \tan \alpha - \left(\frac{3m_u - m_d}{m_u + m_d} \right) \cot \alpha \right] = \xi [1.6 \tan \alpha - 0.4 \cot \alpha] \approx 1.2\xi, \\ \xi_\eta &= \xi \left[\sqrt{3} \tan \alpha + \frac{1}{\sqrt{3}} \cot \alpha \right] \approx 2.3\xi, \\ \xi &= \frac{1}{4} \frac{F_\pi}{v} \approx 0.9 \cdot 10^{-4}.\end{aligned}\tag{5.56}$$

The numerical values are obtained with the previously stated values for the quark mass ratios and angle α .

The mixing will give the physical axion e.g. a possibility to decay to two photons, $a_0 \rightarrow 2\gamma$, with a rate of around [23]

$$\Gamma(a_0 \rightarrow 2\gamma) \approx \left(\frac{4N}{3} \right)^2 \left(\frac{m_a}{m_\pi} \right)^3 \xi^2 \Gamma(\pi^0 \rightarrow 2\gamma) \approx 10^4 \left(\frac{m_a}{1\text{MeV}} \right)^3 \text{s}^{-1}.\tag{5.57}$$

Also production rates of axions in the decay of excited nuclear states where derived in this way [25].

Numerous experiments have been done trying to find the standard axion. These are all summarized by the Particle Data Group in [6]. The most prominent exclusions of the standard axion come from nuclear reactor experiments, e.g. [26] and [27], or beam dump experiments, e.g. [28].

5.5 Alternative models with a Peccei-Quinn symmetry

In the standard axion model, the Peccei-Quinn symmetry is broken together with the electroweak symmetry. The mass of the standard axion (eq. (5.55)) and the coupling to ordinary matter (proportional to ξ in eq. (5.56)) are both inversely proportional to the electroweak breaking scale v . This led to the proposal of ‘invisible axion’ models, in which the Peccei-Quinn symmetry is broken at a scale much larger than the electroweak breaking scale.

A spontaneously broken Peccei-Quinn symmetry can be constructed in many different ways, depending on the amount of new particles included in the theory. Most models can be divided into two classes. The first being *hadronic axion models*, in which one or more new (heavy) quarks, carrying a Peccei-Quinn charge, are introduced. The second class models only introduce more new scalar fields. The common feature is that the Peccei-Quinn symmetry is broken at a scale of at least $F_a > 10^9 \text{GeV}$.

The prototype of the hadronic axion models is the KSVZ [29, 30] model. It extends the standard model with a weak-interaction-singlet quark Q and a weak-interaction-singlet, complex Higgs scalar σ , both with zero weak hypercharge. These new particles have a Peccei-Quinn charge, in contrast to all ordinary particles, which are uncharged under the Peccei-Quinn symmetry. The complex Higgs scalar couples to the Q quark only. It has, in the Lagrangian, a potential, such that it develops a very large ($\sim 10^{15} \text{GeV}$) vacuum expectation value. The VEV of the new complex Higgs scalar gives a very large mass to the Q quark and breaks the Peccei-Quinn symmetry.

The prototype of the *non-hadronic axion models* is the DFSZ [31, 32] model. It has two Higgs doublets ϕ_u and ϕ_d , as in the standard axion model. It has, as an extra ingredient, a scalar field ϕ . The two Higgs doublets and the ordinary particles have Peccei-Quinn charges as in the standard axion model. The scalar field ϕ is also charged under the Peccei-Quinn symmetry. This new scalar field does not couple to quarks directly. In the Lagrangian, there is a potential of the form $\lambda(|\phi|^2 - V^2)^2$, such that ϕ develops a very large ($\sim 10^{15} \text{GeV}$) vacuum expectation value, which breaks the Peccei-Quinn symmetry. The ϕ field has a coupling to the two scalar doublets through a $c(\phi_u^i \epsilon_{ij} \phi_d^j \phi^2 + \text{h.c.})$ term.

The general behavior of all axion couplings is inversely proportional to the breaking scale F_A . The exact couplings to electrons, photons, nucleons, etc. can vary heavily from one specific model to the other. Multiple articles, e.g. [33] and [34], have been written which deal with the couplings of completely general axion models, given just the Peccei-Quinn charges of the different fields and their VEV’s.

The ‘invisible’ axion evades all the experimental results, which ruled out the standard axion. New experiments have been set up, trying to find ‘invisible’ axions. Also, limits on couplings and mass have

been derived from stellar evolution and cosmological limits on the amount of dark matter in the universe. A review on the current bounds on the existence of the axion is given in [6]. A summarizing plot of the bounds on the axion mass (or decay constant) for some specific models is shown in Fig. 5.2.

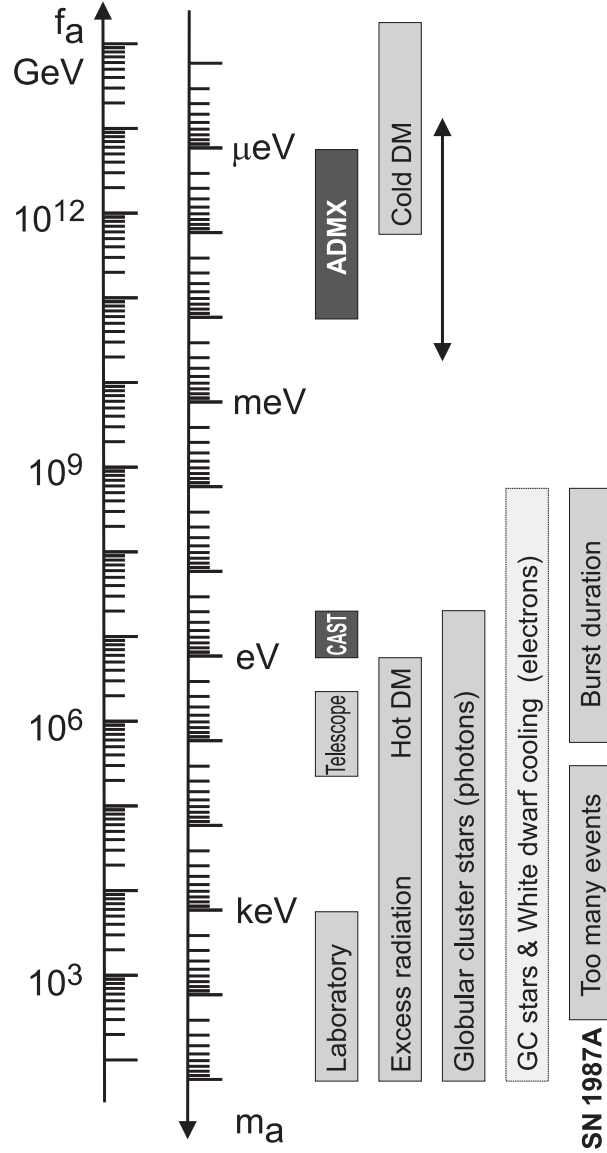
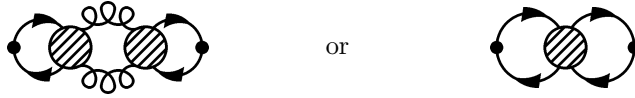


Figure 5.2: Exclusion (light colored) and experimental search ranges (dark colored) on axions. Limits on coupling strengths are translated into limits on m_a and F_a using $m_u/m_d = 0.56$ and the KSVZ values for the coupling strengths. The “Laboratory” bar is a rough representation of the exclusion range for standard or variant axions. The “GC stars and white-dwarf cooling” range uses the DFSZ model. The Cold Dark Matter exclusion range is particularly uncertain. Taken from Particle Data Group, [6].

Chapter 6

Low-energy effective models for QCD

In the last chapter the quantum effective potential for a single-flavor QCD model with a Peccei-Quinn symmetry was calculated (eq. (5.40)). This was done to first order in the fermion-scalar coupling G . The result contained a single unevaluated quantity, K , the vacuum expectation value of the fermion bilinear $\bar{\Psi}_L \Psi_R$ in the presence of a $\nu = 1$ instanton. As noted in section 2.1, a perturbative calculation in QCD can only be done in a high energy limit ($> 1\text{GeV}$). It is for this reason that the constant K , having all momentum scales in it, cannot be evaluated perturbatively. Keeping higher powers of G in the quantum effective potential would have resulted in more vacuum expectation values of strings of fermion operators, i.e. more K -type quantities. These would, for example, include



which cannot be evaluated in perturbative QCD either.

To do any calculations involving the strong interaction at low energies one has to resort to other methods. The two most used methods are *lattice QCD* and *effective models*. Lattice QCD does approximate the full path integral by a discretized one. The discretized path integral can then be evaluated numerically. This is, for useful lattice size, very computational intensive. The method of effective models tries to replace the full QCD theory with a new one which is only valid in a particular approximation regime. Usually this is taken as a low-energy regime, resulting in a low-energy effective model.

At low energies only hadrons and mesons are observed, no quarks or gluons. At energies below the mass of the neutron, only mesons exist. In a quark theory, mesons are considered to be bound states of a quark and an anti-quark. The lowest mass mesons are the neutral pion, π^0 , with mass 135MeV and the charged pions, π^\pm , with mass 140MeV . The neutral π^0 has the quark content $\bar{u}\gamma^5 u - \bar{d}\gamma^5 d$ and the charged pions π^\pm have the content $\bar{d}\gamma^5 u$ and $\bar{u}\gamma^5 d$ respectively. All these mesons do not have angular momentum and are pseudoscalars.

A low-energy effective model only needs to have the low-mass meson degrees of freedom. Next to that, it should have all (approximate) symmetries which the underlying theory has in the low-energy limit. From the meson particle spectrum (see table 6.1) it is seen that the flavor $SU(3)_V$ transformations, as discussed in section 2.4.2, are an approximate symmetry. An even better symmetry is given by the flavor $SU(2)_V$ transformations on the light-quark vector containing u and d . There is, however, no sign of a chiral $SU(2)_L \otimes SU(2)_R$ or a chiral $SU(3)_L \otimes SU(3)_R$ symmetry; such a symmetry could transform the pseudoscalar mesons in their scalar parity partners. The scalar mesons are, however, much heavier than the pseudoscalar mesons.

In the QCD Lagrangian both the $SU(2)_V$ flavor and the chiral $SU(2)_L \otimes SU(2)_R$ symmetry are broken by the quark masses (see section 2.4.2). The breaking of the $SU(2)_V$ flavor symmetry is caused by the non-degeneracy of the quark masses and the chiral $SU(2)_L \otimes SU(2)_R$ symmetry is broken by the

meson	quark content	mass (MeV)
π^-	$\bar{u}\gamma^5 d$	139.5702 ± 0.0004
π^0	$\bar{u}\gamma^5 u - \bar{d}\gamma^5 d$	134.9766 ± 0.0006
π^+	$\bar{d}\gamma^5 u$	139.5702 ± 0.0004
K^-	$\bar{u}\gamma^5 s$	493.68 ± 0.02
K^0	$\bar{s}\gamma^5 d$	497.61 ± 0.02
\bar{K}^0	$\bar{d}\gamma^5 s$	497.61 ± 0.02
K^+	$\bar{s}\gamma^5 u$	493.68 ± 0.02
η^0	$\bar{u}\gamma^5 u + \bar{d}\gamma^5 d - 2\bar{s}\gamma^5 s$	547.85 ± 0.02
η'	$\bar{u}\gamma^5 u + \bar{d}\gamma^5 d + \bar{s}\gamma^5 s$	957.7 ± 0.2
a_0	$\bar{u}u - \bar{d}d$	985 ± 1

Table 6.1: Properties of the light mesons [35]

fact that the quark masses are non-zero. The difference in the up- and down-quark mass is of the same size as the masses themselves and so the explicit breaking of both the flavor and chiral symmetry are equal in strength. This would imply that the mass splitting between the charged and neutral pions is of the same size as, e.g. the mass difference between the neutral pion and its parity partner, the a_0^0 . Since this is definitely not the case, the only option is that chiral symmetry is spontaneously broken. This fact is the basis for the effective models.

For an effective model, it is not necessary to know *how* the chiral symmetry is spontaneously broken. The only thing we need to know is that there is an approximate chiral $SU(3)_L \otimes SU(3)_R$ symmetry which is spontaneously broken to an approximate $SU(3)_V$ symmetry. The effective Lagrangian should then include all terms made out of the meson fields which are compatible with this symmetry breaking structure. Models that include all (scalar and pseudoscalar) meson fields are in general called *linear sigma models*. They were first proposed by Schwinger [36], Polkinghorne [37] and Gell-Mann, Lévy [38].

A low-energy effective model can be made by realizing that the spontaneous breakdown of the chiral symmetry will lead to eight (one for every generator of $SU(3)$) pseudo Goldstone bosons. These Goldstone bosons are light and should, therefore, be the only relevant degrees of freedom in a low-energy theory. By removing the non-Goldstone boson fields from the theory, constraint equations on the remaining fields follow, e.g. $UU^\dagger = 1$. The remaining Goldstone boson fields, consequently, transform under a $SU(2)$ chiral transformation non-linearly, hence the name *non-linear sigma models*.

A non-linear sigma model uses the Goldstone boson fields to build an effective Lagrangian. This Lagrangian should have a chiral invariant part which can, due to the constraint equations, only be constructed out of derivative terms. Also a chiral-symmetry breaking part should be included in the Lagrangian. The terms in that part represent the explicit breaking of the chiral symmetry and should, therefore, be proportional to the quark masses. In principle infinitely many terms can be included in an effective Lagrangian, but since the theory has to be valid only at low energies one can make a low-energy expansion, i.e. an expansion in the number of derivatives on the meson fields. Such an expansion will be in powers of p^2/Λ , where the constant $\Lambda \sim 1\text{GeV}$ sets the energy scale above which the low-energy approximation is not valid anymore. Besides that, a power series expansion in the quark masses can be made. That is an expansion in the ‘departure’ from chiral symmetry and is also valid since $m_{u,d,s} \ll 1\text{GeV}$.

6.1 The $SU(3) \times SU(3)$ non-linear sigma model

A non-linear sigma model can be built without any reference to a quark model. It is instructive, though, to make the connection with the quark model in order to see how the quark mass matrix and the chiral condensate come about in the non-linear sigma model.

The three light quark fields u , d and s can be written in terms of Goldstone-boson fields $\xi_a(x)$ and

Goldstone-boson free quark fields $\tilde{q}(x)$ (see e.g. chapter 19.7 in [16]), through

$$q(x) \equiv \begin{pmatrix} u(x) \\ d(x) \\ s(x) \end{pmatrix} = \exp \left(-i\gamma_5 \sum_a \xi_a(x) \lambda_a \right) \tilde{q}(x). \quad (6.1)$$

Where the λ_a 's are the generators of $SU(3)$ in the defining representation, normalized so that $\text{Tr}(\lambda_a \lambda_b) = 2\delta_{ab}$. The Goldstone-boson fields $\xi_a(x)$ can be expressed in terms of the conventionally normalized pseudoscalar meson fields by writing

$$\sum_a \xi_a \lambda_a = \frac{\sqrt{2}}{F_\pi} \begin{bmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta^0 & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta^0 & K^0 \\ \bar{K}^- & \bar{K}^0 & -\sqrt{\frac{2}{3}}\eta^0 \end{bmatrix}. \quad (6.2)$$

The Lagrangian for the non-linear sigma model is usually written down in terms of a unitary unimodular matrix U , which is related to the ξ_a fields through

$$U(x) \equiv \exp \left(2i \sum_a \xi_a(x) \lambda_a \right). \quad (6.3)$$

Knowing how the quark fields $q(x)$ transform under a chiral $SU(3)_L \otimes SU(3)_R$ rotation, parameterized by θ_a^R and θ_a^L , and demanding that $\tilde{q}(x)$ does not chirally transform, the transformation rule for the matrix U can be determined to be (chapter 19.7 in [16])

$$U'(x) = \exp \left(i \sum_a \lambda_a \theta_a^R \right) U(x) \exp \left(-i \sum_a \lambda_a \theta_a^L \right). \quad (6.4)$$

The unique $SU(3)_L \otimes SU(3)_R$ -invariant term of second order in derivatives is

$$\mathcal{L}_{2\text{deriv}} = -\frac{1}{16} F_\pi^2 \text{Tr} [(\partial_\mu U)(\partial^\mu U^\dagger)], \quad (6.5)$$

where the factor in front is chosen such that the kinetic term for the meson fields is of the usual form

$$\mathcal{L}_{\text{kin}} = -\frac{1}{2} \partial_\mu \pi^0 \partial^\mu \pi^0 + \dots \quad (6.6)$$

The chiral symmetry of QCD is broken by the quark mass term (see section 2.4.2), which can be written in terms of the meson fields and the Goldstone-boson free quark fields as

$$\mathcal{L}_{\text{mass}} = -\bar{q} M_q q = -\bar{\tilde{q}} \exp \left(-i\gamma_5 \sum_a \lambda_a \xi_a \right) M_q \exp \left(-i\gamma_5 \sum_a \lambda_a \xi_a \right) \tilde{q}, \quad (6.7)$$

where the quark-mass matrix

$$M_q = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}. \quad (6.8)$$

In our low energy theory for the mesons, the \tilde{q}_a fields will always stay in their ground state, i.e. only mesons can be produced. The VEV of the \tilde{q}_a fields is defined such that it breaks the chiral symmetry down to only a $SU(3)_V$ flavor symmetry,

$$\langle \tilde{q}_a \tilde{q}_b \rangle = -v \delta_{ab}, \quad \langle \tilde{q}_a \gamma_5 \tilde{q}_b \rangle = 0. \quad (6.9)$$

With the use of these vacuum expectation values the chiral-symmetry-breaking mass term in eq. (6.7) can be written as

$$\mathcal{L}_{\text{mass}} = \frac{1}{2} v \text{Tr} [M_q (U^\dagger + U)], \quad (6.10)$$

leading to the full $SU(3) \otimes SU(3)$ non-linear sigma model Lagrangian, to lowest order in derivatives and quark masses,

$$\mathcal{L} = -\frac{1}{16}F_\pi^2 \text{Tr} [(\partial_\mu U)(\partial^\mu U^\dagger)] + v \text{Tr} [M_q(U^\dagger + U)], \quad (6.11)$$

which is also referred to as the chiral Lagrangian.

6.2 The $U(3) \times U(3)$ non-linear sigma model

In the previous section the $SU(3) \otimes SU(3)$ non-linear sigma model was discussed. It treats the dynamics of the eight light pseudoscalar mesons. These mesons were considered to be pseudo Goldstone-bosons coming from the spontaneous breakdown of the chiral $SU(3)_L \otimes SU(3)_R$ symmetry to $SU(3)_V$.

The global symmetry of the QCD Lagrangian is, in the limit $m_u, m_d, m_s \rightarrow 0$, not only $SU(3)_L \otimes SU(3)_R$, but $U(1)_B \otimes SU(3)_L \otimes SU(3)_R \otimes U(1)_A$ or equivalently $U(3)_L \otimes U(3)_R$. Redoing the whole analysis of the last section, one would conclude that there should be nine light pseudoscalar mesons. In addition to the eight flavored mesons from the previous section, there should also be a flavorless one. In nature an unflavored pseudoscalar meson is observed, the η' , but its mass is much larger than the flavored ones. This problem has been known as the $U(1)$ problem.

The fact that the $U(1)_A$ symmetry is anomalous, is not enough to explain the η' mass; one can also define a non-anomalous $U(1)$ symmetry in the massless quark limit. This non-anomalous $U(1)$ has as the only source of explicit breaking the quark mass matrix, which should imply that the η' mass is of the same order as the other pseudoscalar mesons. The problem was solved by 't Hooft [39], who showed that due to both the axial anomaly and the instanton field solutions, there is no $U(1)$ symmetry even in the limit $m_u, m_d, m_s \rightarrow 0$.

After the formal solution to the $U(1)$ problem, Di Vecchia, Veneziano [40] and Witten [41] derived an effective Lagrangian, which includes the η' , the chiral anomaly and the θ -term of QCD. This Lagrangian only incorporates the instanton effects to leading order in a $1/N$ expansion, where N is the number of colors. The Di Vecchia-Veneziano-Witten effective Lagrangian is, written in terms of a unitary matrix U ,

$$\mathcal{L} = \frac{F_\pi^2}{4} \text{Tr} [(\partial_\mu U)(\partial^\mu U^\dagger)] + \frac{F_\pi^2}{4} \text{Tr} [MU + M^\dagger U^\dagger] - \frac{aF_\pi^2}{4N} [\theta - i \log \det U]^2, \quad (6.12)$$

where M is not directly the quark mass matrix M_q , but some scalar multiplication of it such that it has dimension two and a is an unknown constant specifying the instanton interaction strength. The matrix U can be written in terms of the meson fields, defined in eq. (6.2), by

$$U \equiv \exp \left(i \sum_a \xi_a \lambda_a + \frac{\sqrt{2}\eta'}{\sqrt{3}F_\pi} \right), \quad (6.13)$$

where $\eta'(x)$ is the field corresponding to the flavorless pseudoscalar η' meson.

6.3 Dashen's phenomenon and metastable CP-violating states

It was shown already before the invention of QCD by Dashen [42], that a model of the strong interactions, which is based on the assumption of chiral symmetry breaking, could exhibit spontaneous CP -violation. The possibility of spontaneous CP -violation in the strong interaction has consequently become known as *Dashen's phenomenon*. The method of low-energy effective models opened up a way to investigate the low-energy properties of QCD. It was shown by Di Vecchia, Veneziano [40] and Witten [41], that the low-energy effective models *do* have spontaneous CP -violation, but at $\theta = \pi$.

We know from experiments (see section 4.3) that in nature θ is extremely close to zero. In QCD θ is a parameter and cannot change in time. The possibility of spontaneous CP -violation at $\theta = \pi$ seems therefore rather academic. It is, however, still interesting to see what a world at $\theta = \pi$ would look

like. Besides that, it is not ruled out that QCD should be replaced by some other theory, which has the possibility of a varying θ . In fact, in the next chapter we will look at a non-linear sigma model with a Peccei-Quinn symmetry, such that the θ parameter is replaced by an effective ' θ field', which can change in time.

The θ -dependence of QCD is still subject of investigation. The approach discussed before is still being used by e.g. Smilga [43], Tytgat [44] and Creutz [45]. Different approaches include e.g. the use of the NJL model (introduced in [46, 47]) by Fujihara *et. al.* [48] and Boer and Boomsma [49]. The general conclusion includes that at $\theta = \pi$, at least for some range of parameters, spontaneous CP -violation occurs.

Aside from the spontaneous CP -violation at $\theta = \pi$, the low-energy effective Lagrangian also predicts the existence of metastable (CP -violating) states. These are local minima of the effective potential, into which the field can relax for a finite amount of time. The effective potential in a lowest-order single-flavor ($U(1) \otimes U(1)$) non-linear sigma model is given in [44] as

$$E_\theta(\phi) = -m_u \Sigma \cos(\phi) + \frac{\tau}{2}(\theta - \phi)^2, \quad (6.14)$$

where m_u is the up-quark mass, the constant $\Sigma = |\langle \bar{q}q \rangle|$ defines the absolute value of the chiral condensate in a two-flavor model, hence related to v as defined in eq. (6.9) by $\Sigma = 2v$, τ a constant specifying the 'strength' of the anomaly, usually referred to as the topological susceptibility, and $\phi = \eta'/F_\pi$, where η' is the single Goldstone-boson field in this single-flavor theory. The constant τ can in principle be related to the underlying theory, but here it will be considered as just a parameter which has to be fitted to experiment. The metastable states and the spontaneous CP -violation can be made visible by plotting the effective potential as function of ϕ for different values of θ , see figure 6.1. At $\theta = \pi$ there is no explicit CP violation; the CP transformation $\phi \rightarrow -\phi + 2\pi$ leaves the Lagrangian invariant. There are, however, two degenerate minima of the effective potential, which can be transformed into each other by the CP transformation, i.e. spontaneous CP -violation. At $\theta = 0$ the Lagrangian is also CP invariant, but then under the transformation $\phi \rightarrow -\phi$. The global minimum is there CP invariant. There are, however, also two local minima of the action which are not CP invariant, i.e. there are metastable CP -violating states. At $0 \neq \theta \neq \pi$, there are also local minima, but without definite CP transformation properties, since the Lagrangian is not explicitly CP -conserving.

The same effect appears in the two-flavor ($U(2) \otimes U(2)$) non-linear sigma model. The effective potential to lowest order in the degenerate mass case $m_u = m_d = m$ is given in [44] as

$$E_\theta(\phi, \alpha) = -2m\Sigma \cos\left(\frac{\phi}{2}\right) \cos(\alpha) + \frac{\tau}{2}(\theta - \phi)^2, \quad (6.15)$$

where α is related to the neutral pion field by $\alpha = \pi^0/F_\pi$. This potential is plotted in figure 6.2, where it is seen that the same effects appear as in the single-flavor case. At $\theta = \pi$, there is spontaneous CP -violation and for $\theta \neq \pi$ there is a local minimum of the potential, i.e. a metastable state. Although not visible from the figure, there are also two (weak) metastable states at $\theta = 0$. They disappear for values of $\tau \geq 0.3m\Sigma$.

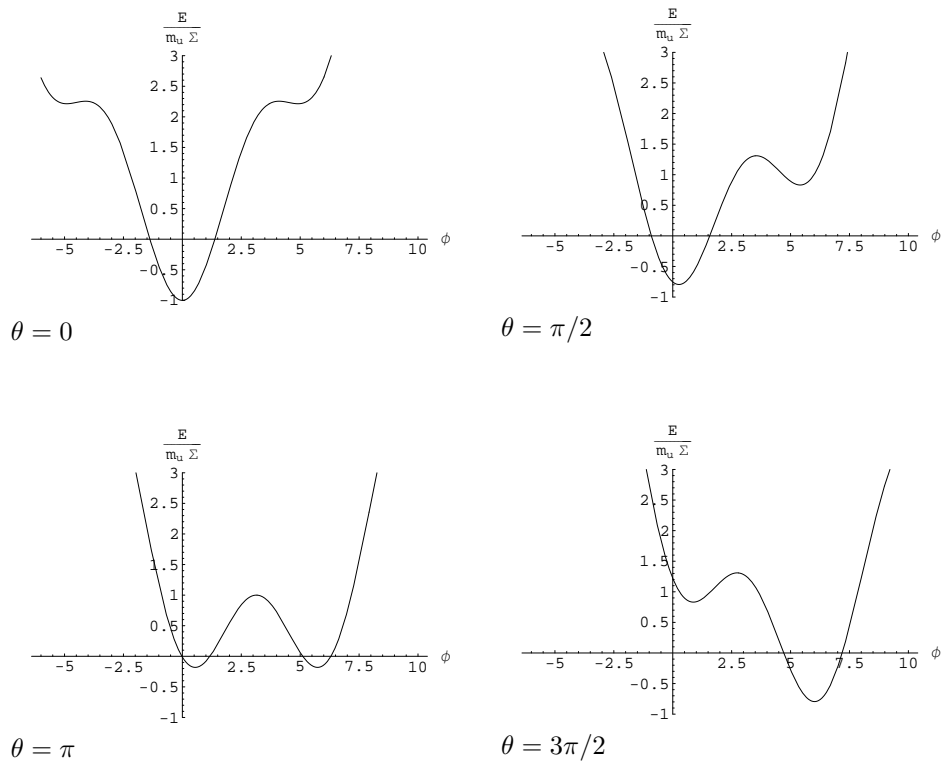


Figure 6.1: Plot of the effective potential in the single-flavor non-linear sigma model to lowest order, for different values of θ and $\tau = 0.2m_u \Sigma$.

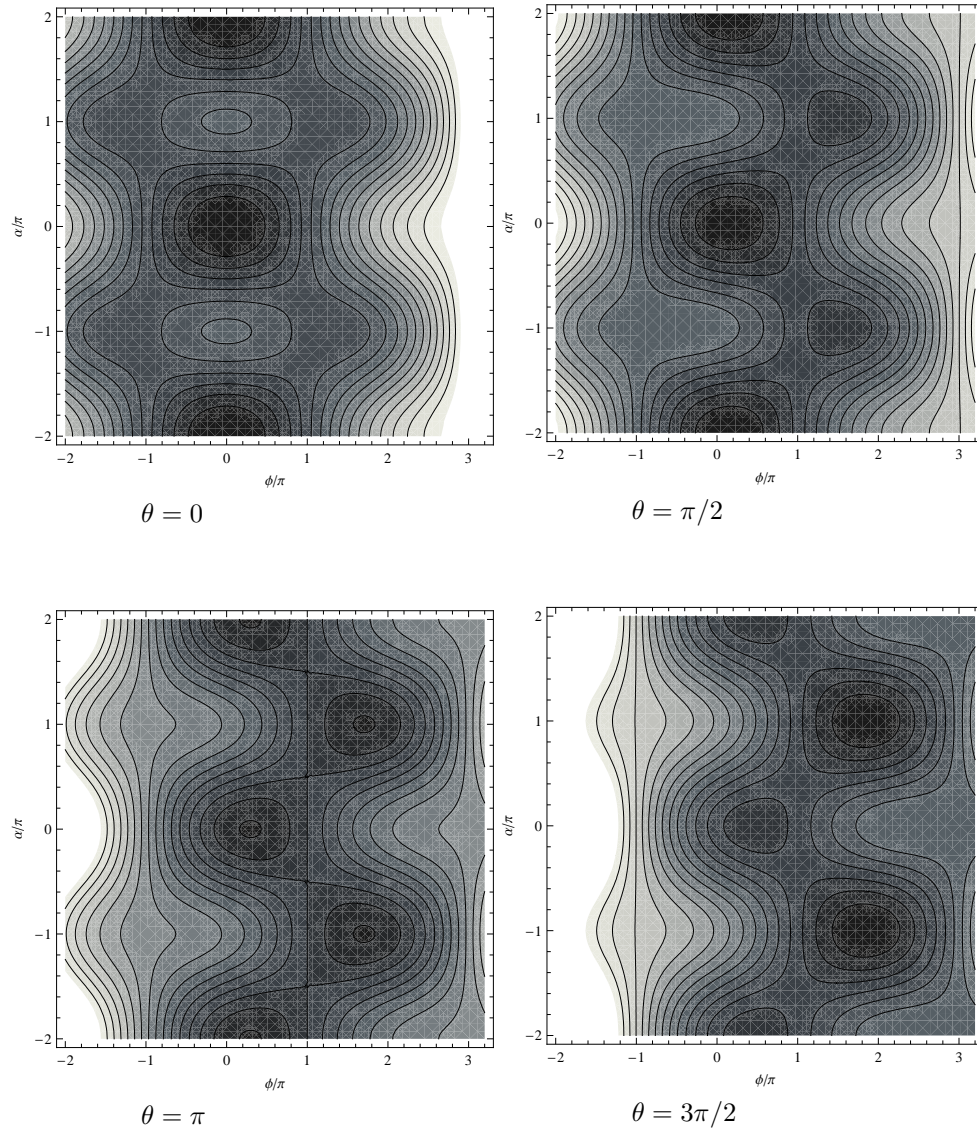


Figure 6.2: Plot of the effective potential $E_\theta(\phi, \alpha)/m\Sigma$ in the two-flavor non-linear sigma model to lowest order, in the degenerate mass case $m_u = m_d = m$ for different values of θ and $\tau = 0.2m_u\Sigma$.

Chapter 7

Non-linear sigma model with a Peccei-Quinn symmetry

In the previous chapter low-energy effective models were introduced as an approximation to QCD. These models predicted some interesting phenomena at $\theta \sim \pi$, such as the possibility of spontaneous CP -violation at $\theta = \pi$ and the possibility of metastable states. These metastable states appeared e.g. at θ around π , but also at $\theta = 0$ for some specific range of τ , leading to metastable CP -violating states.

Only if the correct model for the strong interactions in nature is QCD with a fixed mass matrix, then θ is a parameter and cannot change in time. This is also the case if the mass is generated through the Higgs mechanism of the standard model. Experimentally this value should be very close to zero ($< 10^{-10}$). The low-energy effective models do not predict spontaneous CP -violation in that regime.

It was shown in chapter 5 that the Peccei-Quinn mechanism had the effect of replacing the parameter θ by an effective $\Theta(x)$ field. This opens up the possibility of an effectively varying θ . It would be interesting to see how a low-energy effective model, like the non-linear sigma model, would behave if it had a PQ symmetry. Because the PQ mechanism would allow the effective θ to be around π , one might expect that spontaneous CP -violation could occur. Also the metastable states might be accessible if the effective θ would be allowed to vary.

The questions, which we will try to answer by investigating non-linear sigma models with a Peccei-Quinn symmetry, are: Is the well-established fact of spontaneous CP -violation at $\theta = \pi$ accessible through a PQ symmetry? By accessible we will mean that the theory has either a ground- or metastable state which violates CP -symmetry. The other question will be more general: Does a non-linear sigma model with a PQ symmetry have a ground- or metastable state in which the physical properties, such as e.g. the mass, of the particles are different?

These questions are interesting, because if these states exist they could have experimental consequences, e.g. in heavy-ion collisions. In such a collision a hot ($T > 200\text{MeV}$) quark-gluon plasma is created, which subsequently cools down. In this cooling down process the meson fields and/or axion field could land in a local minimum of the effective potential, i.e. a metastable state, in which it stays for a finite amount of time. If this state is CP -violating, or is in any other way physically different from the true ground state, this could be experimentally observable. A prediction of such states would allow for another way of testing if nature has a PQ symmetry, besides the ongoing search for the axion.

One-flavor quantum chromodynamics with a PQ symmetry was shown to be neither explicitly nor spontaneously CP -symmetry violating. A non-linear sigma model with a PQ symmetry should, being a low energy approximation to QCD, *also* conserve CP -symmetry both explicitly and spontaneously. The absence of spontaneous CP -violation in QCD with a PQ symmetry was derived by showing that, using the same argument as in the Vafa-Witten theorem, the CP -conserving field configuration $\Theta(x) = 0$ is a *global* minimum of the quantum effective potential. So, in the effective theory, the global minimum of the potential should also be a CP -conserving one. The spontaneous CP -violation at $\theta = \pi$ in the effective models is expected, by adding a PQ symmetry, to disappear somehow. The PQ mechanism says, on the other hand, nothing about metastable CP -violating states. It could be that a non-linear sigma model

with a PQ symmetry has metastable CP -violating states.

7.1 Single-flavor non-linear sigma model with a PQ symmetry

In section 6.3 the possibility of spontaneous CP -violation and metastable states was discussed, in specific the single- and dual-flavor case of the non-linear sigma model was discussed. We will now look at a single-flavor model with a Peccei-Quinn mechanism.

The non-linear sigma model that incorporates both the θ -term of QCD and the η' field was discussed in section 6.2. The Lagrangian density of this model was given in eq. (6.12). The matrix M in there can be expressed in terms of the quark-mass matrix by the same line of reasoning as was used in section 6.1. This results in

$$\mathcal{L} = \frac{F_\pi^2}{4} \text{Tr} [(\partial_\mu U)(\partial^\mu U^\dagger)] + 2|v| \text{Re} [\text{Tr}(M_q U^\dagger)] - \frac{\tau}{2} (\theta + i \log \det U)^2, \quad (7.1)$$

where $|v|$ is the size of the chiral condensate, defined as in eq. (6.9), M_q is the quark-mass matrix and τ is the topological susceptibility of the vacuum.

The non-linear sigma model can be given a PQ symmetry by e.g. replacing the mass term by a Yukawa coupling to a scalar field, which develops a vacuum expectation value. This modifies the model as if the underlying quarks get their mass through a Higgs mechanism. Note that this Higgs mechanism is not the same as the one in the standard model. In the standard model three quarks couple to the scalar field ϕ and the other three quarks to the complex conjugate of it, ϕ^* , resulting in no PQ symmetry. The Lagrangian density of the non-linear sigma model to lowest order with a PQ symmetry can thus be written down as

$$\mathcal{L} = \frac{F_\pi^2}{4} \text{Tr} [(\partial_\mu U)(\partial^\mu U^\dagger)] + 2|v| \text{Re} [\text{Tr}(\phi G e^{i\theta/N_f} U^\dagger)] - \frac{\tau}{2} (i \log \det U)^2 + (\partial_\mu \phi)(\partial^\mu \phi^*) - V(|\phi|), \quad (7.2)$$

where G is a Yukawa coupling matrix such that it gives the quarks the right mass.

In the single-flavor case, the unitary matrix U reduces to just a complex number with unit norm. It can be parameterized in terms of the single Goldstone-boson field of this theory, which will be called η' , by

$$U(x) = e^{i\eta'(x)/F_\pi}. \quad (7.3)$$

The field ϕ will be parameterized by just its angular part; the radial part is assumed to be ‘frozen’ in its ground state. This assumption is valid as long as the potential is much steeper in the radial direction than in the angular direction, i.e. the ‘Higgs particle’ in this model is much heavier than the bare axion. The field can thus be written as

$$\phi(x) = \bar{\kappa} e^{ia(x)/F_a} = F_a e^{ia(x)/F_a}, \quad (7.4)$$

where $\bar{\kappa}$ is the norm of the ϕ field at the minimum, $a(x)$ is the bare axion field and the constant F_a will be chosen equal to $\bar{\kappa}$. The Lagrangian density reduces in this case to

$$\mathcal{L} = \frac{1}{4} (\partial_\mu \eta')(\partial^\mu \eta') + (\partial_\mu a)(\partial^\mu a) + m \Sigma \cos \left(\arg G + \frac{a}{F_a} + \theta - \frac{\eta'}{F_\pi} \right) - \frac{\tau}{2} \left(\frac{\eta'}{F_\pi} \right)^2, \quad (7.5)$$

where has been used that $\bar{\kappa}|G| = m$, where m is the mass of the single quark in this theory and $\Sigma = |\langle \bar{q}q \rangle|$ defines the absolute value of the chiral condensate in a two-flavor model, hence related to v as defined in eq. (6.9) by $\Sigma = 2v$.

The CP transformation properties of both fields are defined up to a constant shift. This is because the CP transformation on a fermion or scalar field is defined up to a constant phase rotation. The shifts will be picked such that the Lagrangian is explicitly CP conserving, i.e.

$$\begin{aligned} \eta' &\xrightarrow{P, CP} -\eta' \\ \Theta(x) &\xrightarrow{P, CP} -\Theta(x), \end{aligned} \quad (7.6)$$

where an ‘effective θ -field’, $\Theta(x)$, is introduced, which is defined through

$$\Theta(x) \equiv \arg G + \frac{a(x)}{F_a} + \theta. \quad (7.7)$$

The CP transformation can, of course, also be expressed directly in terms of the a field by

$$\frac{a}{F_a} \xrightarrow{P, CP} -\frac{a}{F_a} - 2 \arg G - 2\theta. \quad (7.8)$$

The quantum effective potential, to tree level, can directly be read off from the Lagrangian; it is just the Lagrangian density without the derivative terms. This potential is plotted in figure 7.1. From the plot and the Lagrangian density in eq. (7.5) it is clear that there are *no local minima* and that the global minima are at

$$\Theta = 0 \mod 2\pi \quad \text{and} \quad \eta' = 0, \quad (7.9)$$

which are CP conserving.

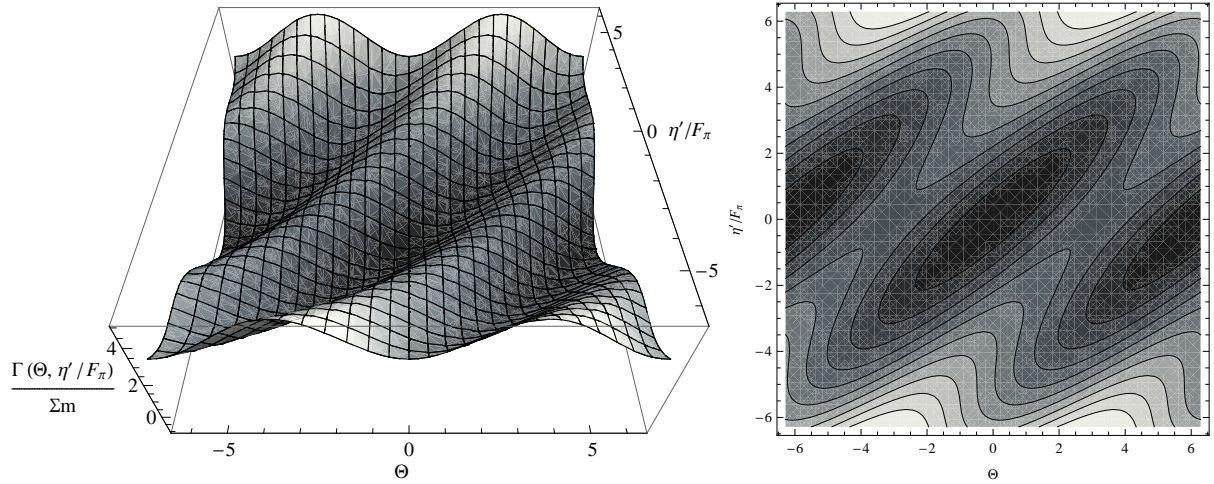


Figure 7.1: The leading order potential for the single-flavor non-linear sigma model with a Peccei-Quinn symmetry, having a topological susceptibility of $\tau = 0.2m_u\Sigma$.

It is also clear from the plot *how* the spontaneous CP -violation at $\theta = \pi$ disappears. If the a field is not there, i.e. no PQ symmetry, then the Θ field degrades from a field to a parameter. It is then nothing but the θ parameter in eq. (6.14). At a fixed $\Theta = \pi$ the potential in this model has two degenerate CP -violating minima at non-zero η' , but when Θ is allowed to vary the field will relax to a global minimum at $\Theta = 0 \mod 2\pi$ and $\eta' = 0$. This can also be made clear by plotting the QEP as a function of just Θ , where for η' the value is chosen which minimizes the potential, this is done in figure 7.2.

7.2 Next-to-leading order non-linear sigma model with a PQ symmetry

The next-to-leading order (in quark masses and derivatives) $SU(3) \otimes SU(3)$, i.e. three-flavor, non-linear sigma model is given in e.g. [50] or chapter 19.7 in [16]. The Lagrangian density can be written as the lowest order contribution, given in eq. (6.11), plus an extra contribution, \mathcal{L}_4 , which is fourth order in derivatives and meson masses (quark masses are considered to be second order in meson mass). For the quantum effective action only the non-derivative part is needed, which can be written as

$$\begin{aligned} \mathcal{L}_4 = 4 \frac{\Sigma^2}{F_\pi^2} & \left(L_6 \left(\text{Tr} [M_q^\dagger U + M_q U^\dagger] \right)^2 + L_7 \left(\text{Tr} [M_q^\dagger U - M_q U^\dagger] \right)^2 + L_8 \text{Tr} \left[(U M_q^\dagger)^2 + (U^\dagger M_q)^2 \right] \right) \\ & + \text{deriv. terms}, \end{aligned} \quad (7.10)$$

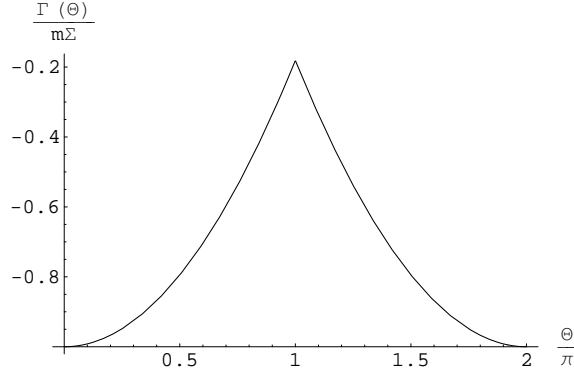


Figure 7.2: The leading-order potential for the single-flavor non-linear sigma model with a Peccei-Quinn symmetry, having a topological susceptibility of $\tau = 0.2m_u\Sigma$, plotted as a function of just Θ , where for η' the value is chosen which minimizes the potential.

where $L_{6,7,8}$ are dimensionless constants to be determined by comparison with experiment. Limiting ourselves again to the single-flavor case and adding a PQ symmetry, by using

$$U = e^{i\eta'/F_\pi} \quad M_q = |G|F_a e^{i(a/F_a + \arg G + \theta)} \equiv m e^{i\Theta}, \quad (7.11)$$

the Lagrangian density reduces to

$$\mathcal{L}_4 = 8 \left(\frac{m\Sigma}{F_\pi^2} \right)^2 (L_6 - L_7 + L_8) \cos \left(2[\Theta - \eta'/F_\pi] \right). \quad (7.12)$$

To calculate amplitudes at next-to-leading order in the non-linear sigma model, all diagrams up to one-loop from the lowest order Lagrangian should be included plus all tree-level diagrams made with vertices from the next-to-leading order Lagrangian. One-loop diagrams are suppressed by at least a power of p^2 or m^2 (see e.g. [51]). We want to calculate the quantum effective potential for a space-time independent background field. The loop diagrams, which are suppressed by a power of p^2 , will always enter with two derivatives on the background field. They are consequently not important for a *constant* background-field. The diagrams that do not come with a p^2 suppression are proportional to $m^{2,4}$, but all terms which obey the symmetries and are proportional to $m^{2,4}$ are already in the Lagrangian. Those loop diagrams do nothing but renormalizing the terms that are already present in the Lagrangian. So the effect of those loop diagrams can be taken into account by using appropriately renormalized values of the coupling constants.

The quantum effective potential at next-to-leading order in the non-linear sigma model is just the potential part of the Lagrangian density in eq. (7.5) plus the one in eq. (7.12),

$$\frac{\Gamma(\Theta, \eta')}{m\Sigma} = -\cos(\Theta - \eta'/F_\pi) + \frac{\tau}{2m\Sigma} (\eta'/F_\pi)^2 - \frac{8m\Sigma}{F_\pi^4} (L_6 - L_7 + L_8) \cos[2(\Theta - \eta'/F_\pi)]. \quad (7.13)$$

The QEP is plotted in figure 7.3 for a specific choice of τ and the constant $L_6 - L_7 + L_8$. It displays a metastable state at $\Theta = \pi$, $\eta' = 0$.

The derivatives of the QEP in eq. (7.13) at the point $\Theta = \pi$, $\eta' = 0$ are zero in both directions. It is therefore a stationary point, but not necessarily a local minimum, it can also be a saddle point. By considering the second-order derivatives, it is found that if

$$\frac{8(L_6 - L_7 + L_8)m\Sigma}{F_\pi^4} > \frac{1}{4}, \quad (7.14)$$

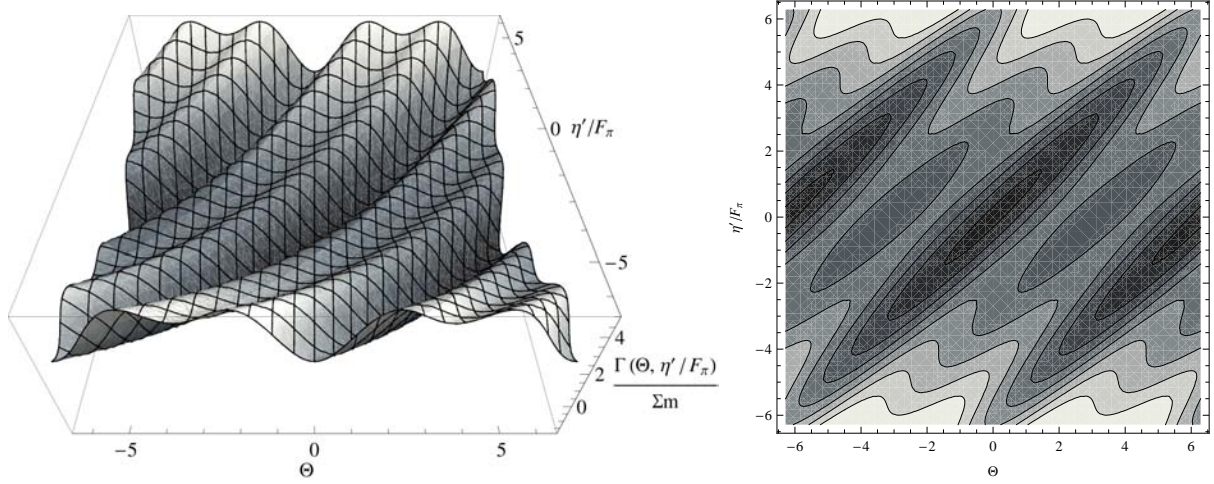


Figure 7.3: The next-to-leading order potential for the single-flavor non-linear sigma model with a Peccei-Quinn symmetry, having a topological susceptibility of $\tau = 0.2m_u\Sigma$ and $8(L_6 - L_7 + L_8)m\Sigma/F_\pi^4 = 0.75$.

then the point is a local minimum.

The masses for the field fluctuations around both the global and local minimum can be calculated by diagonalizing the matrix of second-order derivatives, i.e.

$$m_a^2, m_{\eta'}^2 = \text{Eigenvalues of} \begin{pmatrix} \frac{1}{2} \frac{\partial^2 \Gamma}{\partial a^2} & \frac{1}{2} \frac{\partial^2 \Gamma}{\partial a \partial \eta'} \\ 0 & \frac{1}{2} \frac{\partial^2 \Gamma}{\partial \eta'^2} \end{pmatrix}, \quad (7.15)$$

where the derivatives are evaluated at the global or local minimum. This leads to the masses

$$\begin{aligned} \text{at } \Theta = 0 \quad m_a^2 &= \frac{m\Sigma}{2F_a^2} + \frac{16(L_6 - L_7 + L_8)m^2\Sigma^2}{F_a^2 F_\pi^4} \\ m_{\eta'}^2 &= \frac{m\Sigma}{2F_\pi^2} + \frac{16(L_6 - L_7 + L_8)m^2\Sigma^2}{F_\pi^6} + \frac{\tau}{F_\pi^2} \end{aligned} \quad (7.16)$$

and

$$\begin{aligned} \text{at } \Theta = \pi \quad m_a^2 &= -\frac{m\Sigma}{2F_a^2} + \frac{16(L_6 - L_7 + L_8)m^2\Sigma^2}{F_a^2 F_\pi^4} \\ m_{\eta'}^2 &= -\frac{m\Sigma}{2F_\pi^2} + \frac{16(L_6 - L_7 + L_8)m^2\Sigma^2}{F_\pi^6} + \frac{\tau}{F_\pi^2}. \end{aligned} \quad (7.17)$$

From this it can be concluded that the metastable state differs from the ground state; the particles have different mass if the vacuum ‘hangs’ (temporarily) in the metastable state at $\Theta = \pi, \eta' = 0$.

The most important question concerning these metastable states is whether it is reasonable to assume that condition 7.14 holds. All constants describing this low-energy effective model should be obtained by fitting predictions to experiments. Since the single-flavor toy-model considered here is unrealistic, as we know that there are two light quarks, such a fit to experiments cannot be made. To do such a fit one could work with a two- or even three-flavor non-linear sigma model and add a Peccei-Quinn symmetry in some way.

More can, however, be said about the single-flavor model. From the expression for the masses at $\Theta = 0$, it can be seen that the condition for a local minimum implies that the second-order correction (in the quark mass m) to the mass of the axion and the meson is larger than the lowest-order contribution. The metastable state appears thus only in a regime where this model is not to be trusted; the perturbative expansion in the quark mass seems not to converge. The fact that the model is not to be trusted does, however, not imply that the metastable state is not there in QCD.

Can we tell what the parameters of the underlying theory should be to enter this regime? If the condition 7.14 holds, then there is no metastable state, but whether this condition holds depends completely on the unknown constant $L \equiv L_6 - L_7 + L_8$. This constant cannot easily be related to the ‘fundamental’ constants, which describe the underlying theory (QCD).

We *do* know, however, that the condition 7.14 corresponds with a bad convergence of the expansion in the quark mass. This expansion in the quark mass in the non-linear sigma model depends on the size of the explicit chiral-symmetry breaking compared to the spontaneous chiral-symmetry breaking. So if one wants find a metastable state, then it should be looked for in the limit of large explicit chiral-symmetry breaking, i.e. a large quark mass, or in the limit of small spontaneous breaking. Sufficiently small breaking might happen at a high temperature, near the chiral transition temperature, above which the spontaneous chiral-symmetry breaking is restored.

In conclusion, the single-flavor non-linear sigma model excludes metastable states in the regime of small $Lm\Sigma/F_\pi^4$. On the other hand, for large enough $Lm\Sigma/F_\pi^4$ there might be a metastable state. The parameter $Lm\Sigma/F_\pi^4$ is, however, hard to relate to the parameters of a underlying QCD theory, but we *do* know that this parameter is exactly the measure of how well the non-linear sigma model works and so if there are metastable states, they are in a regime where the non-linear sigma model is not applicable. This might be e.g. when the single quark in this model is heavy, the spontaneous breaking of the chiral symmetry is small, which could be e.g. around the critical temperature, above which the chiral symmetry breaking is restored.

Chapter 8

Discussion

In the previous chapter the single-flavor non-linear sigma model was equipped with a Peccei-Quinn symmetry by replacing the quark mass term with a Yukawa coupling to a scalar field. The scalar field was given a potential such that it develops a non-zero vacuum expectation value, giving a mass to the quark. The result was an explicitly CP -conserving theory, irrespective of the value of the θ parameter and the complex phase of the Yukawa coupling, $\arg G$. It was also shown that the global minimum of quantum effective potential, at next-to-leading order, of this model was CP -conserving. So the model does not have spontaneous CP -violation. This is in accordance with the whole idea behind the Peccei-Quinn mechanism: explaining why there is no CP violation observed in the strong interaction even though quantum chromodynamics could in general be CP -violating. It is also in accordance with the fact that a PQ mechanism changes the θ parameter into an effective Θ field, for which the lowest energy state is the CP -conserving $\Theta = 0$ state, due to the (all-order) Vafa-Witten theorem.

The Vafa-Witten theorem applied to the energy of the θ -vacua does, however, not exclude metastable (possibly CP -violating) states. The next-to-leading order single-flavor non-linear sigma model with a PQ symmetry does, for a specific choice of parameters, have a metastable although not CP -violating state. This metastable state is different from the groundstate in the sense that if the vacuum ‘hangs’ in this state, then the excitations of the vacuum, i.e. the particles, have different mass. This metastable state might have observable consequences in heavy-ion collisions.

The parameter, which determines whether a metastable state occurs or not, cannot be easily expressed in terms of the parameters of the underlying QCD theory. However, this parameter happens to be exactly a measure of how well the expansion in the quark mass in non-linear sigma model converges. This gives an indication in which situation to expect the metastable state. There where the non-linear sigma model is not very well applicable, there could be a metastable state. Whenever the quark-mass expansion in the non-linear sigma model converges fast, there is no metastable state.

The non-linear sigma model is based on the spontaneous breaking of chiral symmetry. It also includes the explicit breaking of chiral symmetry. This explicit breaking should be small in comparison to the spontaneous breaking for the model to work. So this gives us some handles on when a metastable state could occur. This might be in the limit that the quark is heavy, so the explicit breaking is large, or in the limit of small spontaneous breaking, which could occur e.g. at a high temperature, near the chiral symmetry restoration transition. To make predictions in these regimes, one should, however, resort to other models, such as e.g. the linear sigma model or the NJL model. We will discuss the expectations for these models in section 8.2.

8.1 Applicability of a single-flavor model

Another important question is what use is a *single*-flavor model. As we know, multiple light quarks exist; if one wants to describe the physical unflavored η' , one should at least include all other observed particles up to its mass, so all pions and kaons. Therefore, the complete three-flavor $U(3) \otimes U(3)$ non-linear sigma model is the only one which could be well fitted to experiments.

But what low-energy model should one use if the underlying theory has only one heavy quark coupling to the axion, while all light quarks do not? Such a model is e.g. the – not ruled out – KSVZ axion model, discussed in section 5.5. One could argue that, since the axion does not couple directly to the light quarks, a single-flavor model will do the job of describing the axion dynamics. And then, apart from that, another multi-flavor model (without a PQ symmetry) will describe the light-meson dynamics. This does not seem right. Both the light quarks and the heavy quark interact through gluons. One can thus not describe the heavy and light quarks with two different models. One should work with one low-energy theory, obtained by integrating out the high-energy degrees of freedom. Integrating out the high-energy degrees of freedom does not destroy a symmetry of the theory. The PQ symmetry is, however, not exact and the amount of explicit breaking might change due to the removal of the high-energy degrees of freedom. But we know that the PQ symmetry is broken by the axial anomaly and a PQ transformation under an angle α changes the Lagrangian density by an amount $-\alpha g^2 \text{Tr}[F\tilde{F}]/32\pi^2$. This does not depend on the mass of the quark, which is charged under the PQ transformation. It only depends on the strength of the gauge interaction, g . So it is not important how the PQ symmetry is implemented in the high-energy model, the low-energy effective model will always inherit the PQ symmetry, with unaltered strength.

The applicability of a single-flavor model is therefore rather limited. One has to use a multi-flavor model with a PQ symmetry to describe the low-energy dynamics. One may, however, expect that some properties of the single-flavor model are carried over to a multi-flavor model. Just like the single- and two-flavor non-linear sigma model, discussed in section 6.3, have both the spontaneous CP -violation at $\theta = \pi$ and the metastable states, it might be expected that the possibility of a metastable state is also present in a multi-flavor model. To make sure of this, actual calculations have to be performed of course.

8.2 Expectations for other models

To investigate further whether the addition of a PQ symmetry to a QCD theory causes the appearance of a metastable state with properties different from the true ground state, other models than the non-linear sigma model have to be used. The metastable state might occur in a situation where the quark masses are heavy enough compared to the spontaneous breaking to break down the non-linear sigma model. As discussed in the previous section, the low-energy effective model should include just the lightest quarks and not necessarily the quarks to which the axion couples. For the lightest quarks we know that the non-linear sigma model works well and so at low energies, for realistic quark masses and spontaneous breaking, we do not expect a metastable state.

There is, however, a way to lower the amount of spontaneous symmetry breaking, namely by increasing the temperature. The non-linear sigma, having a fixed amount of spontaneous chiral symmetry breaking, is not able to describe a high temperature situation. There are effective models that can describe the dynamics near the chiral phase transition ($\sim 200\text{MeV}$). These are the linear sigma model and the NJL model.

A calculation of the effective potential in the linear sigma model or NJL model with a PQ symmetry for different temperatures goes beyond the scope of this thesis. But we have learned that a PQ symmetry in the underlying theory has as only effect in the low-energy effective theory that the θ parameter changes into a field, i.e. it becomes dynamic. So if in those models the dependence on the temperature and the θ parameter is known, then the effect of a PQ symmetry is trivially deduced. The ground and metastable states are given by a minimization of not only the meson fields at fixed θ , but one also has to minimize in the θ direction.

The θ dependence of the thermal effective potential in the linear sigma model is investigated by e.g. Mizher and Fraga in [52]. Their result of the potential around the chiral transition temperature for $\theta = 0, \pi$ is plotted in figure 8.1. The linear sigma model also includes the scalar mesons, in contradiction to the non-linear sigma model. Looking at the potential below the transition temperature (the left pictures) we see a very clear resemblance to the potential of the non-linear sigma model *with* the metastable state, figure 7.3. The global CP -violating minima are at $\eta = \pm C$, where C is just a constant with in the middle, at $\eta = 0$, a CP -conserving metastable state. This is the same as in figure 7.3 at constant $\Theta = \pi$. At $\theta = 0$ there is just one global minimum, which is lower than the two at $\theta = \pi$, exactly as in figure

7.3. Having only two sample points, at $\theta = 0, \pi$, where both theories agree does not imply that they are the same at other values of θ . But it is most likely that, just as in the next-to-leading order non-linear sigma mode, by lowering θ from π one of the two global minima becomes the global minimum and by raising θ the other. The most important question is, of course, what the metastable state does under a change in θ . If it disappears, as in figure 7.3, then it means that it is also a metastable state if θ is allowed to vary. When it merges with one of the two other minima, then by making θ dynamic it is not a metastable state anymore. It would seem strange if the global minimum tends to $\eta = 0$ when $\theta \rightarrow 0$, but the local minimum moves away from $\eta = 0$ when $\theta \rightarrow 0$. Plots for other values of θ are needed to give a definitive answer.

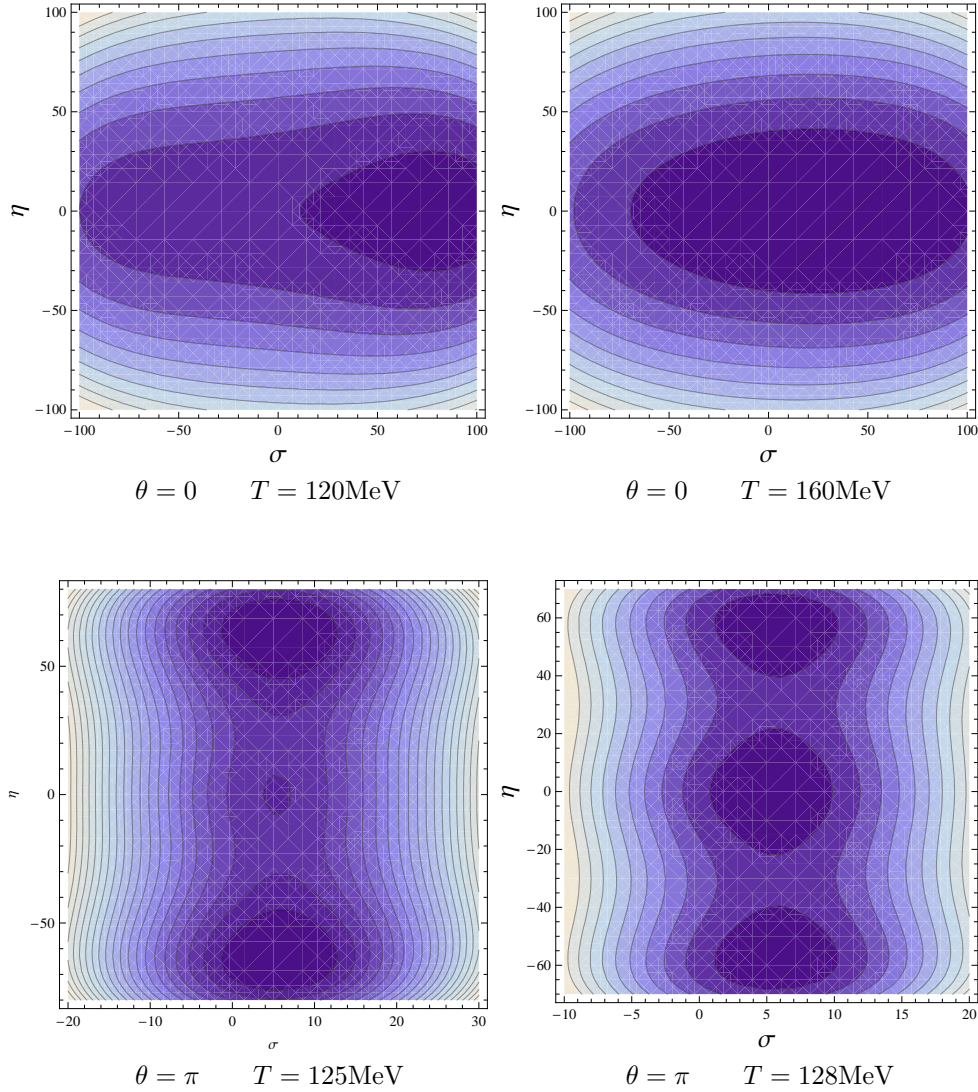


Figure 8.1: Contourplot of the thermal effective potential as a function of the unflavored η (pseudoscalar) and σ (scalar) field in the dual-flavor linear sigma model. At $\theta = \pi$, for a sufficiently high temperature a metastable state is formed before the global minimum changes to the chiral restored $\eta = 0$, $\sigma = 0$ state, i.e. a first-order phase transition. At $\theta = 0$ no such metastable state is formed and the transition is a cross-over. Taken from [52].

It seems, in any case, that adding a PQ symmetry leads to the creation of a CP -conserving metastable state at $\Theta = \pi$ for temperatures just below the transition temperature. This might have observable

consequences in high-energy heavy-ion collisions. If an accurate prediction could be made, then this would open up another way to test whether nature has a PQ symmetry or not, *irrespective* of the precise way in which the PQ symmetry is implemented. It is not clear if these effects could be observed, but if they are, then they can confirm or rule out axion models in general, also the ‘invisible’ axion models.

The fact that, for fixed $\theta = \pi$ a metastable state appears in the linear sigma model, is related to the order of the chiral phase transition. In the linear sigma model this is a first order transition, which always comes about with first a metastable state, which then, by increasing e.g. the temperature becomes the global minimum.

The NJL model, investigated by Boer and Boomsma [49], predicts a different type of phase transition at fixed $\theta = \pi$. It predicts a second-order phase transition, which implies there will not be a metastable state at $\theta = \pi$, $\eta = 0$. This would then imply that, even near the chiral transition, a PQ symmetry would not cause a metastable state at $\Theta = \pi$. It is not clear why these models differ in their predictions.

Chapter 9

Conclusion

There is a general understanding that in QCD there is a possibility of spontaneous CP -violation and metastable (CP -violating or not) states around $\theta = \pi$. Fact is that θ is a parameter and it is experimentally determined to be very close to zero ($< 10^{-10}$). Adding a Peccei-Quinn symmetry to the model describing the strong interactions, is a – not yet ruled out – way to explain why there is no CP violation in the strong interaction, i.e. why θ is so small. It does this by effectively making θ a field, i.e. dynamic, and showing that it will relax to the CP -conserving value $\theta = 0$.

The question arises whether an effectively varying θ opens up a way of accessing the previously discussed phenomena around $\theta = \pi$. In this thesis this is investigated with the use of a low-energy approximation to QCD, the single-flavor non-linear sigma model, equipped with a PQ symmetry. It is shown that this model is indeed explicitly CP -conserving and, to next-to-leading order, it is shown that the ground state of the model does not violate CP . The model does thus not have spontaneous CP -violation, in accordance with what the Peccei-Quinn mechanism is supposed to do.

The possibility of metastable states is, at next-to-leading order, ruled out in the case of realistic quark masses and strength of the spontaneous chiral symmetry breaking. However, in the limit of small spontaneous chiral symmetry breaking, which can be caused by e.g. a high temperature, the model is not trustworthy, but it seems that a CP -conserving metastable state at $\theta = \pi$, with properties different from the ground state, is formed in that case. Other models, which are valid in the regime of the chiral transition, such as the linear sigma model and the NJL model should be used to investigate this further. The results presented in the literature suggest that the linear sigma model will predict a metastable state, whereas the NJL model will not.

Appendix A

Conventions used

Units

Natural units are used throughout this thesis, where

$$\hbar = c = k_B = 1, \quad (\text{A.1})$$

such that

$$[\text{energy}] = [\text{mass}] = [\text{temperature}] = [\text{length}]^{-1} = [\text{time}]^{-1}. \quad (\text{A.2})$$

Wick rotation

A Wick rotation transforms a theory defined in Minkowski space to one defined in Euclidean space. We write objects in Minkowski space without any special label and Euclidean objects with an additional E sub- or super-script. We demand that

$$x_\mu x^\mu = -x_\mu^E x_E^\mu, \quad \text{where} \quad g^{\mu\nu} = \text{diag}(1, -1, -1, -1) \quad \text{and} \quad g_E^{\mu\nu} = \text{diag}(1, 1, 1, 1). \quad (\text{A.3})$$

A substitution certainly satisfying the above equation is

$$x_0^E = ix_0, \quad x_i^E = x_i, \quad (\text{A.4})$$

which will be taken as the transformation of a 4-vector under a Wick rotation.

Most of the time we will not bother writing Euclidean labels on 4-vectors, only on Lagrangians or Actions. A Wick rotation can then be specified in a transformation style, transforming the variables to Euclidean space. An inverse Wick rotation then transforms Euclidean variables back to Minkowski space.

So under a Wick rotation (W.r.) the zero-component of a 4-vector receives a factor $-i$:

$$x_0 \xrightarrow{\text{W.r.}} -ix_0, \quad \text{because} \quad x_0 = -ix_0^E. \quad (\text{A.5})$$

And under an inverse Wick rotation the zero-component receives a factor i :

$$x_0 \xrightarrow{\text{inv.W.r.}} ix_0, \quad \text{because} \quad x_0^E = ix_0. \quad (\text{A.6})$$

In this way, inner products are straightforwardly assigned a transformation, in both a forward and a backward Wick rotation these are

$$\begin{aligned} x_\mu x^\mu &\rightarrow -x_\mu x^\mu, \\ F_{\mu\nu} F^{\mu\nu} &\rightarrow F_{\mu\nu} F^{\mu\nu}. \end{aligned} \quad (\text{A.7})$$

Combinations including the Levi-Civita symbol transform like

$$\begin{aligned}\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} &\xrightarrow{\text{W.r.}} -i\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}, \\ \epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} &\xrightarrow{\text{inv.W.r.}} i\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}.\end{aligned}\tag{A.8}$$

Integration limits do not change under a Wick rotation. We want x_μ^E to be real in order to obtain a positive Euclidean distance between two arbitrary points, so the integration over x_0^E is still from $-\infty$ to ∞ . The actual path taken in x_0 space does thus change 90° anti-clockwise. This does not only hold for spatial coordinates but also for e.g. the $F_{\mu\nu}^a$ fields.

The Euclidean action, S_E , is defined such that, under a wick rotation,

$$e^{iS} \xrightarrow{\text{W.r.}} e^{-S_E}.\tag{A.9}$$

It implies that the Euclidean action is obtained by Wick rotating $-i$ times the Minkowskian action:

$$-iS \xrightarrow{\text{W.r.}} S_E.\tag{A.10}$$

So if the Minkowskian action, S , is defined in terms of the Minkowskian Lagrangian density, \mathcal{L} , through

$$S = \int d^4x \mathcal{L},\tag{A.11}$$

then the Euclidean action, S_E , is given in terms of the Euclidean Lagrangian density, \mathcal{L}_E , by

$$S_E = - \int d^4x \mathcal{L}_E,\tag{A.12}$$

in which \mathcal{L}_E is the Wick rotation of \mathcal{L} .

Bibliography

- [1] A. D. Sakharov: *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32 [JETP Lett. **5** (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991) 24].
- [2] C. Amsler *et al.*: *Baryons*. Physics Letters B **667**, 1061 (2008).
- [3] R.J. Crewther, P. Di Vecchia, G. Veneziano, E. Witten: *Chiral estimate of the electric dipole moment of the neutron in quantum chromodynamics*. Physics Letters B **88**, 123 (1979).
- [4] R.D. Peccei, Helen R. Quinn: *CP Conservation in the Presence of Pseudoparticles*. Physical Review Letters **38**, 1440 (1977).
- [5] R.D. Peccei, Helen R. Quinn: *Constraints imposed by CP conservation in the presence of pseudoparticles*. Physical Review D **19**, 1791 (1977).
- [6] C. Amsler *et al.*: *Gauge and Higgs Bosons*. Physics Letters B **667**, 385 (2008).
- [7] Michael E. Peskin, Daniel V. Schroeder: *Introduction to quantum field theory*, Westview Press.
- [8] Sidney Coleman: *Aspects of symmetry: selected Erice Lectures of Sidney Coleman*, Cambridge University Press, Cambridge.
- [9] A.A. Belavin, A.M. Polyakov, A.S. Schwartz, Yu.S. Tyupkin: *Pseudoparticle solutions of the Yang-Mills equations*. Physics Letters B **59**, 85 (1975).
- [10] Ta-Pei Cheng, Ling-Fong Li: *Gauge theory of elementary particle physics*, Oxford University Press, Oxford.
- [11] C. Amsler *et al.*: *Standard Model and Related Topics*. Physics Letters B **667**, 116 (2008).
- [12] R.F. Streater, A.S. Wightman: *PCT, Spin and Statistics, and All That*, Princeton University Press, Princeton, New Jersey.
- [13] Eric W. Weisstein: *Homotopy*, From MathWorld – A Wolfram Web Resource. <http://mathworld.wolfram.com/Homotopy.html>.
- [14] Jean Zinn-Justin: *Quantum field theory and critical phenomena, third edition*, Oxford University Press, Oxford.
- [15] G. 't Hooft: *Computation of the quantum effects due to a four-dimensional pseudoparticle*, Physical Review D **14**, 3432 (1976).
- [16] Steven Weinberg: *The Quantum Theory of Fields*, Cambridge University Press, Cambridge.
- [17] C. Vafa, E. Witten: *Parity Conservation in Quantum Chromodynamics*. Physical Review Letters **53**, 535 (1984).
- [18] J. Gasser, H. Leutwyler: *Quark masses*. Physics Reports **87**, 77 (1982).

- [19] P. Langacker, H. Pagels: *Light-quark mass spectrum in quantum chromodynamics*. Physical Review D **19**, 2070 (1979).
- [20] C.A. Dominguez: *Zero-Mass Quarks and the $U(1)$ Problem*. Physical Review Letters **41**, 605 (1978).
- [21] Michael Creutz: *Ambiguities in the Up-Quark Mass*. Physical Review Letters **92**, 162003 (2004).
- [22] J. Goldstone: *Field Theories with Superconductor Solutions*. Il Nuovo Cimento **19**, 154 (1961).
- [23] Steven Weinberg: *A New Light Boson?* Physical Review Letters **40**, 223 (1978).
- [24] F. Wilczek: *Problem of Strong P and T invariance in the Presence of Instantons*. Physical Review Letters **40**, 279 (1978).
- [25] S.B. Treiman, F. Wilczek: *Axion emission in decay of excited nuclear states*, Physics Letters B **74**, 381 (1978).
- [26] A. Zehnder, K. Gabathuler, J.-L. Vuilleumier: *Search for axions in specific nuclear gamma-transitions at a power reactor*. Physics Letters B **110**, 419 (1982).
- [27] V.M. Datar, C.V.K. Baba, M.G. Betigeri and P. Singh: *Search for axions in thermal neutron capture by protons*. Physics Letters B **114**, 63 (1982).
- [28] T.W. Donnelly, S.J. Freedman, R.S. Lytel, R.D. Peccei, M. Schwartz: *Do axions exist?*, Physical Review D **18**, 1607 (1978).
- [29] Jihn E. Kim: *Weak-Interaction Singlet and Strong CP Invariance*, Physical Review Letters **43**, 103 (1979).
- [30] M.A. Shifman, A.I. Vainshtein, V.I. Zakharov: *Can confinement ensure natural CP invariance of strong interactions?*, Nuclear Physics B **166**, 493 (1980).
- [31] Michael Dine, Willy Fischler, Mark Srednicki: *A simple solution to the strong CP problem with a harmless axion*, Physics Letters B **104**, 199 (1981).
- [32] A.R. Zhitnitsky, Soviet Journal of Nuclear Physics **31**, 260 (1980).
- [33] Howard Georgi, David B. Kaplan, Lisa Randall: *Manifesting the invisible axion at low energies*, Physics Letters B **169**, 73 (1986).
- [34] Mark Srednicki: *Axion couplings to matter*, Nuclear Physics B **260**, 689 (1985).
- [35] C. Amsler *et al.*: *Mesons*. Physics Letters B **667**, 583 (2008).
- [36] Julian Schwinger: *A Theory of the Fundamental Interactions*, Annals of Physics **2**, 407 (1957).
- [37] J.C. Polkinghorne: *Renormalization of Axial Vector Coupling (II)*, Il Nuovo Cimento **8**, 179,(781) (1958).
- [38] M. Gell-Mann, M. Lévy: *The Axial Vector Current in Beta Decay*, Il Nuovo Cimento **16N.4**, 705 (1960).
- [39] G. 't Hooft: *Symmetry Breaking through Bell-Jackiw Anomalies*, Physical Review Letters **37**, 8 (1976).
- [40] P. Di Vecchia, G. Veneziano: *Chiral dynamics in the large N limit*, Nuclear Physics B **171**, 253 (1980).
- [41] E. Witten: *Large N Chiral Dynamics*, Annals of Physics **128**, 363 (1980).
- [42] R. Dashen: *Some Features of Chiral Symmetry Breaking*, Physical Review D **3**, 1879 (1971).

- [43] A.V. Smilga: *QCD at $\theta \sim \pi$* , Physical Review D **59**, 114021 (1999).
- [44] Michel H.G. Tytgat: *QCD at $\theta \sim \pi$ reexamined: Domain walls and spontaneous CP violation*, Physical Review D **61**, 114009 (2000).
- [45] M. Creutz: *Spontaneous Violation of CP Symmetry in the Strong Interactions*, Physical Review Letters **92**, 201601 (2004).
- [46] Y. Nambu, G. Jona-Lasinio: *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I*, Physical Review **122**, 345 (1961).
- [47] Y. Nambu, G. Jona-Lasinio: *Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. II*, Physical Review **124**, 246 (1961).
- [48] T. Fujihara, T. Inagaki, D. Kimura: *Influence of QED Corrections on the Orientation of Chiral Symmetry Breaking in the NJL Model*, Progress of Theoretical Physics **117**, 139 (2007).
- [49] Daniël Boer, Jorn K. Boomsma: *Spontaneous CP violation in the strong interaction at $\theta = \pi$* , Physical Review D **78**, 054027 (2008).
- [50] J. Gasser, H. Leutwyler: *Chiral Perturbation Theory: Expansions in the mass of the strange quark*, Nuclear Physics B **250**, 465 (1985).
- [51] J. Gasser, H. Leutwyler: *Chiral Perturbation Theory to One Loop*, Annals of Physics **158**, 142 (1984).
- [52] Ana Júlia Mizher, Eduardo S. Fraga: *CP violation and chiral symmetry restoration in the hot linear sigma model in a strong magnetic background*, arXiv:0810.5162 [hep-ph] (2008).

Acknowledgements

First of all, I want to thank my supervisor Daniël Boer for guiding me through the whole project. Especially for keeping up to good spirit and always providing new ideas and telling me that my work was interesting. Whereas I had, more than once, doubts about where this research would end.

I would like to thank Rick Bethlem for being the second reader. In particular for his ability to do all this on a very short time scale.

Then I would like to thank everyone in the theoretical physics group for providing a nice working environment and, of course, also for all fruitful discussions. In particular I would like to thank Tomáš Ježo, Thomas van Dijk, Thijs Stegeman, Jorn Boomsma, Mathieu Blom, Daniël Boekel and Paul McCullough for all the fun at the coffee and lunch breaks. Especially all their endless attempts to beat me in Mario Kart.

Above all, I would like to thank my girlfriend Hannah van Pijkeren and also my parents for their patience and support.